McGraw-Hill's
ENGINEERING COMPANION

- ✔ Contains tables, formulas, charts, diagrams, figures, and worked-out problems

- ✔ Provides reliable, ready-to-apply solutions to hundreds of engineering problems

Ejup N. Ganic
Tyler G. Hicks

McGRAW-HILL'S ENGINEERING COMPANION

Ejup N. Ganić, Sc.D.   Editor in Chief

Professor of Mechanical Engineering, University of Sarajevo
Academy of Sciences and Arts of Bosnia and Herzegovina
Sarajevo, Bosnia and Herzegovina
Formally with the University of Illinois at Chicago

Tyler G. Hicks, P.E.   Editor

International Engineering Associates
Member: American Society of Mechanical Engineers and
Institute of Electrical and Electronics Engineers

Contributions by Myke Predko

McGRAW-HILL, INC.

New York   Chicago   San Francisco   Lisbon   London   Madrid
Mexico City   Milan   New Delhi   San Juan   Seoul
Singapore   Sydney   Toronto

McGraw-Hill

A Division of The McGraw-Hill Companies

McGraw-Hill eBooks are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. For more information, please contact George Hoare, Special Sales, at george_hoare@mcgraw-hill.com or (212) 904-4069.

TERMS OF USE

CONTENTS

Chapter 5. Mechanics of Rigid Bodies 5.1

Chapter 6. Mechanics of Deformable Bodies 6.1

Chapter 7. Thermodynamics 7.1

Chapter 8. Mechanics of Fluids 8.1

Chapter 9. Heat and Mass Transfer 9.1

Chapter 10. Conservation Equations and Dimensionless Groups 10.1

Chapter 11. Topics in Applied Physics 11.1

Chapter 12. Automatic Control 12.1

Chapter 13. Mechanical Engineering 13.1

Mechanical Design Engineering

Chapter 14. Civil Engineering and Hydraulic Engineering 14.1

Chapter 15. Chemical Engineering, Environmental Engineering, and Petroleum and Gas Engineering 15.1

Chapter 16. Electrical Engineering 16.1

Chapter 17. Electronics Engineering 17.1

Index follows Chapter 20

PREFACE

The scores of literature materials available to engineers of today are both vast and impressive. Specialized studies, textbooks, encyclopedias, pocket books of frequently used formulas, and collections of tables, data and conversion charts and, of course, handbooks clutter the bookshelves of libraries and homes. All of these serve a common purpose: to provide information or necessary solutions to problems various or specific. With the new *McGraw-Hill's Engineering Companion,* we make a bold step toward adding yet another work to this well established list and an even bolder claim that this work will soon stand on a new list on its own: A list of one.

How do we find the information we seek, the solution we need, without conducting a large survey through engineering literature? How do we allow engineers to work more on concepts and less on information quests? The scope of this book is defined precisely in answer to these needs. In one volume, essentials of engineering sciences, concise selection of engineering data, and surveys for use in both design and everyday practice are gathered together to help both students and experts navigate their way through challenges ahead. *McGraw-Hill's Engineering Companion* is one volume of manageable size that will soon become the irreplaceable aid for engineers of the twenty-first century.

While preparing the data for the *McGraw-Hill's Engineering Companion* we considered the everyday needs of engineers working on design, product development, applied research, production, installation service, engineering consulting, sales and regulations. We gathered material ready to use for most application-oriented problems and/or theory based calculations. In result, this volume addresses engineering needs both where the user possesses abstract and/or concrete competence. It is in the true sense *the engineer's companion* in every situation.

For students of engineering, the *Companion* illustrates all avenues of engineering world as well as the essence of each topic in the field of engineering. It is an extremely useful tool in completion of student design projects (especially if one holds a summer job!). The *Companion* further allows consulting engineers to quickly position themselves toward the solution of an imposed problem.

Covering every frequently used topic in engineering science, and presenting all engineering data needed in most common applications (including key methods and tools), basic core material of every engineering branch has also been given. This book is a compact but comprehensive source for both the ''old-school'' and the new generation of engineers, working across the traditional discipline lines.

As is the case with all professional activities including engineering, nothing can take the place of common sense and an inquiring mind. This reference is geared to provide the necessary information, in order to save time for professional judgment and creativity. The *Companion* is the place to *turn first* for all practical advice and quick answers through all phases of working on a specific task in engineering.

The art of writing lies in the ability to edit your work. A competent writer is not the one who manages to include all, but the one who chooses well. After thoroughly researching the acclaimed collection of McGraw-Hill publications in the

field of engineering and following that, related literature from publishers worldwide; after utilizing the expertise of many authors, contributors, and editors; we can say that we chose well. Many worked-out examples, a systematic collection of data in the form of tables and diagrams, a collection of clear illustrations, key methods, and tools in every field of engineering are contained in this volume. It should be noted that from the bulk of materials from a variety of sources, only *the tested facts* of engineering and engineering science were included in this book.

McGraw-Hill's Engineering Companion is as easy to use as it is well laid out, with clear chapter division, clear and uniformly outlined contents for each chapter, and in addition, most chapters ending with a nomenclature section that includes a list of all symbols and their SI units.

We did our best to achieve an error-free publication, despite the magnitude of the volume and the amount of information contained within it. Accordingly, we would appreciate being informed of any errors and receiving opinions on deficiencies detected and possible improvements so that these may be amended/included in subsequent printings and future editions.

Ejup N. Ganić
Tyler G. Hicks

ACKNOWLEDGMENTS

Parts of this book were drawn from McGraw-Hill books published over the past several decades. Priority and emphasis was given to the most recent editions of these works. Collections of McGraw-Hill handbooks, specialized technical books, textbooks, pocket books, manuals, encyclopedias, etc., have been thoroughly surveyed for material analysis and selections for this book. This was the main source of literature used. The material drawn is updated and revised.

Proper acknowledgment and/or reference for every source used are given either on the specific page on which the source is used, as an end-of-chapter note, or in the reference list.

Related literature from other publishers worldwide was also surveyed in our attempt to provide the reader with the most competent and complete presentation of material. These works were used for comparison and in order to aid clarity of presentation. All instances where the works in question were used were also properly cited and referenced.

We are grateful to many authors, contributors and editors whose work was used both directly and indirectly within this book. The new generation of engineers will certainly benefit from the contribution of these experts.

We wish to acknowledge specifically the work of Myke Predko, to whom we owe all sections of this volume concerned with electrical and electronic engineering.

We also wish to thank the professional staff at McGraw-Hill, who were involved with the production of the book at various stages of the project, for their outstanding cooperation and continued support.

CHAPTER 1
ENGINEERING UNITS

DIMENSIONS AND UNITS

There are as many *dimensions* as there are kinds of physical quantities. Each new physical quantity gives rise to a new dimension. There can be only one dimension for each physical quantity.

A *unit* is a particular amount of the physical quantity. There are infinite possibilities for choosing a unit of a single physical quantity. All the possible units of the same physical quantity must be related by purely numerical factors.

Derived units are algebraic combinations of base units with some of the combinations being assigned special names and symbols.

SYSTEMS OF UNITS

There are still several systems of common units in use throughout the world. Transition from the others to Système International d'Unités (International System of Units), or SI, will proceed at a rational pace to accommodate the needs of the professions or industries involved and the public. The transition period will be long and complex, and duality of units probably will be demanded for at least a decade after the change is introduced.

1. SI Units. In October 1960, the Eleventh General (International) Conference on Weights and Measures redefined some of the original metric units and expanded the SI system to include other physical and engineering units.

The Metric Conversion Act of 1975 codifies the voluntary conversion of the United States to the SI system. It is expected that in time all units used in the United States will be in SI. For that reason, this chapter includes tables showing SI units, prefixes, and equivalents.

SI consists of seven base units, two supplementary units, a series of derived units consistent with the base and supplementary units, and a series of approved prefixes for the formation of multiples and submultiples of the various units. Multiple and submultiple prefixes in steps of 1000 are recommended.

Table 1.1 shows SI base and supplementary quantities (dimensions) and units. Tables 1.2 and 1.3 and Fig. 1.1 include additional derived units of SI. Table 1.4 shows SI prefixes.

TABLE 1.1 SI Base and Supplementary Quantities and Units

Quantity or "dimension"	SI unit	SI unit symbol ("abbreviation");
Base quantity or "dimension"		
length	meter	m
mass	kilogram	kg
time	second	s
electric current	ampere	A
thermodynamic temperature	kelvin	K
amount of substance	mole	mol
luminous intensity	candela	cd
Supplementary quantity or "dimension"		
plane angle	radian	rad
solid angle	steradian	sr

Source: From Perry, Green, and Maloney.[1]

TABLE 1.2 Derived Units of SI which Have Special Names

Quantity	Unit	Symbol	Formula
frequency (of a periodic phenomenon)	hertz	Hz	$1/s$
force	newton	N	$(kg{\cdot}m)/s^2$
pressure, stress	pascal	Pa	N/m^2
energy, work, quantity of heat	joule	J	$N{\cdot}m$
power, radiant flux	watt	W	J/s
quantity of electricity, electric charge	coulomb	C	$A{\cdot}s$
electric potential, potential difference, electromotive force	volt	V	W/A
capacitance	farad	F	C/V
electric resistance	ohm	Ω	V/A
conductance	siemens	S	A/V
magnetic flux	weber	Wb	$V{\cdot}s$
magnetic-flux density	tesla	T	Wb/m^2
inductance	henry	H	Wb/A
luminous flux	lumen	lm	$cd{\cdot}sr$
illuminance	lux	lx	lm/m^2
activity (of radionuclides)	becquerel	Bq	$1/s$
absorbed dose	gray	Gy	J/kg

Source: From Perry, Green, and Maloney.[1]

2. U.S. Customary System Units. The U.S. Customary System (USCS) is the system of units most commonly used for measures of weight and length in the United States. They are identical for practical purposes with the corresponding English units (see Sec. 4), but the capacity measures differ from those now in use in the British Commonwealth, the U.S. gallon being defined as 231 in^3 and the bushel as 2150.42 in^3, whereas the corresponding British Imperial units are, respectively, 277.42 in^3, and 2219.36 in^3 (1 Imp gal = 1.2 U.S. gal, approx.; 1 Imp bu = 1.03 U.S. bu, approx.). Table 1.5*a* shows USCS units, the corresponding SI

TABLE 1.3 Additional Common Derived Units of SI

Quantity	Unit	Symbol
acceleration	meter per second squared	m/s^2
angular acceleration	radian per second squared	rad/s^2
angular velocity	radian per second	rad/s
area	square meter	m^2
concentration (of amount of substance)	mole per cubic meter	mol/m^3
current density	ampere per square meter	A/m^2
density, mass	kilogram per cubic meter	kg/m^3
electric-charge density	coulomb per cubic meter	C/m^3
electric-field strength	volt per meter	V/m
electric-flux density	coulomb per square meter	C/m^2
energy density	joule per cubic meter	J/m^3
entropy	joule per kelvin	J/K
heat capacity	joule per kelvin	J/K
heat-flux density irradiance	watt per square meter	W/m^2
luminance	candela per square meter	cd/m^2
magnetic-field strength	ampere per meter	A/m
molar energy	joule per mole	J/mol
molar entropy	joule per mole-kelvin	$J/(mol{\cdot}K)$
molar-heat capacity	joule per mole-kelvin	$J/(mol{\cdot}K)$
moment of force	newton-meter	$N{\cdot}m$
permeability	henry per meter	H/m
permittivity	farad per meter	F/m
radiance	watt per square-meter steradian	$W/(m^2{\cdot}sr)$
radian intensity	watt per steradian	W/sr
specific-heat capacity	joule per kilogram-kelvin	$J/(kg{\cdot}K)$
specific energy	joule per kilogram	J/kg
specific entropy	joule per kilogram-kelvin	$J/(kg{\cdot}K)$
specific volume	cubic meter per kilogram	m^3/kg
surface tension	newton per meter	N/m
thermal conductivity	watt per meter-kelvin	$W/(m{\cdot}K)$
velocity	meter per second	m/s
viscosity, dynamic	pascal-second	$Pa{\cdot}s$
viscosity, kinematic	square meter per second	m^2/s
volume	cubic meter	m^3
wave number	1 per meter	$1/m$

Source: From Perry, Green, and Maloney.[1]

units, and the numerical factors used to convert USCS values into SI. Table 1.5*b* defines the abbreviations used.

3. Metric System of Units. In the United States, the *metric* is commonly taken to refer to a system of length and mass units developed in France about 1800. The unit of length was equal to 1/10,000,000 of a quarter meridian (north pole to equator) and named the *meter.* A cube 0.1 meter on a side was the *liter,* the unit of volume. The mass of water filling this cube was the *kilogram,* or standard of mass; i.e., 1 liter of water = 1 kilogram of mass. Metal bars and weights were constructed conforming to these prescriptions for the meter and kilogram. One bar
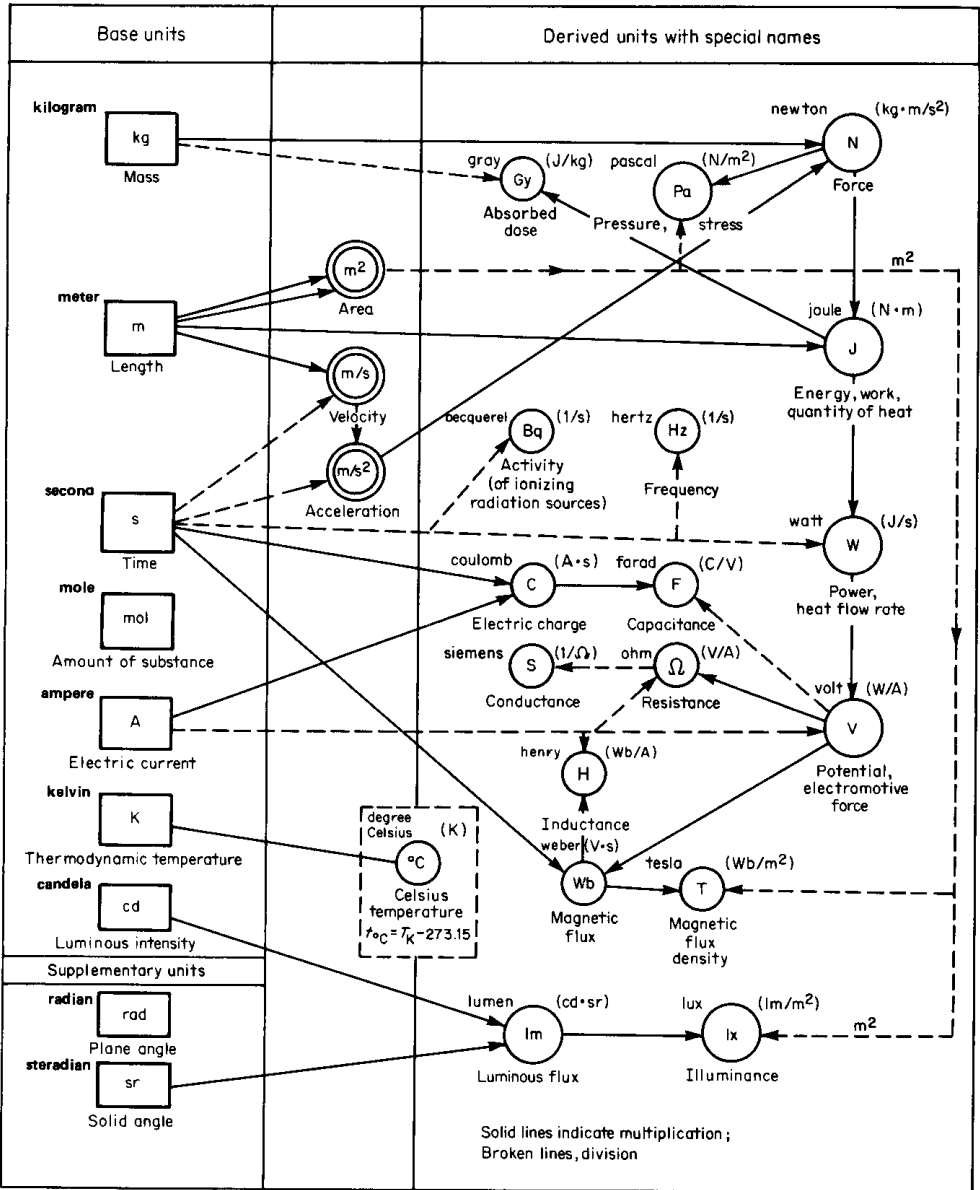
FIGURE 1.1 Graphic relationships of SI units with names. (*From Perry, Green, and Maloney.*[1])

and one weight were selected to be the primary representations. The kilogram and the meter are now defined independently, and the liter, although for many years defined as the volume of a kilogram of water at the temperature of its maximum density, 4°C, and under a pressure of 76 cm of mercury, is now equal to 1 cubic decimeter (1 dim³).

In 1866, the U.S. Congress formally recognized metric units as a legal system, thereby making their use permissible in the United States. In 1893, the Office of

TABLE 1.4   SI Prefixes

Factor	Prefix	Symbol
10^{18}	exa	E
10^{15}	peta	P
10^{12}	tera	T
10^9	giga	G
10^6	mega	M
10^3	kilo	k
10^2	hecto*	h
10	deka*	da
10^{-1}	deci*	d
10^{-2}	centi	c
10^{-3}	milli	m
10^{-6}	micro	μ
10^{-9}	nano	n
10^{-12}	pico	p
10^{-15}	femto	f
10^{-18}	atto	a

 *Generally to be avoided.
 Source: From Rohsenow, Hartnett, and Ganić.[2]

Weights and Measures (now the National Bureau of Standards), by executive order, fixed the values of the U.S. yard and pound in terms of the meter and kilogram, respectively, as 1 yd 3600/3937 m and 1 lb = 0.453 592 4277 kg. By agreement in 1959 among the national standards laboratories of the English-speaking nations, the relations in use now are: 1 yd = 0.9144 m, whence 1 in = 25.4 mm exactly; and 1 lb = 0.453 592 37 kg, or 1 lb = 453.59 g (nearly).

4. English Units.   The base units for the English engineering units are given in Table 1.6 (third column). The unit of force in English units is the pound force (lb_f). However, the use of the pound mass (lb) and pound force in engineering work causes considerable confusion in the proper use of these two fundamentally different units. A clear understanding of the units of mass and force can be gained by examining Newton's second law of motion. With any system of units, a conversion factor g_c must be introduced into the newtonian dynamics equation so that both sides of the equation will have the same units. Thus,

$$F = \frac{ma}{g_c}$$

in which the numerical value and units of g_c depend on the units chosen for mass, force, length, and time.

The units of pound mass and pound force are related by the standard gravitational acceleration, which has a value of 32.174 ft/s². When a 1-lb mass is held at a location on the earth's surface where the gravitational acceleration is 32.174 ft/s², the mass weighs 1 lb_f. With this system of units, the value of g_c is determined as follows:

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†	
		Space,‡time			
Length	naut ml	km		1.852*	E + 00
	mi	km		1.609 344*	E + 00
	yd	m		9.144*	E − 01
	ft	m		3.048*	E − 01
		cm		3.048*	E + 01
	in	mm		2.54*	E + 01
	in	cm		2.54	E + 00
	mil	μm		2.54*	E + 01
Area	mi^2	km^2		2.589 988	E + 00
	acre	ha		4.046 856	E − 01
	ha	m^2		1.000 000*	E + 04
	yd^2	m^2		8.361 274	E − 01
	ft^2	m^2		9.290 304*	E − 02
	in^2	mm^2		6.451 6	E + 02
		cm^2		6.451 6*	E + 00
Volume	cubem	km^3		4.168 182	E + 00
	acre · ft	m^3		1.233 482	E + 03
		ha · m		1.233 482	E − 01
	yd^3	m^3		7.645 549	E − 01
	bbl (42 U.S. gal)	m^3		1.589 873	E − 02
	ft^3	m^3		2.831 685	E − 02
		dm^3	L	2.831 685	E + 01
	U.K. gal	m^3		4.546 092	E − 03
		dm^3	L	4.546 092	E + 00
	U.S. gal	m^3		3.785 412	E − 03
		dm^3	L	3.785 412	E + 00
	U.K. qt	dm^3	L	1.136 523	E + 00
	U.S. qt	dm^3	L	9.463 529	E − 01
	U.S. pt	dm^3	L	4.731 765	E − 01
	U.K. fl oz	cm^3		2.841 307	E + 01
	U.S. fl oz	cm^3		2.957 353	E + 01
	in^3	cm^3		1.638 706	E + 01
Plane angle	rad	rad		1	
	deg (°)	rad		1.745 329	E − 02
	min (′)	rad		2.908 882	E − 04
	sec (″)	rad		4.848 137	E − 06
Solid angle	sr	sr		1	
Time	year	a		1	
	week	d		7.0*	E + 00
	h	s		3.6*	E + 03
		min		6.0*	E + 01
	min	s		6.0*	E + 01
		h		1.666 667	E − 02
	mμs	ns		1	

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units (*Continued*)

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†	
Mass, amount of substance					
Mass	U.K. ton	Mg	t	1.016 047	E + 00
	U.S. ton	Mg	t	9.071 847	E − 01
	U.K. cwt	kg		5.080 234	E + 01
	U.S. cwt	kg		4.535 924	E + 01
	lb	kg		4.535 924	E − 01
	oz (troy)	g		3.110 348	E + 01
	oz (av)	g		2.834 952	E + 01
	gr	mg		6.479 891	E + 01
Amount of substance	lb · mol	kmol		4.535 924	E − 01
	std m³(0°C, 1 atm)	kmol		4.461 58	E − 02
	std ft³(60°F, 1 atm)	kmol		1.195 30	E − 03
Enthalpy, calorific value, heat, entropy, heat capacity					
Calorific value, enthalpy (mass basis)	Btu/lb	MJ/kg		2.326 000	E − 03
		kJ/kg	J/g	2.326 000	E + 00
		kWh/kg		6.461 112	E − 04
	cal/g	kJ/kg	J/g	4.184*	E + 00
	cal/lb	J/kg		9.224 141	E + 00
Caloric value, enthalpy (mole basis)	kcal/(g · mol)	kJ/kmol		4.184*	E + 03
	Btu/(lb · mol)	kJ/kmol		2.326 000	E + 00
Calorific value (volume basis—solids and liquids)	Btu/U.S. gal	MJ/m³	kJ/dm³	2.787 163	E − 01
		kJ/m³		2.787 163	E + 02
		kWh/m³		7.742 119	E − 02
	Btu/U.K. gal	MJ/m³	kJ/dm³	2.320 800	E − 01
		kJ/m³		2.320 800	E + 02
		kWh/m³		6.446 667	E − 02
	Btu/ft³	MJ/m³	kJ/dm³	3.725 895	E − 02
		kJ/m³		3.725 895	E + 01
		kWh/rm³		1.034 971	E − 02
	cal/mL	MJ/m³		4.184*	E + 00
	(ft · lb_f)/U.S. gal	kJ/m³		3.581 692	E − 01
Calorific value (volume basis—gases)	cal/mL	kJ/m³	J/dm³	4.184*	E + 03
	kcal/m³	kJ/m³	J/dm³	4.184*	E + 00
	Btu/ft³	kJ/m³	J/dm³	3.725 895	E + 01
		kWh/m³		1.034 971	E − 02
Specific entropy	Btu/(lb · °R)	kJ/(kg · K)	J/(g · K)	4.186 8*	E + 00
	cal/(g · K)	kJ/(kg · K)	J/(g · K)	4.184*	E + 00
	kcal/(kg · °C)	kJ/(kg · K)	J/(g · K)	4.184*	E + 00
Specific-heat capacity (mass basis)	Btu/(lb · °F)	kJ/(kg · K)	J/(g · K)	4.186 8*	E + 00
	kcal/(kg · °C)	kJ/(kg · K)	J/(g · K)	4.184*	E + 00
Specific heat capacity (mole basis)	Btu/(lb mol · °F)	kJ/(kmol · K)		4.186 8*	E + 00
	cal/(g · mol · °C)	kJ/(kmol · K)		4.184*	E + 00

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units (*Continued*)

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†
Temperature, pressure, vacuum				
Temperature (absolute)	°R	K		5/9
	K	K		1
Temperature (traditional)	°F	°C		5/9(°F − 32)
Temperature (difference)	°F	K, °C		5/9
Pressure	atm (760 mmHg at 0°C or 14.696 lb_f/in^2)	MPa		1.013 250* E − 01
		kPa		1.013 250* E + 02
		bar		1.013 250* E + 00
	bar	MPa		1.0* E − 01
		kPa		1.0* E + 02
	mmHg (0°C) = torr	MPa		6.894 757 E − 03
		kPa		6.894 757 E + 00
		bar		6.894 757 E − 02
	μmHg (0°C)	kPa		3.376 85 E + 00
	μ bar	kPa		2.488 4 E − 01
	mmHg = torr (0°C)	kPa		1.333 224 E − 01
	cmH_2O (4°C)	kPa		9.806 38 E − 02
	lb_f/ft^2 (psf)	kPa		4.788 026 E − 02
	mHg (0°C)	Pa		1.333 224 E − 01
	bar	Pa		1.0* E − 01
	dyn/cm^2	Pa		1.0* E − 01
Vacuum, draft	inHg (60°F)	kPa		3.376 85 E + 00
	inH_2O (39.2°F)	kPa		2.490 82 E − 01
	inH_2O (60°F)	kPa		2.488 4 E − 01
	mmHg (0°C) = torr	kPa		1.333 224 E − 01
	cmH_2O (4°C)	kPa		9.806 38 E − 02
Liquid head	ft	m		3.048* E − 01
	in	mm		2.54* E + 01
		cm		2.54* E + 00
Pressure drop/length	$lb_f/in^2/ft$	kPa/m		2.262 059 E + 01
Density, specific volume, concentration, dosage				
Density	lb/ft^3	kg/m^3		1.601 846 E + 01
		g/m^3		1.601 846 E + 04
	lb/U.S. gal	kg/m^3		1.198 264 E + 02
		g/cm^3		1.198 264 E − 01
	lb/U.K. gal	kg/m^3		9.977 633 E + 01

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units (*Continued*)

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†	
Density, specific volume, concentration, dosage (Continued)					
	lb/ft³	kg/m³		1.601 846	E + 01
		g/cm³		1.601 846	E − 02
	g/cm³	kg/m³		1.0*	E + 03
	lb/ft³	kg/m³		1.601 846	E + 01
Specific volume	ft³/lb	m³/kg		6.242 796	E − 02
		m³/g		6.242 796	E − 05
	ft³/lb	dm³/kg		6.242 796	E + 01
	U.K. gal/lb	dm³/kg	cm³/g	1.002 242	E + 01
	U.S. gal/lb	dm³/kg	cm³/g	8.345 404	E + 00
Specific volume (mole basis)	L/(g · mol)	m³/kmol		1	
	ft³/(lb · mol)	m³/kmol		6.242 796	E − 02
Concentration (mass/volume)	lb/bbl	kg/m³	g/dm³	2.853 010	E + 00
	g/U.S. gal	kg/m³		2.641 720	E − 01
	g/U.K. gal	kg/m³	g/L	2.199 692	E − 01
	lb/1000 U.S. gal	g/m³	mg/dm³	1.198 264	E + 02
	lb/1000 U.K. gal	g/m³	mg/dm³	9.977 633	E + 01
	gr/U.S. gal	g/m³	mg/dm³	1.711 806	E + 01
	gr/ft³	mg/m³		2.288 351	E + 03
	lb/1000 bbl	g/m³	mg/dm³	2.853 010	E + 00
	mg/U.S. gal	g/m³	mg/dm³	2.641 720	E − 01
	gr/100 ft³	mg/m³		2.288 351	E + 01
Concentration (mole/volume)	(lb · mol)/ U.S. gal	kmol/m³		1.198 264	E + 02
	(lb · mol)/ U.K. gal	kmol/m³		9.977 644	E + 01
	(lb · mol)/ft³	kmol/m³		1.601 846	E + 01
Flow rate					
Flow rate (mass basis)	lb/s	kg/s		4.535 924	E − 01
	lb/mm	kg/s		7.559 873	E − 03
	lb/h	kg/s		1.259 979	E − 04
Flow rate (volume basis)	bbl/day	m³/d		1.589 873	E − 01
		L/s		1.840 131	E − 03
	ft³/day	m³/d		2.831 685	E − 02
	bbl/h	L/s		3.277 413	E − 04
		m³/s		4.416 314	E − 05
	ft³/h	L/s		4.416 314	E − 02
		m³/s		7.865 791	E − 06
	ft³/min	dm³/s	L/s	6.309 020	E − 02
	ft³/s	dm³/s	L/s	4.719 474	E − 01
		dm³/s	L/s	2.831 685	E + 01
Flow rate (mole basis)	(lb · mol)/s	kmol/s		4.535 924	E − 01
	(lb · mol)/h	kmol/s		1.259 979	E − 04

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units (*Continued*)

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†	
Flow rate (*Continued*)					
Flow rate/area (mass basis)	$lb/(s \cdot ft^2)$	$kg/(s \cdot in^2)$		4.882 428	E + 00
	$lb/(h \cdot ft^2)$	$kg/(s \cdot m^2)$		1.356 230	E − 03
Flow rate/area (volume basis)	$ft^3/(s \cdot ft^2)$	m/s	$m^3/(s \cdot m^2)$	3.0 048*	E − 01
	$ft^3/(min \cdot ft^2)$	m/s	$m^3/(s \cdot m^2)$	5.08*	E − 03
	U.K. $gal/(h \cdot ft^2)$	m/s	$m^3/(s \cdot m^2)$	1.359 270	E − 05
	U.S. $gal/(h \cdot ft^2)$	m/s	$m^3/(s \cdot m^2)$	1.131 829	E − 05
Energy, work, power					
Energy, work	therm	MJ		1.055 056	E + 02
		kJ		1.055 056	E + 05
		kWh		2.930 711	E + 01
	$hp \cdot h$	MJ		2.684 520	E + 00
		kJ		2.684 520	E + 03
		kWh		7.456 999	E − 01
	$ch \cdot h$ or $CV \cdot h$	MJ		2.647 780	E + 00
		kJ		2.647 780	E + 03
		kWh		7.354 999	E − 01
	kWh	MJ		3.6*	E + 00
		kJ		3.6*	E + 03
	Chu	kJ		1.899 101	E + 00
		kWh		5.275 280	E − 04
	Btu	kJ		1.055 056	E + 00
		kWh		2.930 711	E − 04
	kcal	kJ		4.184*	E + 00
	cal	kJ		4.184*	E − 03
	$ft \cdot lb_f$	kJ		1.355 818	E − 03
	$lb_f \cdot ft$	kJ		1.355 818	E − 03
	J	kJ		1.0*	E − 03
	$(lb_f \cdot ft^2)/s^2$	kJ		4.214 011	E − 05
	erg	J		1.0*	E − 07
Impact energy	$kg_f \cdot m$	J		9.806 650*	E + 00
	$lb_f \cdot ft$	J		1.355 818	E + 00
Surface energy	erg/cm^2	mJ/m^2		1.0*	E + 00
Specific-impact energy	$(kg_f \cdot m)/cm^2$	J/cm^2		9.806 650*	E − 02
	$(lb_f \cdot ft)/in^2$	J/cm^2		2.101 522	E − 03
Power	million Btu/h	MW		2.930 711	E − 01
	ton of refrigeration	kW		3.516 853	E + 00
	Btu/s	kW		1.055 056	E + 00
	hydraulic horsepower—hhp	kW		7.460 43	E − 01
	hp (electric)	kW		7.46*	E − 01
	hp $[(550 ft \cdot lb_f)/s]$	kW		7.456 999	E − 01

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units (*Continued*)

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†	
		Energy, work, power (*Continued*)			
	Btu/min	kW		1.758 427	E − 02
	$(ft \cdot lb_f)/s$	kW		1.355 818	E − 03
	kcal/h	W		1.162 222	E + 00
	Btu/h	W		2.930 711	E − 01
	$(ft \cdot lb_f)/min$	W		2.259 697	E − 02
Power/area	$Btu/(s \cdot ft^2)$	kW/m^2		1.135 653	E + 01
	$cal/(h \cdot cm^2)$	kW/m^2		1.162 222	E − 02
	$Btu/(h \cdot ft^2)$	kW/m^2		3.154 591	E − 03
Heat-release rate, mixing power	hp/ft^3	kW/m^3		2.633 414	E + 01
	$cal/(h \cdot cm^3)$	kW/m^3		1.162 222	E + 00
	$Btu/(s \cdot ft^3)$	kW/m^3		3.725 895	E + 01
	$Btu/(h \cdot ft^3)$	kW/m^3		1.034 971	E − 02
Cooling duty (machinery)	$Btu/(bhp \cdot h)$	W/kW		3.930 148	E − 01
Specific fuel consumption (mass basis)	$1b/(hp \cdot h)$	mg/J	kg/MJ	1.689 659	E − 01
		kg/kWh		6.082 774	E − 01
Specific fuel consumption (volume basis)	m^3/kWh	dm^3/MJ	mm^3/J	2.777 778	E + 02
	U.S. $gal/(hp \cdot h)$	dm^3/MJ	mm^3/J	1.410 089	E + 00
Fuel consumption	U.K. gal/mi	$dm^3/100$ km	L/100 km	2.824 807	E + 02
	U.S. gal/mi	$dm^3/100$ km	L/100 km	2.352 146	E + 02
	mi/U.S. gal	km/dm^3	km/L	4.251 437	E − 01
	mi/U.K. gal	km/dm^3	km/L	3.540 064	E − 01
Velocity (linear), speed	knot	km/h		1.852*	E + 00
	mi/h	km/h		1.609 344*	E + 00
	ft/s	m/s		3.048*	E − 01
		cm/s		3.048*	E + 01
	ft/min	m/s		5.08*	E − 03
	ft/h	mm/s		8.466 667	E − 02
	ft/day	mm/s		3.527 778	E − 03
		m/d		3.048*	E − 01
	in/s	mm/s		2.54*	E + 01
	in/min	mm/s		4.233 333	E − 01
Corrosion rate	in/year (ipy)	mm/a		2.54*	E + 01
	mil/year	mm/a		2.54*	E − 02
Rotational frequency	r/min	r/s		1.666 667	E − 02
		rad/s		1.047 198	E − 01
Acceleration (linear)	ft/s^2	m/s^2		3.048*	E − 01
		cm/s^2		3.048*	E + 01

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units (*Continued*)

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†	
		Energy, work, power (*Continued*)			
Acceleration (rotational)	rpm/s	rad/s^2		1.047 198	E − 01
Momentum	(lb · ft)/s	(kg · m)/s		1.382 550	E − 01
Force	U.K. ton$_f$	kN		9.964 016	E + 00
	U.S. ton$_f$	kN		8.896 443	E + 00
	kg$_f$ (kp)	N		9.806 650*	E + 00
	lb$_f$	N		4.448 222	E + 00
	dyn	mN		1.0	E − 02
Bending moment, torque	U.S. ton$_f$ · ft	kN · m		2.711 636	E + 00
	kg$_f$ · m	N · m		9.806 650*	E + 00
	lb$_1$ft	N · m		1.355 818	E + 00
	lb$_f$ · in	N · m		1.129 848	E − 01
Bending moment/ length	(lb$_f$ · ft)/in	(N · m)/m		5.337 866	E + 01
	(lb$_f$ · in)/in	(N · m)/m		4.448 222	E + 00
Moment of inertia	lb · ft^2	kg · m^2		4.214 011	E − 02
Stress	U.S. ton$_f$ /in^2	MPa	N/mm^2	1.378 951	E + 01
	kg$_f$/mm^2	MPa	N/mm^2	9.806 650*	E + 00
	U.S. ton$_f$ /ft^2	MPa	N/mm^2	9.576 052	E − 02
	lb$_f$/in^2 (psi)	MPa	N/mm^2	6.894 757	E − 03
	lb$_f$/ft^2 (psf)	kPa		4.788 026	E − 02
	dyn/cm^2	Pa		1.0*	E − 01
Mass/length	lb/ft	kg/m		1.488 164	E + 00
Mass area structural loading, bearing capacity (mass basis)	U.S. ton/ft^2	Mg/m^2		9.764 855	E + 00
	lb/ft^2	kg/m^2		4.882 428	E + 00
		Miscellaneous transport properties			
Diffusivity	ft^2/s	m^2/s		9.290 304*	E − 02
	m^2/s	mm^2/s		1.0*	E + 06
	ft^2/h	m^2/s		2.580 64*	E − 05
Thermal resistance	(°C · m^2 · h)/ kcal	(K · m^2)/kW		8.604 208	E + 02
	(°F · ft^2 · h)/Btu	(K · m^2)/kW		1.761 102	E + 02
Heat flux	Btu/(h · ft^2)	kW/m^2		3.154 591	E − 03
Thermal conductivity	(cal · cm)/ (s · cm^2 · °C)	W/(m · K)		4.184*	E + 02
	(Btu · ft)/ (h · ft^2 · °F)	W/(m · K)		1.730 735	E + 00
		(kJ · m)/ (h · m^2 · K)		6.230 646	E + 00
	(kcal · m)/ (h · m^2 · °C)	W/(m · K)		1.6 162 22	E + 00

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units (*Continued*)

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†	
	Miscellaneous transport properties (*Continued*)				
	(Btu · in)/ (h · ft² · °F)	W/(m · K)		1.442 270	E − 01
	(cal · cm)/ (h · cm² · °C)	W/(m · K)		1.162 222	E − 01
Heat-transfer coefficient	cal/(h · cm² · °C)	kW/(m² · K)		1.162 222	E − 02
	Btu/(h · ft² · °F)	kW/(m² · K)		5.678 263	E − 03
		kJ/(h · m² · K)		2.044 175	E + 01
	Btu/(h · ft² · °R)	kW/(m² · K)		5.678 263	E − 03
	kcal/(h · m² · °C)	kW/(m² · K)		1.162 222	E − 03
Volumetric heat-transfer coefficient	Btu/(s · ft³ · °F)	kW/(m³ · K)		6.706 611	E + 01
	Btu/(h · ft³ · °F)	kW/(m³ · K)		1.862 947	E − 02
Surface tension	dyn/cm	N/m		1.	E + 03
Viscosity (dynamic)	(lb$_f$ · s)/in²	Pa · s	(N · s)/m²	6.894 757	E + 03
	(lb$_f$ · s)/ft²	Pa · s	(N · s)/m²	4.788 026	E + 01
	(kg$_f$ · s)/m²	Pa · s	(N · s)/m²	9.806 650*	E + 00
	lb/(ft · s)	Pa · s	(N · s)/m²	1.488 164	E + 00
	(dyn · s)/cm²	Pa · s	(N · s)/m²	1.0*	E − 01
	cP	Pa · s	(N · s)/m²	1.0*	E − 03
	lb/(ft · h)	Pa · s	(N · s)/m²	4.133 789	E − 04
Viscosity (kinematic)	ft²/s	m²/s		9.290 304*	E 02
	in²/s	mm²/s		6.451 6*	E + 02
	m²/h	mm²/s		2.777 778	E + 02
	ft²/h	m²/s		2.580 64*	E − 05
	cSt	mm²/s		1	
Permeability	darcy	μm²		9.869 233	E − 01
	millidarcy	μm²		9.869 233	E − 04
Thermal flux	Btu/(h · ft²)	W/m²		3.152	E + 00
	Btu/(s · ft²)	W/m²		1.135	E + 04
	cal/(s · cm²)	W/m²		4.184	E + 04
Mass-transfer coefficient	(lb · mol)/[h · ft² (lb · mol/ft³)]	m/s		8.467	E − 05
	(g · mol)/[s · m² (g · mol/L)]	m/s		1.0	E + 01
	Electricity, magnetism				
Admittance	S	S		1	
Capacitance	μF	μF		1	
Charge density	C/mm³	C/mm³		1	
Conductance	S	S		1	
	Ω(mho)	S		1	

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units (*Continued*)

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†
	Electricity, magnetism (*Continued*)			
Conductivity	S/m	S/m		1
	Ω/m	S/m		1
	mΩ/m	mS/m		1
Current density	A/mm²	A/mm²		1
Displacement	C/cm²	C/cm²		1
Electric charge	C	C		1
Electric current	A	A		1
Electric-dipole moment	C · m	C · m		1
Electric-field strength	V/m	V/m		1
Electric flux	C	C		1
Electric polarization	C/cm²	C/cm²		1
Electric potential	V	V		1
	mV	mV		1
Electromagnetic moment	A · m²	A · m²		1
Electromotive force	V	V		1
Flux of displacement	C	C		1
Frequency	cycles/s	Hz		1
Impedance	Ω	Ω		1
Linear-current density	A/mm	A/mm		1
Magnetic-dipole moment	Wb · m	Wb · m		1
Magnetic-field strength	A/mm	A/mm		1
	Oe	A/m		7.957 747 E + 01
	gamma	A/m		7.957 747 E − 04
Magnetic flux	mWb	mWb		1
Magnetic-flux density	mT	mT		1
	G	T		1.0* E − 04
	gamma	nT		1
Magnetic induction	mT	mT		1
Magnetic moment	A · m²	A · m²		1
Magnetic polarization	mT	mT		1

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units
(*Continued*)

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†
		Electricity, magnetism (*Continued*)		
Magnetic potential difference	A	A		1
Magnetic-vector potential	Wb/mm	Wb/mm		1
Magnetization	A/mm	A/mm		1
Modulus of admittance	S	S		1
Modulus of impedance	Ω	Ω		1
Mutual inductance	H	H		1
Permeability	μH/m	μH/m		1
Permeance	H	H		1
Permittivity	μF/m	μF/m		1
Potential difference	V	V		1
Quantity of eletricity	C	C		1
Reactance	Ω	Ω		1
Reluctance	H^{-1}	H^{-1}		1
Resistance	Ω	Ω		1
Resistivity	$\Omega \cdot$ cm $\Omega \cdot$ m	$\Omega \cdot$ cm $\Omega \cdot$ m		1 1
Self-inductance	mH	mH		1
Surface density of change	mC/m^2	mC/m^2		1
Susceptance	S	S		1
Volume density of charge	C/mm^3	C/mm^3		1
		Acoustics, light, radiation		
Absorbed dose	rad	Gy		1.0* E $-$ 02
Acoustical energy	J	J		1
Acoustical intensity	W/cm^2	W/m^2		1.0* E $+$ 04
Acoustical power	W	W		1
Sound pressure	N/m^2	N/m^2		1.0*
Illuminance	fc	lx		1.076 391 E $+$ 01

TABLE 1.5a Conversion Factors: U.S. Customary and Commonly Used Units to SI Units (*Continued*)

Quantity	Customary or commonly used unit	SI unit	Alternative SI unit	Conversion factor*; multiply customary unit by factor to obtain SI unit†
Acoustics, light, radiation (Continued)				
Illumination	fc	lx		1.076 391 E + 01
Irradiance	W/m^2	W/m^2		1
Light exposure	fc · s	lx · s		1.076 391 E + 01
Luminance	cd/m^2	cd/m^2		1
Luminous efficacy	lm/W	lm/W		1
Luminous exitance	lm/m^2	lm/m^2		1
Luminous flux	lm	lm		1
Luminous intensity	cd	cd		1
Radiance	$W/m^2 \cdot sr$	$W/m^2 \cdot sr$		1
Radiant energy	J	J		1
Radiant flux	W	W		1
Radiant intensity	W/sr	W/sr		1
Radiant power	W	W		1
Wavelength	Å	nm		1.0* E − 01
Capture unit	$10^{-3}cm^{-1}$	m^{-1}		1.0* E + 01
			$10^{-3}cm^{-1}$	1
	m^{-1}	m^{-1}		1
Radioactivity	Ci	Bq		3.7* E + 10

　*An asterisk indicates that the conversion factor is exact.
　†Or multiply SI unit by reciprocal value of factor to obtain USCS unit.
　‡Conversion factors for length, area, and volume are based on the international foot. The international foot is longer by 2 parts in 1 million than the U.S. Survey foot (land-measurement use).
　Source: From Perry, Green, and Maloney.[1]

$$F = \frac{ma}{g_c}$$

$$1 \ lb_f = \frac{1 \ lb \times 32.174 \ ft/s^2}{g_c}$$

whence

$$g_c = 32.174 \ lb \cdot ft/(lb_f \cdot s^2)$$

TABLE 1.5b Unit Symbols Used in Table 1.5a

Unit symbol	Name	Unit symbol	Name
A	ampere	lm	lumen
a	annum (year)	lx	lux
Bq	becquerel	m	meter
C	coulomb	℧	mho
cd	candela	min	minute
Ci	curie	′	minute
d	day	N	newton
°C	degree Celsius	naut mi	U.S. nautical mile
°	degree	Oe	oersted
dyn	dyne	Ω	ohm
F	farad	P	poise
fc	footcandle	Pa	pascal
G	gauss	rad	radian
g	gram	r	revolution
gr	grain	S	siemens
Gy	gray	s	second
H	henry	″	second
h	hour	sr	steradian
ha	hectare	St	stokes
Hz	hertz	T	tesla
J	Joule	ton_f	ton force
K	kelvin	t	tonne
L, *l*, l	liter	V	volt
lb	pound mass	W	watt
lb_f	pound force	Wb	weber

Thus, g_c is merely a conversion factor and it should not be confused with the gravitational acceleration g. The numerical value of g_c is a constant depending only on the system of units involved, and not on the value of the gravitational acceleration at a particular location. Values of g_c corresponding to different systems of units found in engineering literature are given in Table 1.6.

5. Derived SI Units.   Note that the SI units comprise a rigorously coherent form of the metric system; i.e., all remaining units may be derived from the base units using formulas which do not involve any numerical factors. For example, the unit of force is the newton (N). One newton is one kilogram metre per second squared. The newton is defined as the force which, when applied to a mass of one kilogram, gives it an acceleration of one metre per second squared. Thus, *force $F = m \cdot a$* where m is the mass in kilograms and a is the acceleration in meters per second squared. *Gravitational force,* or *weight,* is $m \cdot g$, where $g = 9.81$ m/s².

   The unit of *charge* is the coulomb (C) where one coulomb is one ampere second. (1 coulomb = 6.24×10^{18} electrons). The coulomb is defined as the quantity of electricity which flows past a given point in an electric circuit when a current of one ampere is maintained for one second. Thus, *charge* in coulombs $Q = i \cdot t$ where i is the current in amperes and t is time in seconds.

   Therefore, derived SI units use combinations of basic units and there are many of them. Further simple examples are:

TABLE 1.6   Conversion Factor g_c for the Common Unit Systems

Quantity	Si	English engineering*	cgs‡	Metric engineering
Mass	kilogram, kg	pound mass, lb	gram, g	kilogram mass, kg
Length	meter, m	foot, ft	centimeter, cm	meter, m
Time	second, s	second, s, or hour, h	second, s	second, s
Force	newton, N	pound force, lb_f	dyne, dyn	kilogram force, kg_f
g_c	1 $kg \cdot m/(N \cdot s^2)$‡	32.174 $lb \cdot ft/(1b_f \cdot s^2)$ or 4.1698×10^2 $lb \cdot ft/(lb_f \cdot h^2)$	1 $g \cdot cm/(dyn \cdot s^2)$	9.80665 $kg \cdot m/(kg_f \cdot s^2)$

* In this system of units the temperature is given in degrees Fahrenheit (°F).
† Centimeter-gram-second: this system of units has been used mostly in scientific work.
‡ Since $1 \text{ kg} \cdot m/s^2 = 1$ N, then $g_c = 1$ in the SI system of units.
Source: From Rohsenow, Hartnett, and Ganić.[2]

velocity—meters per second (m/s)

acceleration—meters per second squared (m/s²)

pressure—Newton per meter squared (N/m²)

The unit of pressure (N/m²) is often referred to as the *pascal* (Pa).

In the SI system, there is one unit of energy, whether the energy is thermal, mechanical, or electrical: the joule (J), ($1 \text{ J} = 1 \text{ N} \cdot m$). The unit for energy rate, or power, is the J/s, where one joule per second is equivalent to one watt (W) ($1 \text{ J/s} = 1$ W).

In the English system of units, it is necessary to relate thermal and mechanical energy via the mechanical equivalent of heat, J_c. Thus

$$J_c \times \text{thermal energy} = \text{mechanical energy}$$

The unit of heat in the English system is the British thermal unit (Btu). When the unit of mechanical energy is the pound-force-foot ($lb_f \cdot ft$), then

$$J_c = 778.16 \text{ } lb_f \cdot ft/\text{Btu}$$

as $1 \text{ Btu} = 778.16 \text{ } lb_f \cdot ft$. Happily, in the SI system the units of heat and work are identical and J_c is unity.

6. SI Learning and Usage.   The technical and scientific community throughout the world accepts SI units for use in both applied and theoretical calculation. With such widespread acceptance, every engineer must become proficient in the use of this system if he or she is to remain up to date. For this reason, most calculation procedures in this handbook are given in both SI and USCS. This will help all engineers become proficient in using both systems. However, in some cases results and tables are given in one system, mostly to save space, and conversion factors are printed at the end of such results (or tables) for the reader's convenience.

Engineers accustomed to working in USCS are often timid about using SI. There are really no sound reasons for these fears. SI is a logical, easily understood, and

readily manipulated group of units. Most engineers grow to prefer SI, once they become familiar with it and overcome their fears.

Overseas engineers who must work in USCS because they have a job requiring its usage will find the dual-unit presentation of calculation procedures most helpful. Knowing SI, they can easily convert to USGS.

An efficient way for the USCS-conversant engineer to learn SI follows these steps:

1. List units of measurement commonly used in one's daily work.
2. Insert, opposite each USGS unit, the usual SI unit used; Table 1.5 shows a variety of commonly used quantities and the corresponding SI units.
3. Find, from a table of conversion factors, such as Table 1.5, the value to use to convert the USGS unit to SI, and insert it in the list. (Most engineers prefer a conversion factor that can be used as a multiplier of the USGS unit to give the SI unit.)
4. Apply the conversion factor whenever the opportunity arises. Think in terms of SI when an USGS unit is encountered.
5. Recognize—here and now—that the most difficult aspect of SI is becoming comfortable with the names and magnitudes of the units. Numerical conversion is simple once a conversion table has been set up. So think *pascal* whenever pounds per square inch pressure are encountered, *newton* whenever a force in pounds is being dealt with, etc.

CONVERSION FACTORS

Conversion factors between SI and USGS units are given in Table 1.5. Note that E indicates an exponent, as in scientific notation, followed by a positive or negative number representing the power of 10 by which the given conversion factor is to be multiplied before use. Thus, for the square feet conversion factor, $9.290\ 304 \times 1/100 = 0.092\ 903\ 04$, the factor to be used to convert square feet to square meters. For a positive exponent, as in converting British thermal units per cubic foot to kilojoules per cubic meter, $3.725\ 895 \times 10 = 37.258\ 95$.

Where a conversion factor cannot be found, simply use the dimensional substitution. Thus, to convert pounds per cubic inch to kilograms per cubic meter, find $1\ \text{lb} = 0.453\ 592\ 4\ \text{kg}$ and $1\ \text{in}^3 = 0.000\ 016\ 387\ 06\ \text{m}^3$. Then, $1\ \text{lb/in}^3 = 0.453\ 592\ 4\ \text{kg}/0.000\ 016\ 387\ 06\ \text{m}^3 = 2.767\ 990\ \text{E} + 04$.

SELECTED PHYSICAL CONSTANTS

A list of selected physical constants is given in Table 1.7.

DIMENSIONAL ANALYSIS

Dimensional analysis is the mathematics of dimensions and quantities and provides procedural techniques whereby the variables that are assumed to be significant in

TABLE 1.7 Fundamental Physical Constants

1 sec. = 1.00273791 sidereal seconds	sec. = mean solar second
g_0 = 9.80665 m./sec.^{2}	Definition: g_0 = standard gravity
1 liter = 0.001 cu. m.	
1 atm. = 101,325 newtons/sq. m.	Definition: atm. = standard atmosphere
1 mm. Hg (pressure = $1/760$) atm.	mm. Hg (pressure) = standard millimeter mercury
$\quad$ = 133.3224 newtons/sq. m.	
1 int. ohm = 1.000495 ± 0.000015 abs. ohm	int. = international; abs. = absolute
1 int. amp. = 0.999835 ± 0.000025 abs. amp.	amp. = ampere
1 int. coul. = 0.999835 ± 0.000025 abs. coul.	coul. = coulomb
1 int. volt = 1.000330 ± 0.000029 abs. volt	
1 int. watt = 1.000165 ± 0.000052 abs. watt	
1 int. joule = 1.000165 ± 0.000052 abs. joule	
$T_{0°C}$ = 273.150 ± 0.010 K.	Absolute temperature of the ice point, 0°C.
R = 8.31439 ± 0.00034 abs. joule/deg. mole	R = gas constant per mole
$\quad$ = 1.98719 ± 0.00013 cal./deg. mole	
$\quad$ = 82.0567 ± 0.0034 cu. cm. atm./deg. mole	
$\quad$ = 0.0820567 ± 0.0000034 liter atm./deg. mole	
ln 10 = 2.302585	ln = natural logarithm (base e)
R ln 10 = 19.14460 ± 0.00078 abs. joule/deg. mole	
$\quad$ = 4.57567 ± 0.00030 cal./deg. mole	
N = (6.02283 ± 0.0022) × 10^{23}/mole	N = Avogadro number
h = (6.6242 ± 0.0044) × 10^{-34} joule sec.	h = Planck constant
c = (2.99776 ± 0.00008) × 10^8 m./sec.	c = velocity of light
$(h^2/8\pi^2 k)$ = (4.0258 ± 0.0037) × 10^{-39} g. sq. cm. deg.	Constant in rotational partition function of gases
$(h/8\pi^2 c)$ = (2.7986 ± 0.0018) × 10^{-39} g. cm.	Constant relating wave number and moment of inertia
$Z = Nhc$ = 11.9600 ± 0.0036 abs. joule cm./mole	Z = constant relating wave number and energy per mole
$\quad$ = 2.85851 ± 0.0009 cal. cm./mole	
$(Z/R) = (hc/k) = c_2$ = 1.43847 ± 0.00045 cm. deg.	C_2 = second radiation constant
$\mathcal{F}$ = 96,501.2 ± 10.0 int. coul./g.-equiv. or int. joule/int. volt g.-equiv.,	$\mathcal{F}$ = Faraday constant
$\quad$ = 96,485.3 ± 10.0 abs. cou./g.-equiv. Or abs. joule/abs. volt g.-equiv.	
$\quad$ = 23,068.1 ± 2.4 cal./int. volt g.-equiv.	
$\quad$ = 23,060.5 ± 2.4 cal./abs. volt g.-equiv.	
e = (1.60199 ± 0.00060) × 10^{-19} abs. coul.	e = electronic charge
$\quad$ = (1.60199 ± 0.00060) × 10^{-20} abs. e.m.u.	
$\quad$ = (4.80239 ± 0.00180) × 10^{-10} abs. e.s.u.	
1 int. electron-volt/molecule = 96,501.2 ± 10 int. joule/mole	
$\quad$ = 23,068.1 ± 2.4 cal./mole	

1 abs. electron-volt/molecule = 96,485.3 ± 10.abs. joule/mole
$\quad$ = 23,060.5 ± 2.4 cal./mole

Constant relating wave number and energy per molecule

1 int. electron-volt = (1.60252 ± 0.00060) × 10^{-12} erg
1 abs. electron-volt = (1.60199 ± 0.00060) × 10^{-12} erg
hc = (1.23916 ± 0.00032) × 10^{-4} int. electron-volt cm.
$\quad$ = (1.23957) ± 0.00032) × 10^{-4} abs. electron-volt cm.
k = (8.61442 ± 0.00100) × 10^{-5} int. electron-volt/deg.
$\quad$ = (8.61727 ± 0.00100) × 10^{-5} abs. electron-volt/deg.
$\quad$ = (R/N) = (1.38048 ± 0.00050) × 10^{-23} joule/deg.

k = Boltzmann constant

1 I.T. cal. = ($\frac{1}{860}$) = 0.00116279 int. watt-hr.
$\quad$ = 4.18605 int. joule
$\quad$ = 4.18674 abs. joule
$\quad$ = 1.000654 cal.

Definition of I.T. cal.: I.T. = International steam tables

1 cal. = 4.1840 abs. joule
$\quad$ = 4.1833 int. joule
$\quad$ = 41.2929 ± 0.0020 cu. cm. atm.
$\quad$ = 0.0412929 ± 0.0000020 liter atm.

cal. = thermochemical calorie
Definition: cal. = thermochemical calorie

1 I.T. cal./g. = 1.8 B.t.u./lb.
1 B.t.u. = 251.996 I.T. cal.
$\quad$ = 0.293018 int. watt-hr.
$\quad$ = 1054.866 int. joule
$\quad$ = 1055.040 abs. joule
$\quad$ = 252.161 cal.

Definition of B.t.u.: B.t.u. = I.T. British Thermal Unit

1 horsepower = 550 ft.-lb. (wt.)/sec.
$\quad$ = 745.578 int. watt
$\quad$ = 745.70 abs. watt

cal. = thermochemical calorie
Definition of horsepower (mechanical): lb. (wt.) = weight* of 1 lb. At standard gravity

1 in. = (1/0.39337) = 2.54 cm.
1 ft = 0.304800610 m.
1 lb. = 453.5924277 g.
1 gal. = 231 cu. in.
$\quad$ = 0.133680555 cu. ft.
$\quad$ = 3.785412 × 10^{-3} cu. m.
$\quad$ = 3.785412 liter

Definition of in.: in. = U.S. inch
ft = U.S. foot (1 ft. = 12 in.)
Definition; lb. = avoirdupois pound
Definition; gal. = U.S. gallon

*lb (wt.) = lb$_f$
Source: From Perry, Green, and Maloney.[1]

a problem can be formed into dimensionless parameters, the number of parameters being less than the number of variables. This is a great advantage, because fewer experimental runs are then required to establish a relationship between the parameter than between the variables. While the user is not presumed to have any knowledge of the fundamental physical equations, the more knowledgeable the user, the better the results. If any significant variable or variables are omitted, the relationship obtained from dimensional analysis will not apply to the physical problem. On the other hand, inclusion of all possible variables will result in losing the principal advantage of dimensional analysis, i.e., the reduction of the amount of experimental data required to establish a relationship. Formal methods of dimensional analysis are given in Chap. 10.

REFERENCES*

1. R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Perry's Chemical Engineers Handbook,* 6th ed., McGraw-Hill, New York, 1984.

2. W. M. Rohsenow, J. P. Hartnett, and E. N. Ganić (eds.), *Handbook of Heat Transfer Fundamentals,* 2d ed., McGraw-Hill, New York, 1985.

3. E. A. Avallone and T. Baumeister III (eds.), *Mark's Standard Handbook for Mechanical Engineers,* 9th ed., McGraw-Hill, New York, 1987.

4. O. W. Eshbach and M. Souders, *Handbook of Engineering Fundamentals,* John Wiley & Sons, New York, 1975.

5. T. G. Hicks, *Standard Handbook of Engineering Calculation,* 2d ed., McGraw-Hill, New York, 1985.

6. R. H. Perry, *Engineering Manual,* 3d ed., McGraw-Hill, New York, 1976.

7. J. Whitaker and B. Benson, *Standard Handbook of Video and Television Engineering,* 3d ed., McGraw-Hill, 2000.

8. R. Walsh, *McGraw-Hill Machining and Metalworking Handbook,* 2d ed., McGraw-Hill, 1999.

9. D. Cristiansen, *Electronics Engineer's Handbook,* 4th ed., McGraw-Hill, 1997.

---

*Those references listed above but not cited in the text were used for comparison between different data sources, clarification, clarity of presentation, and, most important, reader's convenience when further interest in the subject exists.

CHAPTER 2
GENERAL PROPERTIES OF MATERIALS

All materials have properties which must be known in order to promote their proper use. Knowing these properties is also essential to selecting the best material for a given application. This chapter includes general properties widely used in the field of chemical, mechanical, civil, and electrical engineering.

Note that results are given in SI units. Use Table 1.5 of Chap. 1 to obtain results in USGS units.

CHEMICAL PROPERTIES

Every elementary substance is made up of atoms which are all alike and which cannot be further subdivided or broken up by chemical processes. There are as many different classes or families of atoms as there are chemical elements (Table 2.1).

Two or more atoms, either of the same kind or of different kinds, are, in the case of most elements, capable of uniting with one another to form a higher order of distinct particles called *molecules.* If the molecules or atoms of which any given material is composed are all exactly alike, the material is a pure substance. If they are not all alike, the material is a mixture.

If the atoms which compose the molecules of any pure substances are all of the same kind, the substance is, as already stated, an elementary substance. If thc atoms which compose the molecules of a pure chemical substance are not all of the same kind, thc substance is a compound substance.

It appears that some substances which cannot by any available means be decomposed into simpler substances and which must, therefore, be defined as elements, are continually undergoing *spontaneous* changes or radioactive transformation into other substances which can be recognized as physically different from the original substance. The view generally accepted at present is that the atoms of all the chemical elements, including those not known to be radioactive, consist of several kinds of still smaller particles, three of which are known as protons, neutrons, and electrons. The protons are bound together in the atomic nucleus with other particles, including neutrons, and are positively charged. The neutrons are particles having approximately the mass of a proton but no charge. The electrons are negatively charged particles, all alike, external to the nucleus; and sufficient in number to neutralize the nuclear charge in an atom. The differences between the atoms of

TABLE 2.1   Chemical Elements[a]

Element	Symbol	Atomic No.	Atomic weight[b]
Actinium	Ac	89	
Aluminum	Al	13	26.9815
Americium	Am	95	
Antimony	Sb	51	121.75
Argon[c]	Ar	18	39.948
Arsenic[d]	As	33	74.9216
Astatine	At	85	
Barium	Ba	56	137.34
Berkelium	Bk	97	
Beryllium	Be	4	9.0122
Bismuth	Bi	83	208.980
Boron[d]	B	5	10.811[l]
Bromine[e]	Br	35	79.904[m]
Cadmium	Cd	48	112.40
Calcium	Ca	20	40.08
Californium	Cf	98	
Carbon[d]	C	6	12.01115[l]
Cerium	Ce	58	140.12
Cesium[k]	Ca	55	132.905
Chlorine[f]	Cl	17	35.453[m]
Chromium	Cr	24	51.996[m]
Cobalt	Co	27	58.9332
Columbium (see Niobium)			
Copper	Cu	29	63.546[m]
Curium	Cm	96	
Dysprosium	Dy	66	162.50
Einsteinium	Es	99	
Erbium	Er	68	167.26
Europium	Eu	63	151.96
Fermium	Fm	100	
Fluorine[g]	F	9	18.9984
Francium	Fr	87	
Gadolinium	Gd	64	157.25
Gallium[k]	Ga	31	69.72
Germanium	Ge	32	72.59
Gold	Au	79	196.967
Hafnium	Hf	72	178.49
Helium[c]	He	2	4.0026
Holmium	Ho	67	164.930
Hydrogen[h]	H	1	1.00797[l]
Indium	In	49	114.82
Iodine[d]	I	53	126.9044
Iridium	Ir	77	192.2
Iron	Fe	26	55.847[m]
Krypton[c]	Kr	36	83.80
Lanthanum	La	57	138.91
Lead	Pb	82	207.19
Lithium[i]	Li	3	6.939
Lutetium	Lu	71	174.97
Magnesium	Mg	12	24.312

TABLE 2.1   Chemical Elements (*Continued*)

Element	Symbol	Atomic No.	Atomic weight[b]
Manganese	Mn	25	54.9380
Mendelevium	Md	101	
Mercury[e]	Hg	80	200.59
Molybdenum	Mo	42	95.94
Neodymium	Nd	60	144.24
Neon[c]	Ne	10	20.183
Neptunium	Np	93	
Nickel	Ni	28	58.71
Niobium	Nb	41	92.906
Nitrogen[f]	N	7	14.0067
Nobelium	No	102	
Osmium	Os	76	190.2
Oxygen[f]	O	8	15.9994[l]
Palladium	Pd	46	106.4
Phosphorus[d]	P	15	30.9738
Platinum	Pt	78	195.09
Plutonium	Pu	94	
Polonium	Po	84	
Potassium	K	19	39.102
Praseodymium	Pr	59	140.907
Promethium	Pm	61	
Protactinium	Pa	91	
Radium	Ra	88	
Radon[i]	Rn	86	
Rhenium	Re	75	186.2
Rhodium	Rh	45	102.905
Rubidium	Rb	37	85.47
Ruthenium	Ru	44	101.07
Samarium	Sm	62	150.35
Scandium	Sc	21	44.956
Selenium[d]	Se	34	78.96
Silicon[d]	Si	14	28.086[l]
Silver	Ag	47	107.868[m]
Sodium	Na	11	22.9898
Strontium	Sr	38	87.62
Sulphur[d]	S	16	32.064[l]
Tantalum	Ta	73	180.948
Technetium	Tc	43	
Tellurium[d]	Te	52	127.60
Terbium	Tb	65	158.924
Thallium	Tl	81	204.37
Thorium	Th	90	232.038
Thulium	Tm	69	168.934
Tin	Sn	50	118.69
Titanium	Ti	22	47.90
Tungsten	W	74	183.85
Uranium	U	92	238.03
Vanadium	V	23	50.942
Xenon[c]	Xe	54	131.30
Ytterbium	Yb	70	173.04

TABLE 2.1   Chemical Elements (*Continued*)

Element	Symbol	Atomic No.	Atomic weight[b]
Yttrium	Y	39	88.905
Zinc	Zn	30	65.37
Zirconium	Zr	40	91.22

[a]All the elements for which atomic weights are listed are metals, except as otherwise indicated. No atomic weights are listed for most radioactive elements, as these elements have no fixed value.

[b]The atomic weights are based upon nuclidic mass of $C^{12} = 12$.

[c]Inert gas. [d]Metalloid. [e]Liquid. [f]Gas. [g]Most active gas. [h]Lightest gas. [i]Lightest metal. [j]Not placed. [k]Liquid at 25°C.

[l]The atomic weight varies because of natural variations in the isotopic composition of the element. The observed ranges are boron, ± 0.003; carbon, ± 0.00005; hydrogen, ± 0.00001; oxygen, ± 0.0001; silicon, ± 0.001; sulfur, ± 0.003.

[m]The atomic weight is believed to have an experimental uncertainty of the following magnitude: bromine, ± 0.001; chlorine, ± 0.001; chromium, ± 0.001; copper, ± 0.001; iron, ± 0.003; silver, ± 0.001. For other elements, the last digit given is believed to be reliable to ± 0.5.

Source: From Avallone and Baumeister.[1]

different chemical elements are due to the different numbers of these smaller particles composing them.

In a hydrogen atom, there is one proton and one electron; in a radium atom, there are 88 electrons surrounding a nucleus 226 times as massive as the hydrogen nucleus. Only a few, in general the outermost or *valence* electrons of such an atom, are subject to rearrangement within, or ejection from, the atom, thereby enabling it, because of its increased energy, to combine with other atoms to form molecules of either elementary substances or compounds. The atomic number of an element is the number of excess positive charges on the nucleus of the atom. The essential feature that distinguishes one element from another is this charge of thc nucleus. It also determines the position of the element in the periodic table (Table 2.2). Modern research has shown the existence of isotopes, that is, two or more species of atoms having the same atomic number and thus occupying the same place in the periodic system, but differing somewhat in atomic weight. These isotopes are chemically identical and are merely different species of the same chemical element.

Data for solubility of inorganic substances and gases in water are given in Tables 2.3 and 2.4, respectively. Sec Refs. 1 and 3 for information on other chemical properties of materials.

THERMOPHYSICAL PROPERTIES

Most frequently used thermophysical properties in engineering practice are

Density (ρ)

Specific heat (c)

Specific heat at constant pressure (c_p)

Thermal conductivity (k)

TABLE 2.2 Periodic Table of the Elements

Legend (key):
- Atomic number
- Symbol
- Element
- Valence
- Atomic weight based on $C^{12} = 12.00$
- ( ) denotes mass number of most stable known isotope

Hydrogen — 1, H, 1.00797, valence 1

Inert gases
Z	Element	Symbol	Atomic weight	Valence
2	Helium	He	4.0026	0
10	Neon	Ne	20.183	0
18	Argon	Ar	39.948	0
36	Krypton	Kr	83.80	0
54	Xenon	Xe	131.30	0
86	Radon	Rn	(212)	0

Light metals
Z	Element	Symbol	Atomic weight	Valence
3	Lithium	Li	6.939	1
4	Beryllium	Be	9.0122	2
11	Sodium	Na	22.9898	1
12	Magnesium	Mg	24.312	2
19	Potassium	K	39.102	1
20	Calcium	Ca	40.08	2
21	Scandium	Sc	44.956	3
37	Rubidium	Rb	85.47	1
38	Strontium	Sr	87.62	2
39	Yttrium	Y	88.905	3
55	Cesium	Cs	132.905	1
56	Barium	Ba	137.34	2
57	Lanthanum	La	137.91	3
87	Francium	Fr	(223)	1
88	Radium	Ra	(227)	2
89	Actinium	Ac	(227)	3

Brittle metals
Z	Element	Symbol	Atomic weight	Valence
22	Titanium	Ti	47.90	3,4
23	Vanadium	V	50.942	1,2,3,4,5
24	Chromium	Cr	51.996	2,3,6
25	Manganese	Mn	54.938	2,3,4,6,7
40	Zirconium	Zr	91.22	4
41	Niobium	Nb	92.906	2,3,4,5
42	Molybdenum	Mo	95.94	3,4,5,6
43	Technetium	Tc	(99)	
72	Hafnium	Hf	178.49	4
73	Tantalum	Ta	180.948	4,5
74	Tungsten	W	183.85	3,4,5,6
75	Rhenium	Re	186.2	1,4,7

Ductile metals
Z	Element	Symbol	Atomic weight	Valence
26	Iron	Fe	55.847	2,3
27	Cobalt	Co	58.9332	2,3
28	Nickel	Ni	58.71	2,3,4
29	Copper	Cu	63.546	1,2
44	Ruthenium	Ru	101.07	3,4,6,8
45	Rhodium	Rh	103.905	3,4
46	Palladium	Pd	106.4	2,4
47	Silver	Ag	107.868	1
76	Osmium	Os	190.2	2,3,4,6,8
77	Iridium	Ir	192.2	2,3,4,6
78	Platinum	Pt	195.09	2,4
79	Gold	Au	196.967	1,3

Low melting
Z	Element	Symbol	Atomic weight	Valence
30	Zinc	Zn	65.37	2
48	Cadmium	Cd	112.40	2
80	Mercury	Hg	200.59	1,2

Group IIIA
Z	Element	Symbol	Atomic weight	Valence
5	Boron	B	10.811	3
13	Aluminum	Al	26.9815	3
31	Gallium	Ga	69.72	2,3
49	Indium	In	114.82	1,2,3
81	Thallium	Tl	204.37	1,3

Group IVA
Z	Element	Symbol	Atomic weight	Valence
6	Carbon	C	12.01115	2,4
14	Silicon	Si	28.086	4
32	Germanium	Ge	72.59	4
50	Tin	Sn	118.69	2,4
82	Lead	Pb	207.19	2,4

Nonmetallic elements

Group VA
Z	Element	Symbol	Atomic weight	Valence
7	Nitrogen	N	14.0067	3,5
15	Phosphorus	P	30.9738	3,5
33	Arsenic	As	74.9216	3,5
51	Antimony	Sb	121.75	3,5
83	Bismuth	Bi	208.98	3,5

Group VIA
Z	Element	Symbol	Atomic weight	Valence
8	Oxygen	O	15.9994	2
16	Sulphur	S	32.064	2,4,6
34	Selenium	Se	78.96	2,4,6
52	Tellurium	Te	127.60	2,4,6
84	Polonium	Po	(210)	2,4

Group VIIA
Z	Element	Symbol	Atomic weight	Valence
9	Fluorine	F	18.9984	1
17	Chlorine	Cl	35.453	1,3,5,7
35	Bromine	Br	79.904	1,3,5
53	Iodine	I	126.9044	1,3,5,7
85	Astatine	At	(210)	

Rare earth elements — LANTHANIDE SERIES
Z	Element	Symbol	Atomic weight	Valence
58	Cerium	Ce	140.12	3,4
59	Praseodymium	Pr	140.907	3
60	Neodymium	Nd	144.24	3
61	Promethium	Pm	(147)	3
62	Samarium	Sm	150.35	3
63	Europium	Eu	151.96	2,3
64	Gadolinium	Gd	157.25	3
65	Terbium	Tb	158.924	3
66	Dysprosium	Dy	162.50	3
67	Holmium	Ho	164.93	3
68	Erbium	Er	167.26	3
69	Thulium	Tm	168.934	3
70	Ytterbium	Yb	173.04	2,3
71	Lutetium	Lu	174.97	3

Transuranium elements — ACTINIDE SERIES
Z	Element	Symbol	Atomic weight	Valence
90	Thorium	Th	232.038	4
91	Protactinium	Pa	(231)	
92	Uranium	U	238.03	
93	Neptunium	Np	(237)	
94	Plutonium	Pu	(242)	
95	Americium	Am	(243)	
96	Curium	Cm	(245)	
97	Berkelium	Bk	(249)	
98	Californium	Cf	(249)	
99	Einsteinium	Es	(254)	
100	Fermium	Fm	(252)	
101	Mendelevium	Md	(256)	
102	Nobelium	No	(254)	
103	Lawrencium	Lw	(257)	

Source: From Avallone and Baumeister.[1]

2.5

TABLE 2.3 Solubility of Inorganic Substances in Water

(*Number of grams of the anhydrous substance soluble in 1000 g of water. The common name of the substance is given in parentheses.*)

	Composition	Temperature, °F(°C)		
		32 (0)	122 (50)	212 (100)
Aluminum sulfate	$Al_2(SO_4)_3$	313	521	891
Aluminum potassium sulfate (potassium alum)	$Al_2K_2(SO_4)_4 \cdot 24H_2O$	30	170	1540
Ammonium bicarbonate	NH_4HCO_3	119		
Ammonium chloride (sal ammoniac)	NH_4Cl	297	504	760
Ammonium nitrate	NH_4NO_3	1183	3440	8710
Ammonium sulfate	$(NH_4)_2SO_4$	706	847	1033
Barium chloride	$BaCl_2 \cdot 2H_2O$	317	436	587
Barium nitrate	$Ba(NO_3)_2$	50	172	345
Calcium carbonate (calcite)	$CaCO_3$	0.018*		0.88
Calcium chloride	$CaCl_2$	594		1576
Calcium hydroxide (hydrated lime)	$Ca(OH)_2$	1.77		0.67
Calcium nitrate	$Ca(NO_3)_2 \cdot 4H_2O$	931	3561	3626
Calcium sulfate (gypsum)	$CaSO_4 \cdot 2H_2O$	1.76	2.06	1.69
Copper sulfate (blue vitriol)	$CuSO_4 \cdot 5H_2O$	140	334	753
Ferrous chloride	$FeCl_2 \cdot 4H_2O$	644§	820	1060
Ferrous hydroxide	$Fe(OH)_2$	0.0067‡		
Ferrous sulfate (green vitriol or copperas)	$FeSO_4 \cdot 7H_2O$	156	482	
Ferric chloride	$FeCl_3$	730	3160	5369
Lead chloride	$PbCl_2$	6.73	16.7	33.3
Lead nitrate	$Pb(NO_3)_2$	403		1255
Lead sulfate	$PbSO_4$	0.042†		
Magnesium carbonate	$MgCO_3$	0.13‡		
Magnesium chloride	$MgCl_2 \cdot 6H_2O$	524		723
Magnesium hydroxide (milk of magnesia)	$Mg(OH)_2$	0.009‡		
Magnesium nitrate	$Mg(NO_3)_2 \cdot 6H_2O$	665	903	
Magnesium sulfate (Epsom salts)	$MgSO_4 \cdot 7H_2O$	269	500	710
Potassium carbonate (potash)	K_2CO_3	893	1216	1562
Potassium chloride	KCl	284	435	566
Potassium hydroxide (caustic potash)	KOH	971	1414	1773
Potassium nitrate (saltpeter or niter)	KNO_3	131	851	2477
Potassium sulfate	K_2SO_4	74	165	241
Sodium bicarbonate (baking soda)	$NaHCO_3$	69	145	
Sodium carbonate (sal soda or soda ash)	$NaCO_3 \cdot 10H_2O$	204	475	452
Sodium chloride (common salt)	$NaCl$	357	366	392

TABLE 2.3 Solubility of Inorganic Substances in Water (*Continued*)

(*Number of grams of the anhydrous substance soluble in 1000 g of water. The common name of the substance is given in parentheses.*)

	Composition	Temperature, °F(°C)		
		32 (0)	122 (50)	212 (100)
Sodium hydroxide (caustic soda)	NaOH	420	1448	3388
Sodium nitrate (Chile salt-peter)	NaNO$_3$	733	1148	1755
Sodium sulfate (Glauber salts)	Na$_2$SO$_4$·10H$_2$O	49	466	422
Zinc chloride	ZnCl$_2$	2044	4702	6147
Zinc nitrate	Zn(NO$_3$)$_2$·6H$_2$O	947		
Zinc sulfate	ZnSO$_4$·7H$_2$O	419	768	807

*59°F (15°C).
§50°F (10°C).
‡In cold water.
†68°F (20°C).
Source: From Avallonc and Baumeister.[1]

TABLE 2.4 Solubility of Gases in Water

(*By volume, at atmospheric pressure*)

t, °F (°C)	32 (0)	68 (20)	212 (100)
Air	0.032	0.020	0.012
Acetylene	1.89	1.12	
Ammonia	1250	700	
Carbon dioxide	1.87	0.96	0.26
Carbon monoxide	0.039	0.025	
Chlorine	5.0	2.5	0.00
Hydrogen	0.023	0.020	0.018
Hydrogen sulfide	5.0	2.8	0.87
Hydrochloric acid	560	480	
Nitrogen	0.026	0.017	0.0105
Oxygen	0.053	0.034	0.0185
Sulfuric acid	87	43	

Source: From Avallone and Baumeister.[1]

Thermal diffusivity (α)

Dynamic viscosity (μ)

Kinematic viscosity (ν)

Surface tension (σ)

Coefficient of thermal expansion (β)

The kinematic viscosity of a fluid is its dynamic viscosity divided by its density, or $\nu = \mu/\rho$. Its units are m^2/s. The surface tension of a fluid is the work done in

TABLE 2.5 Properties of Metallic Solids

Metal	Properties at 20°C				Thermal conductivity, k(W/m·K)									
	ρ (kg/m³)	c_p (J/kg·K)	k (W/m·K)	α (10⁻⁶ m²/s)	−170°C	−100°C	0°C	100°C	200°C	300°C	400°C	600°C	800°C	1000°C
Aluminum														
Pure	2,707	905	237	9.61	302	242	236	240	238	234	228	215	~95 (liq)	
99% pure			211		220	206	209							
Duralumin (~4% Cu)	2,787	883	164	6.66		126	164	182	194					
Chromium	7,190	453	90	2.77	158	120	95	88	85	82	77	69	64	62
Copper and Cu alloys														
Pure	8,954	384	398	11.57	483	420	401	391	389	384	378	366	352	336
Brass (30% Zn)	8,522	385	109	3.32	73	89	106	133	143	146	147			
Bronze (25% Sn)	8,666	343	26	0.86	(Data on this and other bronzes vary by a factor of about 2)									
Constantan (40% Ni)	8,922	410	22	0.61	17	19	22	26	35					
German silver (15% Ni, 22% Zn)	8,618	394	25	0.73	18	19	24	31	40	45	48			
Gold	19,320	129	315	12.64			318		309					
Ferrous metals														
Pure iron	7,897	447	80	2.26	132	98	84	72	63	56	50	39	30	29.5
Cast iron (0.4% C)	7,272	420	52	1.70										
Steels (C < 1.5%)														
0.5% carbon (mild)	7,833	465	54	1.47			55	52	48	45	42	35	31	29
1.0% carbon	7,801	473	43	1.17			43	43	42	40	36	33	29	28
1.5% carbon	7,753	486	36	0.97			36	36	36	35	33	31	28	28

Metal	ρ (kg/m³)	c_p	α	k										
Stainless steel, type:														
304	8,000	400	0.4	13.8				15	17	19	21	25	24	26
316	8,000	460	0.37	13.5				15	16	17	19	21	26	28
347	8,000	420	0.44	15		12		16	18	19	20	23		
410	7,700	460	0.7	25		13		25	26	27	27	28		
414	7,700	460	0.7	25							29	29		
Lead	11,373	130	2.34	35	40	37	36	34	33	31		20 (liq.)	17 (liq.)	
Magnesium	1,746	1023	8.76	156	17	16	157	154	152	150	148	90 (liq.)		
Mercury (polycrystalline)					32	30	7.8 (liq.)							
Nickel														
Pure	8,906	445	2.30	91	156	114	94	83	74	67	64	69	73	78
Nichrome (24% Fe, 16% Cr)	8,250	448	0.34					13						
Nichrome V (20% Cr)	8,410	466	0.33	13			12	14	15	17	19			
Platinum	21,450	133	2.50	71	78	73	72	72	72	73	74	77	80	84
Silver														
99.99% pure	10,524	236	17.19	427	449	431	428	422	417	409	401	386	370	176 (liq.)
99.9% pure	10,524	236	16.55	411	422	405	373	367	364					
Tin (polycrystalline)	7,304	~220	4.17	67	85	76	68	63						
Titanium (polycrystalline)	4,540	523	0.93	22	31	26	22	21	20	20	19	21	21	22
Tungsten	19,350	133	6.92	178	235	223	182	166	153	141	134	125	122	114
Uranium	18,700	116	1.29	28	22	24	27	29	31	33	36	41	46	
Zinc	7,144	388	4.37	121	124	122	117	110	106	100	60 (liq.)			

Source: From Lienhard.[2] Portions of the original table have been omitted where not relevant to this chapter. The data can also be found in Refs. 1, 2, and 4 through 9.

extending the surface of a liquid one unit of area or work per unit area. Its units are N/m. Also, note that $\alpha = k/\rho c$ and

$$\beta = -\frac{1}{\rho}\left(\frac{\partial \rho}{\partial t}\right)_p = \text{const}$$

In general, all thermophysical properties are strong functions of temperature.

Table 2.5 shows properties of metallic solids. Table 2.6 shows properties of nonmetallic solids. Table 2.7 shows properties of saturated liquids. (Note that the Prandtl number $\text{Pr} = \nu/\alpha$.) Table 2.8 shows properties of gases at atmospheric pressure. Table 2.9 shows data of surface tension of various liquids. Approximate relations for thermal expansions are given in Table 2.14.

MECHANICAL PROPERTIES

Mechanical properties commonly used by engineers are

Ultimate tensile strength

Tensile yield strength

Elongation

Modulus of elasticity

Compressive strength

Shear strength

Endurance limit

Ultimate tensile strength is defined as the maximum load per unit of original cross-sectional area sustained by a material during a tension test. It is also called *ultimate strength.*

Tensile yield strength is defined as the stress corresponding to some permanent deformation from the modulus slope, e.g., 0.2 percent offset in the case of heat-treated alloy steels.

Elongation is defined as the amount of permanent extension in a ruptured tensile test specimen; it is usually expressed as a percentage of the original gage length. Elongation is usually taken as a measure of ductility.

Modulus of elasticity is the property of a material which indicates its rigidity. This property is the ratio of stress to strain within the elastic range.

On a stress-strain diagram, the modulus of elasticity is usually represented by the straight portion of the curve when the stress is directly proportional to the strain. The steeper the curve, the higher the modulus of elasticity and the stiffer the material.

Compressive strength is defined as the maximum compressive stress that a material is capable of developing based on the original cross-sectional area. The general design practice is to assume the compressive strength of a steel is equal to its tensile strength, although it is actually somewhat greater.

Shear strength is defined as the stress required to produce fracture in the plane of cross section, the conditions of loading being such that the directions of force and of resistance are parallel and opposite although their paths are offset a specified minimum amount. The ultimate shear strength is generally assumed to be three-fourths the material's ultimate tensile strength.

TABLE 2.6 Properties of Nonmetallic Solids

Material	Temperature range, °C	Density ρ, kg/m³	Specific heat c, J/kg · °C	Thermal conductivity k, W/m · °C	Thermal diffusivity α, m²/s
Asbestos					
Cement board	20			0.6	
Fiber (properties vary	20	1930		0.8	
with packing)	20	980		0.14	
Asphalt	20–25			0.75	
Beef	25				1.35×10^{07}
Brick					
B&W, K-28 insulating	300			0.3	
B&W, K-28 insulating	1000			0.4	
Cement	10	720		0.34	
Common	0–1000			0.7	
Chrome	100			1.9	
Firebrick	300	2000	960	0.1	5.4×10^{-8}
Firebrick	1000			0.2	
Carbon					
Diamond (type II b)	20	~3250	510	1350.	8.1×10^{-4}
Graphite	20	~2100	~2090	Highly variable structure	
Cardboard	0–20	790		0.14	
Clay					
Fireclay	500–750			1.	
Sandy clay	20	1780		0.9	
Coal					
Anthracite	900	~1500		~0.2	
Brown coal	900			~0.1	
Bituminous in situ		~1300		0.5–0.7	$3 \text{ to } 4 \times 10^{-7}$
Concrete					
Limestone gravel	20	1850		0.6	
Portland cement	90	2300		1.7	
Sand:cement (3:1)	230			0.1	
Slag cement	20			0.14	
Corkboard (medium ρ)	30	170		0.04	
Egg white	20				1.37×10^{-7}
Glass					
Lead	36			1.2	
Plate	20			1.3	
Pyrex	60–100	2210	753	1.3	7.8×10^{-7}
Soda	20			0.7	
Window	46			1.3	
Glass wool	20	64–160		0.04	
Ice	0	917	2100	2.215	1.15×10^{-6}
Ivory	80			0.5	
Kapok	30			0.035	

TABLE 2.6 Properties of Nonmetallic Solids (*Continued*)

Material	Tem-perature range, °C	Density ρ, kg/m³	Specific heat c, J/kg · °C	Thermal conduc-tivity k, W/m · °C	Thermal diffusivity α, m²/s
Magnesia (85%)	38			0.067	
	93			0.071	
	150			0.074	
	204			0.08	
Lunar surface dust (high vacuum)	250	1500 ± 300	~600	~0.0006	~7 × 10⁻¹⁰
Rock wool	−5	~130		0.03	
	93			0.05	
Rubber (hard)	0	1200	2010	0.15	6.2 × 10⁻⁸
Silica aerogel	0	140		0.024	
	120	136		0.022	
Silo-cel (diatomaceous earth)	0	320		0.061	
Soil (mineral)					
Dry	15	1500	1840	1.	4 × 10⁻⁷
Wet	15	1930		2.	
Stone					
Granite (NTS)	20	~2640	~820	1.6	~7.4 × 10⁻⁷
Limestone (Indiana)	100	2300	~900	1.1	~5.3 × 10⁻⁷
Sandstone (Berea)	25			~3	
Slate	100			1.5	
Wood (perpendicular to grain)					
Ash	15	740		0.15–0.3	
Balsa	15	100		0.05	
Cedar	15	480		0.11	
Fir	15	600	2720	0.12	7.4 × 10⁻⁸
Mahogany	20	700		0.16	
Oak	20	600	2390	0.1–0.4	(0.7–2.8) × 10⁻⁷
Pitch pine	20	450		0.14	
Sawdust (dry)	17	128		0.14	
Spruce	20	410		0.11	
Wool (sheep)	20	145		0.05	

Source: From Lienhard.[2] Portions of the original table have been omitted where not relevant to this chapter. The data can also be found in Refs. 1, 2, and 4 through 9.

Endurance limit is defined as the maximum stress to which the material can be subjected for an indefinite service life. Although the standards vary for various types of members and different industries, it is common practice to assume that carrying a certain load for several million cycles of stress reversals indicates that the load can be carried for an indefinite time.

Hardness measures the resistance of the material to indentation. Hardness tests measure the plastic deformation (the size or depth) of an indentation. Brinell hard-

TABLE 2.7 Thermophysical Properties of Saturated Liquids

Temp., K	°C	ρ, kg/m³	c_p, J/kg·K	k, W/m·K	α, m²/s	ν, m²/s	Pr	β, K⁻¹
Ammonia (there is considerable disagreement among sources)
220	−53	706	4426	0.66	2.11×10^{-7}			
240	−33	682	4484	0.61	2.00	4.17×10^{-7}	2.09	
260	−13	656	4547	0.57	1.91	3.27	1.71	
280	7	629	4625	0.52	1.79	2.68	1.50	0.00025
300	27	600	4736	0.470	1.65	2.32	1.41	
320	47	568	4962	0.424	1.50	2.06	1.37	
340	67	533	5214	0.379	1.36	1.79	1.32	
360	87	490	5635	0.335	1.21	1.55	1.28	
380	107	436		0.289		1.34		
400	127	345		0.245		1.19		
CO₂
250	−23	1046	1990	0.135	6.49×10^{-8}			
260	−13	998	2110	0.123	5.84	1.15×10^{-7}	1.97	
270	−3	944	2390	0.113	5.09	1.08	2.12	
280	7	883	2760	0.102	4.19	1.04	2.48	
290	17	805	3630	0.090	3.08	0.99	3.20	0.014
300	27	676	7690	0.076	1.46	0.88	6.04	
303	30	604						
D₂O (heavy water)
589	316	740	2034	0.0509	0.978×10^{-7}	1.23×10^{-7}	1.257	
Freon 11 (trichlorofluoromethane)
220	−53		829	0.110				
240	−33	1607	841	0.105	7.8×10^{-8}	4.78×10^{-7}	6.1	
260	−13	1564	855	0.099	7.4	4.10	5.5	
280	7	1518	871	0.093	7.0	3.81	5.4	0.00154
300	27	1472	888	0.088	6.7	2.82	4.2	0.00163
320	47	1421	906	0.082	6.4			
340	67	1369	927	0.076	6.0			
Freon 12 (dichlorodifluoromethane)
160	−113			0.133				
180	−93	1664	834	0.124	8.935×10^{-8}			
200	−73	1610	856	0.1148	8.33			
220	−53	1555	873	0.1057	7.79	3.2×10^{-7}	4.11	0.00263
240	−33	1498	892	0.0965	7.22	2.60	3.60	
260	−13	1438	914	0.0874	6.65	2.26	3.40	
280	7	1374	942	0.0782	6.04	2.06	3.41	
300	27	1305	980	0.0690	5.39	1.95	3.62	
320	47	1229	1031	0.0599	4.72	1.9	4.03	
340	67		1097	0.0507				
Glycerin (or glycerol)
273	0	1276	2200	0.282	1.00×10^{-7}	0.0083	83,000	
293	20	1261	2350	0.285	0.962	0.001120	11,630	0.00048
303	30	1255	2400	0.285	0.946	0.000488	5,161	0.00049
313	40	1249	2460	0.285	0.928	0.000227	2,451	0.00049
323	50	1243	2520	0.285	0.910	0.000114	1,254	0.00050

TABLE 2.7 Thermophysical Properties of Saturated Liquids (*Continued*)

Temp., K	Temp., °C	ρ, kg/m³	c_p, J/kg · K	k, W/m · K	α, m²/s	ν, m²/s	Pr	β, K⁻¹
Glycerin (or glycerol) (*Continued*)								
644	371	10,540	159	16.1	1.084×10^{-5}	2.276×10^{-7}	0.024	
755	482	10,442	155	15.6	1.223	1.85	0.017	
811	538	10,348	145	15.3	1.02	1.68	0.017	
Mercury								
234	−39		141.5	6.97	3.62×10^{-6}			
250	−23		140.5	7.32	3.83			
300	27	13,611	139.1	8.34	4.41	1.2×10^{-7}	0.027	
350	77	3,489	137.7	5.29	4.91	1.0	0.020	
400	127	13,367	136.7	5.69	5.83×10^{-6}	0.95×10^{-7}	0.016	
500	277	13,128	135.6	6.36	6.00	0.80	0.013	
600	327		135.4	6.93	6.55	0.68	0.010	
700	427		136.1	7.34				
800	527			7.40				
Methyl alcohol (methanol)								
260	−13	823	2336	0.2164	1.126×10^{-7}	-1.3×10^{-6}	~11.5	
280	7	804	2423	0.2078	1.021	~0.9	~8.8	0.00114
300	27	785	2534	0.2022	1.016	~0.7	~6.9	
320	47	767	2672	0.1965	0.959	~0.6	~6.3	
340	67	748	2856	0.1908	0.893	~0.44	~4.9	
360	87	729	3036	0.1851	0.836	~0.36	~4.3	
380	107	710	3265	0.1794	0.774	~0.30	~4.1	
Oxygen								
54		1276	1648	0.191	9.08×10^{-8}	6.5×10^{-7}	7.15	
60	−213		1649	0.185				
80	−193		1653	0.1623				
90	−183	1114	1655	0.1501	8.14×10^{-8}	1.75×10^{-7}	2.15	
120	−153			0.1096				
150	−123			0.061				
Oils (some approximate viscosities)								
273	0	MS-20				0.0076	100,000	
339	66	California crude (heavy)				0.00008		
289	16	California crude (light)				0.00005		
339	66	California crude (light)				0.000010		
289	16	Light machine oil				0.0007		
339	66	Light machine oil				0.00004		
289	16	Light machine oil ($\rho = 907$)				0.00016		
339	66	Light machine oil ($\rho = 907$)				0.000013		
289	16	SAE 30				0.00044	−5,000	
339	66	SAE 30				0.00003		
289	16	SAE 30 (Eastern)				0.00011		
339	66	SAE 30 (Eastern)				0.00001		
289	16	Spindle oil ($\rho = 885$)				0.00005		
339	66	Spindle oil ($\rho = 885$)				0.000007		

TABLE 2.7 Thermophysical Properties of Saturated Liquids (*Continued*)

Temp.,		ρ,	c_p,	k,				β,
K	°C	kg/m³	J/kg · K	W/m · K	α, m²/s	ν, m²/s	Pr	K⁻¹
					Water			
273	0	999.8	4205	0.5750	1.368×10^{-7}	1.753×10^{-6}	12.81	
280	7	999.9	4196	0.5818	1.386	1.422	10.26	
300	27	996.6	4177	0.6084	1.462	0.826	5.65	0.000275
320	47	989.3	4177	0.6367	1.541×10^{-7}	0.566×10^{-6}	3.67	0.000435
340	67	979.5	4187	0.6587	1.606	0.420	2.61	
360	87	967.4	4206	0.6743	1.657	0.330	1.99	
373	100	957.2	4219	0.6811	1.683	0.290	1.72	
400	127	937.5	4241	0.6864	1.726	0.229	1.33	
420	147	919.9	4306	0.6836	1.726×10^{-7}	2.000×10^{-7}	1.16	
440	167	900.5	4391	0.6774	1.713	1.786	1.04	
460	187	879.5	4456	0.6672	1.703	1.626	0.955	
480	207	856.6	4534	0.6530	1.681	1.504	0.894	
500	227	831.5	4647	0.6348	1.463	1.412	0.859	
520	247	803.9	4831	0.6123	1.577×10^{-7}	1.345×10^{-7}	0.853	
540	267	773.0	5099	0.5857	1.486	1.298	0.873	
560	287	738.2	5487	0.555	1.370	1.269	0.926	
580	307	697.6	6010	0.520	1.240	1.240	1.000	
600	327	648.8	6691	0.481	1.108	1.215	1.097	
620	347	586.3				1.213×10^{-7}		
640	367	482.1				1.218		
647.3	374.2	306.8				1.356		

Source: From Lienhard. Portions of the original table have been omitted where not relevant to this chapter. The data can also be found in Refs. 1. 2, and 4 through 9.

ness tests use spheres as indenters; the Vickers test uses pyramids. Rockwell tests use cones or spheres. Microhardness tests for specimens are also available, using the Knoop method with miniature pyramid indenters. Another hardness scale is Mohs' scale, which lists materials in order of their hardness, beginning with talc and ending with diamond. Table 2.15 shows typical Brinell hardness number (BHN) for metals.

Table 2.10 shows typical mechanical properties of some metals and alloys. See Ref. 11 for more data.

Note on Hardness Testing Method

Hardness tests on materials consist of pressing a hardened ball point into a specimen and measuring the size of the resulting indentation (see Figure 2.1). The method shown is the Brinell method, which utilizes a ball. The ball size is 10 mm for most cases or 1 mm for light work.

Let:

$$D = \text{diameter of indentation (mm)}$$

$$D_b = \text{diameter of baIl (mm)}$$

$$F = \text{force on ball (kg}_\text{f})$$

TABLE 2.8 Thermophysical Properties of Gases at Atmospheric Pressure

T, K	ρ, kg/m^3	c_p, J/kg·K	μ, kg/m·s	ν, m^2/s	k, W/m·K	α, m^2/s	Pr
Air							
100		1009	0.706×10^{-5}	$\sim 0.2 \times 10^{-5}$	0.00922		0.74
150		1005	1.038	~ 0.4	0.01375		0.724
200	1.79	1003	1.336	0.746	0.01810	1.01×10^{-5}	0.711
250	1.43	1003	1.606	1.123	0.02226	1.55	0.706
300	1.183	1003	1.853	1.566	0.02614	2.203	0.703
350	1.009	1008	2.081	2.062	0.02970	2.920	0.706
400	0.8826	1013	2.294	2.599	0.03305	3.697	0.703
450	0.7846	1020	2.493	3.177	0.03633	4.540	0.700
500	0.7061	1029	2.682×10^{-5}	3.798×10^{-5}	0.03951	5.438×10^{-5}	0.698
550	0.6419	1039	2.860	4.456	0.0426	6.387	0.698
600	0.5884	1051	3.030	5.150	0.0456	7.374	0.698
650	0.5432	1063	3.193	5.878	0.0484	8.382	0.701
700	0.5044	1075	3.349	6.64	0.0513	9.461	0.702
750	0.4707	1087	3.498	7.43	0.0541	10.57	0.703
800	0.4413	1099	3.643	8.26	0.0569	11.73	0.704
850	0.4154	1110	3.783	9.11	0.0597	12.95	0.704
900	0.3923	1121	3.918	9.99	0.0625	14.22	0.703
950	0.3716	1131	4.049	10.90	0.0649	15.44	0.706
1000	0.3531	1142	4.177×10^{-5}	11.83×10^{-5}	0.0672	16.67×10^{-5}	0.710
1100	0.3210	1159	4.42	13.8	0.0717	19.27	0.716
1200	0.2942	1175	4.65	15.8	0.0759	29.96	0.720
1300	0.2716	1189	4.88	18.0	0.0797	24.7	0.729
1400	0.2522		5.09	20.2	0.0835	27.5	0.734
1500	0.2354		5.30	22.5	0.0870	30.3	0.74
Ammonia (NH$_3$)							
220	0.3828	2.198×10^3	7.255×10^{-6}	1.90×10^{-5}	0.0171	0.2054×10^{-4}	0.93
273	0.7929	2.177	9.353	1.18	0.0220	0.1308	0.90
323	0.6487	2.177	11.035	1.70	0.0270	0.1920	0.88
373	0.5590	2.236	12.886	2.30	0.0327	0.2619	0.87
423	0.4934	2.315	14.672	2.97	0.0391	0.3432	0.87
473	0.4405	2.395	16.49	3.74	0.0467	0.4421	0.84

Carbon dioxide

T	ρ	c_p	μ	ν	k	α	Pr
220	2.4733	0.783×10^3	11.105×10^{-6}	4.490×10^{-6}	0.010805	0.05920×10^{-4}	0.818
250	2.1657	0.804	12.590	5.813	0.012884	0.07401	0.793
300	1.7973	0.871	14.958	8.321	0.016572	0.10588	0.770
350	1.5362	0.900	17.205	11.19	0.02047	0.14808	0.755
400	1.3424	0.942	19.32	14.39	0.02461	0.19463	0.738
450	1.1918	0.980	21.34	17.90	0.02897	0.24813	0.721
500	1.0732	1.013	23.26	21.67	0.03352	0.3084	0.702
550	0.9739	1.047	25.08	25.74	0.03821	0.3750	0.685
600	0.8938	1.076	26.83	30.02	0.04311	0.4483	0.668

Helium

T	ρ	c_p	μ	ν	k	α	Pr
3	1.4657	5.200×10^3	8.42×10^{-7}	3.42×10^{-6}	0.0106	0.04625×10^{-4}	0.74
33	3.3799	5.200	50.2	37.11	0.0353	0.5275	0.70
144	0.2435	5.200	125.5	64.38	0.0928	0.9288	0.694
200	0.1906	5.200	156.6	95.50	0.1177	1.3675	0.70
255	0.13280	5.200	181.7	173.6	0.1357	2.449	0.71
366	0.10204	5.200	230.5	269.3	0.1691	3.716	0.72
477	0.08282	5.200	275.0	375.8	0.197	5.215	0.72
589	0.07032	5.200	311.3	494.2	0.225	6.661	0.72
700	0.06023	5.200	347.5	634.1	0.251	8.774	0.72
800	0.05286	5.200	381.7	781.3	0.275	10.834	0.72
900		5.200	413.6		0.298		0.72

Hydrogen

T	ρ	c_p	μ	ν	k	α	Pr
30	0.84722	10.840×10^3	1.606×10^{-6}	1.805×10^{-6}	0.0228	0.0249×10^{-4}	0.759
50	0.50955	10.501	2.516	4.880	0.0362	0.0676	0.721
100	0.24572	11.229	4.212	17.14	0.0665	0.2408	0.7122
150	0.16371	12.602	5.595	34.18	0.0981	0.475	0.718
200	0.12270	13.540	6.813	55.53	0.1282	0.772	0.719
250	0.09819	14.059	7.919	80.64	0.1561	1.130	0.713
300	0.08185	14.314	8.963	109.5	0.182	1.554	0.706

2.17

TABLE 2.8 Thermophysical Properties of Gases at Atmospheric Pressure (*Continued*)

× T, K	ρ, kg/m³	c_p, J/kg·K	μ, kg/m·s	ν, m²/s	k, W/m·K	α, m²/s	Pr
			Hydrogen (*Continued*)				
350	0.07016	14.436	9.954	141.9	0.206	2.031	0.697
400	0.06135	14.491	10.864	177.1	0.228	2.568	0.690
450	0.05462	14.499	11.779	215.6	0.251	3.164	0.682
500	0.04918	14.507	12.636	257.0	0.272	3.817	0.675
550	0.04469	14.532	13.475	301.6	0.292	4.516	0.668
600	0.04085	14.537×10^{-3}	14.285×10^{-6}	349.7×10^{-6}	0.315	5.306×10^{-4}	0.664
700	0.03492	14.574	15.89	455.1	0.351	6.903	0.659
800	0.03060	14.675	17.40	569	0.384	8.563	0.664
900	0.02723	14.821	18.78	690	0.412	10.21	0.675
1000	0.02451	14.968	20.16	822	0.445	12.13	0.678
1100	0.02227	15.165	21.46	965	0.488	14.45	0.668
1200	0.02050	15.366	22.75	1107	0.528	16.76	0.661
1300	0.01890	15.575	24.08	1273	0.568	19.3	0.660
1333	0.01842	15.638	24.44	1328	0.58	20.1	0.661
			Nitrogen				
100	3.439	1.0722×10^{3}	6.862×10^{-6}	1.995×10^{-6}	0.00958	0.026×10^{-4}	0.767
200	1.688	1.0429	12.947	7.67	0.0183	0.104	0.738
300	1.1233	1.0408	17.84	15.88	0.0259	0.222	0.715
400	0.8425	1.0459	21.98	26.1	0.0327	0.371	0.704
500	0.6739	1.0555	25.70	38.1	0.0389	0.547	0.696
600	0.5615	1.0756	29.11	51.8	0.0446	0.738	0.702
700	0.4812	1.0969	32.13	66.8	0.0499	0.945	0.707
800	0.4211	1.1225	34.84	82.7	0.0548	1.16	0.713
900	0.3743	1.1464	37.49	100.2	0.0597	1.39	0.721
1000	0.3368	1.1677	40.00	119.	0.0647	1.65	0.723
1100	0.3062	1.1857	42.28	138.	0.0700	1.93	0.716
1200	0.2807	1.2037	44.50	158.	0.0758	2.24	0.704

Oxygen

100	3.9918	0.9479×10^2	7.768×10^{-6}	1.946×10^{-6}	0.00903	0.0239×10^{-4}	0.815
150	2.6190	0.9178	11.490	4.387	0.01367	0.0569	0.773
200	1.9559	0.9131	14.850	7.593	0.01824	0.1021	0.745
250	1.5618	0.9157	17.87	11.45	0.02259	0.1579	0.725
300	1.3007	0.9203	20.63	15.86	0.02676	0.2235	0.709
350	1.1133	0.9291	23.16	20.80	0.03070	0.2968	0.702
400	0.9755	0.9420	25.54	26.18	0.03161	0.3768	0.695
450	0.8682	0.9567	27.77	31.99	0.03828	0.4609	0.694
500	0.7801	0.9722	29.91	38.34	0.04173	0.5502	0.697
550	0.7096	0.9881	31.97	45.05	0.04517	0.6441	0.700
600	0.6504	1.0044	33.92	52.15	0.04832	0.7399	0.704

Steam (H_2O vapor)

373.15	0.597	2030	12.28×10^{-6}	21.28×10^{-6}	0.0237	2.023×10^{-5}	1.052
393.15	0.547	1997	13.04	23.85	0.0251	2.298	1.038
413.15	0.520	1980	13.81	26.56	0.0265	2.574	1.032
433.15	0.494	1972	14.59	29.53	0.0280	2.874	1.027
453.15	0.473	1963	15.38	32.52	0.0294	3.166	1.027
473.15	0.452	1963	16.18	35.80	0.0309	3.483	1.029
493.15	0.433	1968	17.00	39.25	0.0323	3.790	1.036
513.15	0.416	1972	17.81	42.82	0.0338	4.120	1.039
533.15	0.400	1976	18.63	46.58	0.0354	4.479	1.040
553.15	0.386	1985	19.46×10^{-6}	50.42×10^{-6}	0.0369	4.816×10^{-5}	1.047
573.15	0.372	1997	20.29	54.54	0.0385	5.183	1.052
593.15	0.359	2010	21.12	58.84	0.0401	5.557	1.059
613.15	0.348	2022	21.95	63.09	0.0416	5.912	1.067

TABLE 2.9 Surface Tension of Liquids

Substance	Temperature, K										
	250	260	270	280	290	300	320	340	360	380	400
Acetone	0.0292	0.0280	0.0267	0.0254	0.0241	0.0229	0.0203	0.0178	0.016	0.014	0.012
Benzene			0.0321	0.0307	0.0293	0.0279	0.0253	0.0228	0.0204	0.0180	0.0156
Bromine	0.047	0.046	0.045	0.044	0.0425	0.041	0.038	0.035	0.032	0.030	0.027
Butane	0.0176	0.0164	0.0152	0.0140	0.0128	0.0116	0.0092	0.0069	0.0049	0.0031	0.0016
Chlorine	0.0243	0.0227	0.0212	0.0197	0.0182	0.0167	0.0137	0.0107	0.0079	0.0051	0.0037
Decane	0.0278	0.0269	0.0260	0.0251	0.0241	0.0233	0.0215	0.0196	0.0178	0.0161	0.0145
Diphenyl						0.0416	0.0388	0.0362	0.0338	0.0316	0.0295
Ethane	0.0061	0.0049	0.0037	0.0026	0.0015	0.0007	—	—	—	—	—
Ethanol			0.0247	0.0239	0.0231	0.0222	0.0204	0.0186	0.0167	0.0148	0.0126
Ethylene	0.0033	0.0020	0.0009	0.0002	—	—	—	—	—	—	—
Heptane	0.0242	0.0233	0.0224	0.0214	0.0204	0.0194	0.0175	0.0156	0.0137	0.0118	0.0100
Hexane	0.0230	0.0219	0.0207	0.0198	0.0187	0.0176	0.0154	0.0134	0.0116	0.0096	0.0077
Methanol	0.0266	0.0257	0.0248	0.0238	0.0229	0.0221	0.0204	0.0187	0.0169	0.0150	0.0129
Nonane	0.0270	0.0261	0.0251	0.0242	0.0232	0.0223	0.0204	0.0186	0.0167	0.0148	0.0129
Octane	0.0256	0.0247	0.0237	0.0228	0.0219	0.0210	0.0191	0.0173	0.0155	0.0138	0.0123
Pentane	0.0210	0.0198	0.0186	0.0175	0.0164	0.0153	0.0131	0.0108	0.0088	0.0069	0.0053
Propane	0.0128	0.0114	0.0101	0.0088	0.0076	0.0064	0.0043	0.0025	0.0007	—	—
Propanol	0.0274	0.0266	0.0258	0.0249	0.0241	0.0232	0.0214	0.0198	0.0182	0.0168	0.0155
Propylene	0.0132	0.0119	0.0105	0.0090	0.0077	0.0064	0.0041	0.0022	0.0005	—	—
R 12	0.0147	0.0134	0.0121	0.0108	0.0095	0.0082	0.0057	0.0034			
Toluene	0.0345	0.0330	0.0315	0.0301	0.0288	0.0275	0.0251	0.0227	0.0205	0.0185	0.0165
Water	—	—	—	0.0747	0.0733	0.0717	0.0685	0.0651	0.0615	0.0576	0.0536

Tabular values in N/m.
Source: From Rohsenow, Hartnett, and Ganić.[3]

TABLE 2.10 Mechanical Properties of Some Metals and Alloys at Room Temperature

Material	Composition	Condition	Yield strength (or 0.2% proof stress), MPa	Ultimate tensile stress, MPa	Elongation on 2 in, %	Hardness*
Ag	99.9	Annealed 600–650°C	7.6†	137	50	26 VHN
		Hard	...	380	4	90 VHN
Al	99.95	Rolled rod 0	...	55	61	17 B
Au	99.99	Soft, cast	0	121	30	33 B
		Hard, 60% red.	21.2	228	4	58 B
Co	99.9	Soft	190	240	4–8	124 B
		Hard	...	675	2–8	165 B
Cu	99.997	Annealed	340	351	60	
		Rod, cold-drawn	34	213	14	37 RB
Ni	>99.0	Annealed	138	482	40	100 B
		Cold-drawn	482	654	25	170 B
Pt	99.99	Annealed	...	120–130	25–40	30–40 VHN
		50% cold-rolled	180	200	3	92 VHN
Pd	99.9	Annealed	35	190	40	37 VHN
		50% cold-drawn	...	320	1.5	106 VHN
Ta	99.98	Annealed	180	200	36	90 VHN
	>99.95	Cold-rolled	330	410	5	160 VHN
W		Swaged, recrys.	195	405	16	(200 VHN)
	99.9	Swaged	...	1,750	1–4	450–490 VHN
Aluminum alloys:						
1100	1% Si	O	27.5	69	45	19 B
1100		H18	124	131	15	35 B
3003	1–1.5% Mn	O	41	110	40	28 B
3003		H18	186	200	10	55 B
5056	4.5–5.6% Mg	O	152	290	35	65 B
5056		H38	345	415	15	100 B
7075	1.2–2% Cu, 2.1–2.9% Mg 5.1–5.6% Z	O	104	227	...	60 B
7075		T6	505	570	...	150 B
Copper alloys:						
OHFC, copper		Wire, soft	...	241	35	
		Wire, hard	...	380	1.5	
Gilding metal	5% Zn	Annealed, strip	69	234	45	46 RF
		Extra hard	390	435	4	73 RB
Red brass	15% Zn	Annealed	70–130	280–320	48	
		Half hard	350	405	12	65 RB
		Extra hard	435	555	4	83 RB
Yellow brass	35% Zn	Annealed	100–157	326–376	54–65	58–78 RF
		Half hard	426	526	8	80 RB
		Extra hard	44	605	5	87 RB
Phosphor bronze	5% Sn	Annealed	150–220	340–390	57–48	33–46 RB
		Hard	560	575	8	89 RB
		Extra spring	710	725	3	98 RB
Beryllium copper	1.9% Be	Annealed	...	430–550	35	45–78 RB
		HT (cold worked and precipta-tion hardened)	1,070	1,420	2	42 RC

TABLE 2.10 Mechanical Properties of Some Metals and Alloys at Room Temperature (*Continued*)

Material	Composition	Condition	Yield strength (or 0.2% proof stress), MPa	Ultimate tensile stress, MPa	Elongation on 2 in, %	Hardness*
Steels:						
C1010	0.08–0.13 C	Hot-rolled	188	340	28	95 B
		Cold-drawn	318	383	20	105 B
C1080	0.77–0.88%	Hot-rolled	440	810	10	229 B
12% Mn steel	12% Mn	Tempered 600°F	1,480	1,520	10	
Stainless steel	9% Ni, 19% Cr	Annealed	340	590	60	160 B
Type 304		Cold-rolled	1,110	1,280	...	400 B

*VHN = Vickers hardness number; B = Brinell; RB = Rockwell B; RC = Rockwell C; RF = Rockwell F.

†0.01% proof stress.
Source: From Fink and Christiansen.[4]
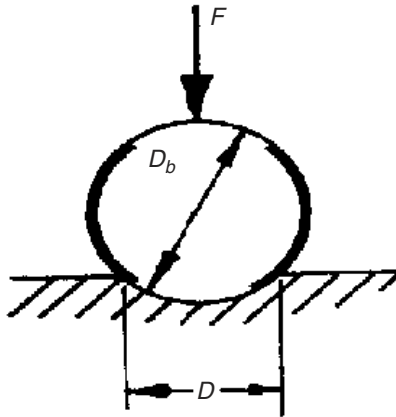
FIGURE 2.1

Values of F: steel, $F = 3OD_b^2$; copper, $F = 10D_b^2$; aluminium, $F = 5D_b^2$

Hardness BHN $= \dfrac{0.64F}{D_b(D_b - \sqrt{D_b^2 - D^2})}$

Source: Ref. 1, 13 and 15.

ELECTRICAL PROPERTIES

The most frequently used properties of materials in the field of electrical and electronics engineering are specific resistance (ρ_s) and the temperature coefficient of resistance (α_t).

Specific resistance (or *resistivity*) is the resistance of a sample of material having both a length and cross section of unity. The two most common resistivity samples are the centimeter cube and the cir mil · ft. If l is the length of a conductor of uniform cross section A, then the resistance is

$$R = \rho_s \cdot \frac{l}{A}$$

where ρ_s is the resistivity. A *circular mil* is a unit of area equal to that of a circle whose diameter is 1 mil (0.001 in).

The resistance of the pure metals increases with temperature as

$$R = R_0(1 + \alpha_t \cdot t)$$

where R_0 is the resistance at 0°C and α_t is the *temperature coefficient of resistance.* Over a narrow range of temperature, the electrical resistivity changes approximately linearly with temperature:

$$\rho_s(t_2) = \rho_s(t_1)[1 + \alpha_{t_1}(t_2 - t_1)]$$

where $\rho_s(t_1)$ is the resistivity at t_1, etc.

When the temperature of reference t_1 is changed to some other value of t, the coefficient α_{t_1} will change to a new value α_t:

$$\alpha_t = \frac{\alpha_{t_1}}{1 + \alpha_{t_1}(t - t_1)}$$

Values of ρ_s are given in Tables 2.11 to 2.13 and Fig. 2.3. Values of α_t are given in Table 2.11. Table 2.11 also includes values of melting point and density.

OTHER ENGINEERING MATERIAL DATA

Composites

A composite is a material consisting of two (or more) different materials bonded together, one forming a 'matrix' in which are embedded fibers or particles that increase the strength and stiffness of the matrix material (see Figure 2.2). Concrete is a composite in which particles of stone add strength, with a further increase in strength provided by steel reinforcing rods. Vehicle tires consist of rubber reinforced with woven cords. Plastics are reinforced with glass, carbon and other fibers. The fibers may be unidirectional, woven or random chopped. Metals, carbon and ceramics are also used as matrix materials. So-called 'whiskers', which are single crystals of silicon carbide, silicon nitride, sapphire, etc., give extremely high strength.

Elastic Modulus of a Composite (continuous fibers in direction of load)
Let:

TABLE 2.11 Electrical Resistivity and Temperature Coefficients
of Resistivity of Pure Metals

Metal	mp, °C	Density at 20°C, $kg/m^3 \times 10^{-3}$	Resistivity at 20°C, $\Omega m \times 10^8$	Temp. coeff. of resistivity, 0–100°C,[a] $K^{-1} \times 10^3$
Aluminum	660.1	2.70	2.67	4.5
Antimony	630.5	6.68	40.1	5.1
Barium	729	3.5	60[b]	...
Beryllium	1287	1.848	3.3	9.0
Bismuth	271	9.80	117	4.6
Cadmium	320.9	8.64	7.3	4.3
Calcium	839	1.54	3.7	4.57
Cerium	798	6.75	85.4	8.7
Cesium	28.5	1.87	20	4.8
Chromium	1860	7.1	13.2	2.14
Cobalt	1492	8.9	6.34	6.6
Copper	1083.4	8.96	1.694[c]	4.3
Gallium	29.7	5.91		...
Germanium	937	5.32	...	...
Gold	1063	19.3	2.20	4.0
Hafnium	2227	13.1	32.2	4.4
Indium	156.4	7.3	8.8	5.2
Iridium	2454	22.4	5.1	4.5
Iron	1536	7.87	10.1	6.5
Lead	327.4	11.68	20.6	4.2
Lithium	181	0.534	9.29	4.35
Magnesium	649	1.74	4.2	4.25
Manganese	1244	7.4	160(α)	...
Mercury	−38.87	13.546	95.9	1.0
Molybdenum	2615	10.2	5.7	4.35
Nickel	1455	8.9	6.9	6.8
Niobium	2467	8.6	16.0	2.6
Osmium	3030	22.5	8.8	4.1
Palladium	1552	12.0	10.8	4.2
Platinum	1769	21.45	10.58	3.92
Potassium	63.2	0.86	6.8	5.7
Radium	700	5		
Rhenium	3180	21.0	18.7	4.5
Rhodium	1966	12.4	4.7	4.4
Rubidium	38.8	1.53	12.1	4.8
Ruthenium	2310	12.2	7.7	4.1
Silicon	1412	2.34	...	...
Silver	960.8	10.5	1.63	4.1
Sodium	97.8	0.97	4.7	5.5
Strontium	770	2.6	23[b]	...
Tantalum	2980	16.6	13.5	3.5
Tellurium	450	6.24	1.6×10^{5b}	...
Thallium	304	11.85	16.6	5.2
Thorium	1755	11.5	14	4.0
Tin	231.9	7.3	12.6	4.6

TABLE 2.11 Electrical Resistivity and Temperature Coefficients of Resistivity of Pure Metals (*Continued*)

Metal	mp, °C	Density at 20°C, $kg/m^3 \times 10^{-3}$	Resistivity at 20°C, $\Omega m \times 10^8$	Temp. coeff. of resistivity, 0–100°C,[a] $K^{-1} \times 10^3$
Titanium	1667	4.5	54	3.8
Tungsten	3400	19.3	5.4	4.8
Uranium	1132	19.05(α), 18.89(β)	27	3.4
Vanadium	1902	6.1	19.6	3.9
Zinc	419.5	7.14	5.96	4.2
Zirconium	1852	6.49	44	4.4

[a]By convention, the signs on orders of magnitude are reversed when they are moved from column to stub (leftmost column) or from column to column heading in a table; i.e., the first value in the temperature-coefficient-of-resistivity column is read $4.5 \times 10^{-3} K^{-1}$. The convention is not always observed by different authors, but it has been followed throughout this table.
[b]0°C
[c]17.4 | *a* axis, 8.1 | *b* axis, 54.3 | *c* axis.
Source: From Fink and Christiansen.[4]

TABLE 2.12 Electrical Resistivity of Some Alloys and Compounds

Material	Resistivity, $\Omega m \times 10^8$	Material	Resistivity, $\Omega m \times 10^8$
92.5 Ag–7.5 Cu	2		
60 Ag–40 Pd	23	TaN	135
97 Ag–3 Pt	3.5	ZrN	13.6
90 Pt–b Ir	25	TiN	21.7
96 Pt–4 W	36	VN	85.9
70 Pd–30 Ag	40	TaC	30
9O Au–10 Cu	10.8	ZrC	63.4
75 Au–25 Ag	10.8	SiC	100–200
78.5 Ni–20 Cr–1.5 Si	108.05	WC	12
71 Ni–29 Fe	19.95	$MoSi_2$	37
80 Ni–20 Cr	112.2	$CbSi_2$	6.5
Carbon steel 0.65% C	18	ZrB_2	9
Electrical Si sheet steels	18–52	LaB_6	15
Stainless steel, type 302	72	TiB_2	20
Type 316	74		

Source: From Fink and Christiansen.[4]

$$E_f = \text{modulus of fibers}$$

$$E_m = \text{modulus of matrix}$$

$$E_c = \text{modulus of composite}$$

TABLE 2.13   Electrical Resistivity of Insulating
Materials

Material	Volume resistivity, $M\Omega \cdot cm$
Asbestos board (ebonized)	10^7
Bakelite	5–30×10^{11}
Epoxy	10^{14}
Fluorocarbons:	
Fluorinated ethylene propylene	10^{18}
Polytetrafluoroethylene	10^{18}
Glass	17×10^9
Magnesium oxide	
Mica	10^{14}–10^{17}
Nylon	10^{14}–10^{17}
Neoprene	
Oils:	
Mineral	21×10^6
Paraffin	10^{15}
Paper	
Paper, treated	
Phenolic (glass fillcd)	10^{12}–10^{13}
Polyethylene	10^{15}–10^{18}
Polyimide	10^{16}–10^{17}
Polyvinyl chloride (flexible)	10^{11}–10^{15}
Porcelain	3×10^8
Rubber	10^{14}–10^{16}
Rubber (butyl)	10^{18}

Source: From Avallone and Baumeister.[1]

TABLE 2.14   Approximate Relations for Thermal Expansions

Let:
α = coefficient of linear expansion (°C)
β = coefficient of superficial expansion (°C)
γ = coefficient of cubical expansion (°C)
T = temperature change (°C)
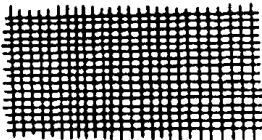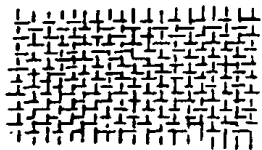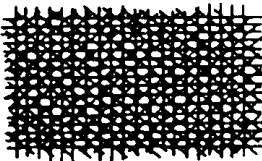
and:	Then:
L = initial length	$L' = L(1 + \alpha T)$
A = initial area	$A' = A(1 + \beta T)$
V = initial volume	$V' = V(1 + \gamma T)$
L' = final length	Note:
A' = final area	$\beta \approx 2\alpha$
V' = final volume	$\gamma \approx 3\alpha$

Source: Ref. 13 and 18.

TABLE 2.15 Typical Brinell Hardness Number (BHN) for Metals

Material	BHN
Soft brass	60
Mild steel	130
Annealed chisel steel	230
White cast iron	410
Nitrided surface	750

TABLE 2.16 Arrangement of Fibers in Composites

Type	Arrangement
1. Unidirectional Load taken in direction of fibers. Weak at right angles to fibers	
2. Bidirectional Takes equal load in both directions. Weaker since only half the fibers used in each direction	
3. Woven mat Similar to bidirectional type but easy to handle	
4. Multidirectional Load capacity much reduced but can take load in any direction in plane of fibers	
5. Random, chopped Low in strength but multidirectional. Has handling advantages	

Source: Refs. 13 and 19.

$$\alpha = \text{(cross-sectional area of fibers)}/\text{(total cross-sectional area)}$$

$$E_c = \alpha E_f + (1-\alpha)E_m$$

Plastics

Plastics are materials based on polymers to which other materials are added to give the desired properties. *Thermoplastic polymers* soften when heated and can be re-

TABLE 2.17 Properties of Plastics

Typical Physical Properties of Plastics

Properties of plastics	ρ (kg/m³)	Tensile strength (N/mm²)	Elongation (%)	E (GN/m²)	BHN	Machinability
Thermoplastics						
PVC rigid	1330	48	200	3.4	20	Excellent
Polystyrene	1300	48	3	3.4	25	Fair
PTFE	2100	13	100	0.3	—	Excellent
Polypropylene	1200	27	200–700	1.3	10	Excellent
Nylon	1160	60	90	2.4	10	Excellent
Cellulose nitrate	1350	48	40	1.4	10	Excellent
Cellulose acetate	1300	40	10–60	1.4	12	Excellent
Acrylic (Perspex)	1190	74	6	3.0	34	Excellent
Polythene (high density)	1450	20–30	20–100	0.7	2	Excellent
Thermoset ring plastics						
Epoxy resin (glass filled)	1600–2000	68–200	4	20	38	Good
Melamine formaldehyde (fabric filled)	1800–2000	60–90	—	7	38	Fair
Urea formaldehyde (cellulose filled)	1500	38	1	7	51	Fair
Phenol formaldehyde (mica filled)	1600	38	0.5	17	36	Good
Acetals (glass filled)	1600	58	2–7	7	27	Good

BHN = Brinell hardness number, ρ = density, E = Young's modulus.

Relative Properties of Plastics

Material	Tensile strength	Compressive strength	Machining properties	Chemical resistance
Thermoplastics				
Nylon	E	G	E	G
PTFE	F	G	E	O
Polypropylene	F	F	E	E
Polystyrene	E	G	F	F
Rigid PVC	E	G	E	G
Flexible PVC	F	P	P	G
Thermosetting plastics				
Epoxy resin (glass-fiber filled)	O	E	G	E
Formaldehyde (asbestos filled)	G	G	F	G
Phenol formaldehyde (Bakelite)	G	G	F	F
Polyester (glass-fiber filled)	E	G	G	F
Silicone (asbestos filled)	O	G	F	F

O = outstanding, E = excellent, G = good. F = fair, P = poor.
Tensile strength (typical): E = 55 N/mm²; P = 21 N/mm²
Compressive strength (typical): E = 210 N/mm²; P = 35 N/mm²
Source: Refs. 13, 15 and 19.
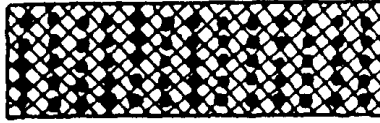
FIGURE 2.2   Matrix with fibers (Refs. 13 and 19)
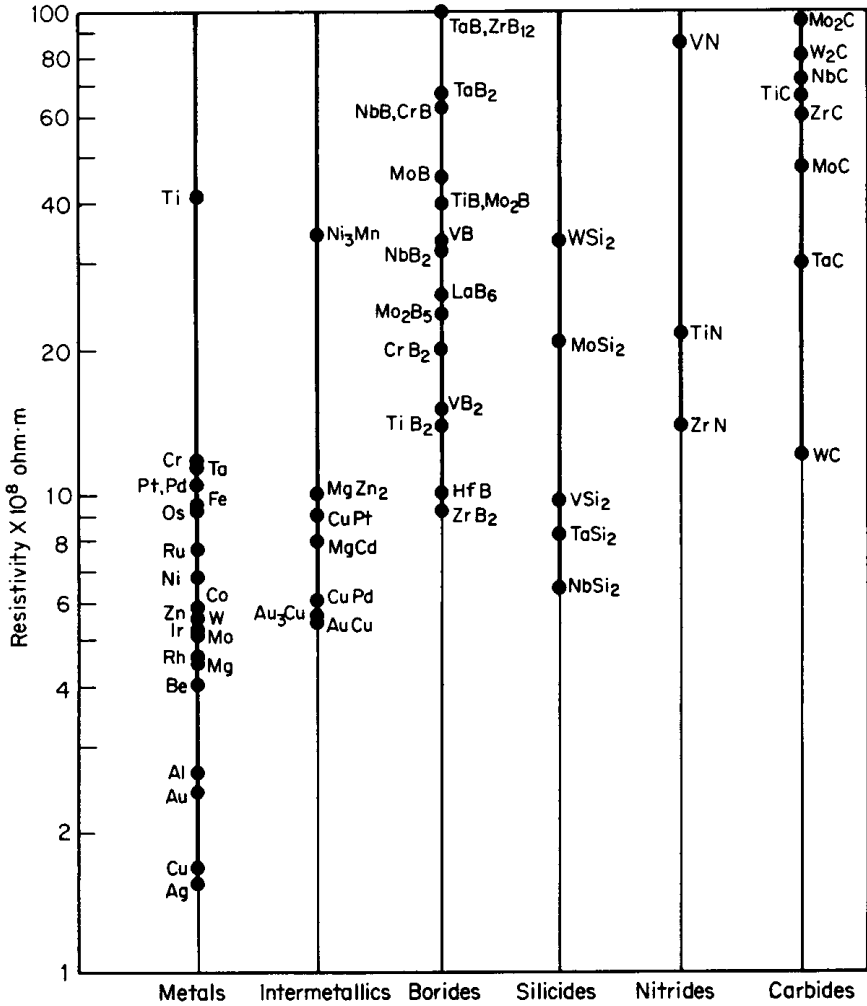
FIGURE 2.3   Electrical resistivity of various materials. (*From Fink and Christiansen.*[4])

shaped, the new shape being retained on cooling. *Thermosetting polymers* (or thermosets) cannot be softened and reshaped by heating. They are plastic at same stage of processing but finally set and cannot be resoftened.

Ceramics

Basic data for ceramics are given in Table 2.18.

Cermets

Typical cermets classification and applications chart data are given in Table 2.19.

Note. Cermets consist of powdered ceramic material in a matrix of metal, combining the hardness and strength of ceramic with the ductility of the metal to produce a hard, strong, yet tough, combination; the process involves compaction and sintering.

Wood

Basic data for wood are given in Table 2.20.

TABLE 2.18 Ceramics

Aluminum oxide (alumina)			
	% Al_2O_3		
	75	86–94	94–98	>98
Density (kg/m³)	3200	3300	3500	3700
Hardness (Moh scale)	8.5	9.0	9.0	9.0
Compressive strength (N/mm²)	1250	1750	1750	1750
Flexural strength (N/mm²)	270	290	350	380
Max. working temperature (°C)	800	1100	1500	1600

Silicon nitride		
	Type	
	Reaction sintered	Hot pressed
Density (kg/m³)	2300–2600	3120–3180
Open porosity (%)	18–28	0.1
Hardness (Moh scale)	9	9
Young's modulus (N/mm²)	160 000	290 000
Flexural strength (N/mm²)		
at 20°C	110–185	550–680
at 1200°C	210	350–480

Source: Refs. 13, 14, 15 and 19.

TABLE 2.19 Cermets

Typical cermets and applications		
Ceramic	Matrix	Applications
Tungsten carbide	Cobalt	Cutting-tool bits
Titanium carbide	Molybdenum, cobalt or tungsten	
Molybdenum carbide	Cobalt	Dies
Silicon carbide	Cobalt or chromium	
Aluminum oxide	Cobalt, iron or chromium	High-temperature components
Magnesium oxide	Magnesium, aluminum, cobalt, iron or nickel	Rocket and jet engine parts
Chromium oxide	Chromium	Disposable tool bits
Uranium oxide	Stainless steel	Nuclear fuel elements
Titanium boride	Cobalt or nickel	
Chromium boride	Nickel	Mainly as cutting tool tips
Molybdenum boride	Nickel or nickel-chromium alloy	

Source: Refs. 13, 14 and 19.

TABLE 2.20 Wood

Permitted stresses in structural timbers (N/mm²)							
Timber	Bending			Compression			
	Stress in extreme fiber		Horizontal shear stress	Stress parallel to grain		Stress perpendicular to grain	
	Outside location	Dry location	All locations	Outside location	Dry location	Outside location	Dry location
Oak	8.3	9.7	0.9	6.0	6.9	1.6	3.5
Douglas fir	7.6	9.0	0.6	6.0	6.9	1.6	2.1
Norway spruce	6.9	7.6	0.6	5.5	5.5	1.2	2.1

Mechanical properties of timbers							
Wood	Moisture (%)	Density, ρ (kg/m³)	Fiber stress at elastic limit (N/mm²)	Modulus of elasticity, E (N/mm²)	Modulus of rupture (N/mm²)	Compressive strength parallel to grain (N/mm²)	Shear strength (N/mm²)
Ash	15	657	60	10070	103	48	10
Beech	—	740	60–110	10350	—	27–54	8.3–14
Birch	9	710	85	15170	130	67	13
Elm, English	—	560	40–54	11790	—	17–32	8–11.3
Elm, Dutch	—	560	42–60	7720	—	18–32	7.2–10
Elm, Wych	—	690	65–100	7860	—	29–47	7.3–11.4
Fir, Douglas	6–9	530	45–73	10340–15170	71–97	49–74	7.4–8.8
Mahogany	15	545	60	8690	80	45	6.0
Oak	—	740	56–87	14550	—	27–50	8–12
Pine, Scots	—	530	41–83	8550–10340	—	21–42	5.2–9.7
Poplar	—	450	40–43	7240	—	20	4.8
Spruce, Norway	—	430	36–62	7380–8620	—	18–39	4.3–8
Sycamore	—	625	62–106	8970–13450	—	26–46	8.8–15

Source: Refs. 13, 19, 20 and 21.

TABLE 2.21 Chart of Materials for Special Requirements

High-strength metals	Solders
High carbon steel	Lead–tin
Tool steel, carbon or alloy	Pure tin
Spring steel	Lead–tin–cadmium
Nickel steel	Lead–tin–antimony
High tensile steel	Silver solder
Chrome–molybdenum steel	Aluminium solder
Nickel–chrome–molybdenum steel	**Metals with high electrical resistance**
18% nickel maraging steels	Advance (Cu, Ni)
Phosphor bronze	Constantan or Eureka (Cu, Ni)
Aluminium bronze	Manganin (Cu, Mn, Ni)
Beryllium copper	Nichrome (Ni, Cr)
High-strength aluminium alloys	Platinoid
High temperature metals	Mercury
Tungsten	Bismuth
Tantalum	**Good electrical insulators**
Molybdenum	Thermoplastics
Chromium	Thermosetting plastics
Vanadium	Glass
Titanium	Mica
Nimonic alloys	Transformer oil
Stellite	Quartz
Hastelloy	Ceramics
Inconel	Soft natural and synthetic rubber
Stainless steel	Hard rubber
Nichrome	Silicone rubber
Heat-resisting alloy steels	Shellac
Coating metals	Paxolin
Copper	Tufnol
Cadmium	Ebonite
Chromium	Insulating papers, silks, etc.
Nickel	Gases
Gold	**Good conductors of electricity**
Silver	Silver
Platinum	Copper
Tin	Gold
Zinc	Aluminium
Brass	Magnesium
Bronze	Brass
Lead	Copper
Corrosion-resistant metals	Phosphor bronze
Stainless steels (especially austenitic)	Beryllium copper
Cupronickel	**Permanent-magnet materials**
Monel	Alnico I
Titanium and alloys	Alnico II
Pure aluminium	Alnico V
Nickel	Cobalt steel 35%
Lead	Tungsten steel 6%
Tin	Chrome steel 3%
Meehanite (cast iron)	Electrical sheet steel 1% Si
	Barium ferrite

TABLE 2.21 Chart of Materials for Special Requirements (*Continued*)

Good conductors of heat	Low-loss magnetic materials
Aluminium	Pure iron
Bronze	Permalloy
Copper	Mumetal
Duralumin	Silicon sheet steel 4.5%
Gold	Silicon sheet steel 1%
Magnesium	Permendur
Molybdenum	Annealed cast iron
Silver	Ferrite
Tungsten	
Zinc	

Good heat insulators	Sound-absorbing materials
Asbestos cloth	Acoustic tiles and boards:
Balsa wood	Cellulose
Calcium silicate	Mineral
Compressed straw	Acoustic plasters
Cork	Blanket materials:
Cotton wool	Rock wool
Diatomaceous earth	Glass wool
Diatomite	Wood wool
Expanded polystyrene	Perforated panels with absorbent backing
Felt	Suspended absorbers
Glass fiber and foam	

	Lubricants
Glass wool	Mineral oils
Hardboard	Vegetable oils
Insulating wallboard	Mineral grease
Magnesia	Tallow
Mineral wool	Silicone oil
Plywood	Silicone grease
Polyurethane foam	Flaked graphite
Rock wool	Colloidal graphite
Rubber	Graphite grease
Sawdust	Molybdenum disulphide
Slag wool	Water
Urea formaldehyde foam	Gases
Wood	
Wood wool	

Semiconductors
Silicon
Germanium
Gallium arsenide
Gallium phosphide
Gallium arsenide phosphide
Cadmium sulphide
Zinc sulphide
Indium antimonide

Source: Refs. 1, 13–19.

SPECIAL REQUIREMENTS

Materials for special requirements are listed in Table 2.21.

*REFERENCES**

1. E. A. Avallone and T. Baumeister III (eds.), *Mark's Standard Handbook for Mechanical Engineers,* 9th ed., McGraw-Hill, New York, 1987.

2. J. H. Lienhard, *A Heat Transfer Textbook,* Prentice-Hall, Englewood Cliffs, N.J., 1981.

3. W. M. Rohsenow, J. P. Hartnett, and E. N. Ganic (eds.), *Handbook of Heat Transfer Fundamentals,* McGraw-Hill, New York, 1985, chap. 3.

4. D. G. Fink and D. Christiansen (eds.), *Electronics Engineers' Handbook,* 3d ed., McGraw-Hill, New York, 1989.

5. E. R. G. Eckert and R. M. Drake, Jr., *Analysis of Heat and Mass Transfer,* McGraw-Hill, New York, 1972.

6. H. Gartmann (ed.), *Dc Laval Engineering Handbook,* 3d ed., McGraw-Hill, New York, 1970.

7. T. F. Irvine, Jr., and J. P. Hartnett (eds.), *Steam and Air Tables in S.I. Units,* Hemisphere Publishing Corp/McGraw-Hill, Washington, D.C., 1976.

8. R. H. Perry, *Engineering Manual,* 3d ed., McGraw-Hill, New York, 1976.

9. R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Pet-my's Chemical Engineers' Handbook,* 6th cd., McGraw-Hill, New York, 1984.

10. K. Raznjevic, *Handbook of Thermodynamic Tables and Charts,* McGraw-Hill, New York, 1976.

11. R. C. Reid, J. M. Prausnitz, and T. K. Sherwood, *The Properties of Gases and Liquids,* 3d ed., McGraw-Hill, New York, 1977.

12. Y. S. Touloukian, *Thermophysical Properties of Matter,* vols. 1–6, 10, and 11, Purdue University, West Lafayette, Ind., 1970–1975.

13. J. Carvill, *Mechanical Engineer's Data Handbook,* Butterworth-Heinemann, 2001.

14. F. S. Merritt, *Standard Handbook for Civil Engineers*, 4th ed., McGraw-Hill, 1996.

15. R. A. Walsh, *McGraw-Hill Machining and Metalworking Handbook*, 2d ed., McGraw-Hill, 1999.

16. D. Christiansen, *Electronics Engineers' Handbook*, 4th ed., McGraw-Hill, 1997.

17. R. A. Corbitt, *Standard Handbook of Environmental Engineering*, 2d. ed., McGraw-Hill, 1999.

18. B. A. Grigorev and B. M. Zorin, *Handbook of Heat and Mass Transfer Experiments,* Energoizdat, 1982 (in Russian).

19. E. N. Ganic, Tehnicka hemija (Chemistry for Engineers—Class Notes) Dept. of Mechanical Engineering, University of Sarajevo, 1984–2002.

20. G. Giordano, Technologia Del Legno, Unione Tipografico, Torino, 1981.

21. F. Kollmann and W. A. Cote, Jr., Principle of Wood Science and Technology, Springer-Verlag, 1968.

---

*Those references listed above but not cited in the text were used for comparison between different data sources, clarification, clarity of presentation, and, most importantly, reader's convenience when further interest in subject exists.

CHAPTER 3
ENGINEERING MATHEMATICS

ALGEBRA

1. Basic Laws

Commutative law:
$$a + b = b + a \quad ab = ba$$

Associative law:
$$a + (b + c) = (a + b) + c \quad a(bc) = (ab)c$$

Distributive law:
$$a(b + c) = ab + ac$$

2. Sums of Numbers

The sum of the first n numbers:
$$\sum_{1}^{n} (n) = \frac{n(n + 1)}{2}$$

The sum of the squares of the first n numbers:
$$\sum_{1}^{n} (n^2) = \frac{n(n + 1)(2n + 1)}{6}$$

The sum of the cubes of the first n numbers:
$$\sum_{1}^{n} (n^3) = \frac{n^2(n + 1)^2}{4}$$

3. Progressions

Arithmetic Progression
$$a, a + d, a + 2d, a + 3d, \ldots$$

where
$$a = \text{first term}$$
$$d = \text{common difference}$$
$$n = \text{number of terms}$$

3.1

$$S = \text{sum of } n \text{ terms}$$
$$l = \text{last term}$$
$$l = a + (n - 1)d$$
$$S = (n/2)(a + l)$$
$$(a + b)/2 = \text{arithmetic mean of } a \text{ and } b$$

Geometric Progression

$$a, ar, ar^2, ar^3, \ldots$$

where a = first term
$\quad r$ = common ratio
$\quad n$ = number of terms
$\quad S$ = sum of n terms
$\quad l$ = last term

$$l = ar^{n-1}$$

$$S = a\frac{r^n - 1}{r - 1} = \frac{rl - a}{r - 1}$$

$$S = \frac{a}{1 - r} \qquad \text{for } r^2 < 1 \text{ and } n = x$$

$$\sqrt{ab} = \text{geometric mean of } a \text{ and } b$$

4. Powers and Roots

$$a^x a^y = a^{x+y}$$

$$\frac{a^x}{a^y} = a^{x-y}$$

$$(ab)^x = a^x b^x$$

$$(a^x)^y = a^{xy}$$

$$a^0 = 1 \text{ if } a \neq 0$$

$$a^{-x} = 1/a^x$$

$$a^{x/y} = \sqrt[y]{a^x}$$

$$a^{1/y} = \sqrt[y]{a}$$

$$\sqrt[x]{ab} = \sqrt[x]{a}\sqrt[x]{b}$$

$$\sqrt[x]{a/b} = \sqrt[x]{a}/\sqrt[x]{b}$$

5. Binomial Theorem

$$(a \pm b)^n = a^n \pm na^{n-1}b + \frac{n(n - 1)}{2!} a^{n-2}b^2 \pm \frac{n(n - 1)(n - 2)}{3!} a^{n-3}b^3 + \cdots$$

$$+ (\pm 1)^m \frac{n(n - 1) \cdots (n - m + 1)}{m!} a^{n-m}b^m + \cdots$$

where $m! = 1 \cdot 2 \cdot 3 \cdots (m - 1)m$

The series is finite if n is a positive integer. If n is negative or fractional, the series is infinite and will converge for $|b| < |a|$ only.

6. Absolute Values. The numerical or absolute value of a number n is denoted by $|n|$ and represents the magnitude of the number without regard to algebraic sign. For example, $|-3| = |+3| = 3$.

7. Logarithms. Definition of a logarithm: If $N = b^x$, the exponent x is the logarithm of N to the base b and is written $x = \log_b N$. The number b must be positive, finite, and different from unity. The base of common, or briggsian, logarithms is 10. The base of natural, napierian, or hyperbolic logarithms is 2.71 828 18 . . . , denoted by e.

Laws of Logarithms

$$\log_b MN = \log_b M + \log_b N \qquad \log_b 1 = 0$$

$$\log_b \frac{M}{N} = \log_b M - \log_b N \qquad \log_b b = 1$$

$$\log_b N^m = m \log_b N \qquad \log_b 0 = +\infty, \; 0 < b < 1$$

$$\log_b \sqrt[r]{N^m} = m/r \log_b N \qquad \log_b 0 = -\infty, \; 1 < b < \infty$$

Important Constants

$$\log_{10} e = 0.434\ 294\ 481\ 9$$

$$\log_{10} x = 0.434\ 3 \log_e x = 0.434\ 3 \ln x$$

$$\ln 10 = \log_e 10 = 2.302\ 585\ 093\ 0$$

$$\ln x = \log_e x = 2.302\ 6 \log_{10} x$$

8. Permutations. The number of possible permutations or arrangments of n different elements is $1 \times 2 \times 3 \times \cdots \times n = n!$ (read: "n factorial").

If among the n elements there are p equal ones of one sort, q equal ones of another sort, r equal ones of a third sort, etc., then the number of possible permutations is $(n!)/(p! \times q! \times r! \times \cdots)$, where $p + q + r + \cdots = n$.

9. Combinations. The number of possible combinations or groups of n elements taken r at a time (without repetition of any element within any group) is $[n(n - 1)(n - 2)(n - 3) \cdots (n - r + 1)]/(r!) = (n)_r$. If repetitions are allowed, so that a group, for example, may contain as many as r equal elements, then the number of combinations of n elements taken r at a time is $(m)_r$, where $m_r = n + r - 1$.

Note that $(n)_1 + (n)_2 + \cdots + (n)_n = 2^n - 1$.

10. Equations in One Unknown*

Roots of an Equation. An equation containing a single variable x will in general be true for some values of x and false for other values. Any value of x for which the equation is true is called a *root of the equation.* To *solve* an equation means to find all its roots. Any root of an equation, when substituted therein for x, will *satisfy* the equation. An equation which is true for all values of x, like $(x + 1)^2 = x^2 + 2x + 1$, is called an *identity* [often written $(x + 1)^2 \equiv x^2 + 2x + 1$].

Types of Equations

1. Algebraic equations
   Of the first degree (linear), e.g., $2x + 6 = 0$ (root: $x = -3$)
   Of the second degree (quadratic), e.g., $x^2 - 2x - 3 = 0$ (roots: $-1, 3$)
   Of the third degree (cubic), e.g., $x^3 - 6x^2 + 5x + 12 = 0$ (roots: $-1, 3, 4$)
2. Transcendental equations
   Exponential equations, e.g., $2^x = 32$ (root: $x = 5$); $2^x = -32$ (no real root)
   Trigonometric equations, e.g., $10 \sin x - \sin 3x = 3$ (roots: $30°, 150°$).

Equations of First Degree. These are linear equations. *Solution:* Collect all the terms involving x on one side of the equation, thus: $ax = b$, where a and b are known numbers. Then divide through by the coefficient of x, obtaining $x = b/a$ as the root.

Equations of Second Degree. These are quadratic equations. *Solution:* Throw the equation into the standard form $ax^2 + bx + c = 0$. Then the two roots are

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \qquad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

The roots are real and distinct, coincident, or imaginary, according as $b^2 - 4ac$ is positive, zero, or negative. The sum of the roots is $x_1 + x_2 = -b/a$; the product of the roots is $x_1 x_2 = c/a$.

Graphical solution: Write the equation in the form $x^2 = px + q$, and plot the parabola $y_1 = x^2$, and the straight line $y_2 = px + q$. The abscissas of the points of intersection will be the roots of the equation. If the line does not cut the parabola, the roots are imaginary.

Equations of Third Degree. The general cubic equation $x^3 + bx^2 + cx + d = 0$ is reducible (substitute $x = y - b/3$) to the form $y^3 + vy + w = 0$, where $v = (3c - b^2)/3$ and $w = (2b^3 - 9bc + 27d)/27$. The roots of the reduced equation are

---

*This section is in part taken from Ref. 1, *Marks' Standard Handbook for Mechanical Engineers,* 9th ed., by E. A. Avallone and T. Baumeister III (cds.). Copyright © 1987. Used by permission of McGraw-Hill, Inc. All rights reserved.

$$x_1 = A + B \quad \text{and} \quad x_{2,3} = -\frac{1}{2}(A + B) \pm \frac{\sqrt{3}}{2}(A - B)i$$

where

$$A = \left[-\frac{w}{2} + \left(\frac{w^2}{4} + \frac{v^3}{27} \right)^{1/2} \right]^{1/3}$$

$$B = \left[-\frac{w}{2} - \left(\frac{w^2}{4} + \frac{v^2}{27} \right)^{1/2} \right]^{1/2}$$

If $(w^2/4 + v^3/27) > 0$, there are one real root and two conjugate complex roots.

11. Determinants*

Evaluation of Determinants. Of the second order:

$$\begin{vmatrix} a_1 b_1 \\ a_2 b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1$$

Of the third order:

$$\begin{vmatrix} a_1 b_1 c_1 \\ a_2 b_2 c_2 \\ a_3 b_3 c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 c_2 \\ b_3 c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 c_1 \\ b_3 c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 c_1 \\ b_2 c_2 \end{vmatrix}$$

$$= a_1(b_2 c_3 - b_3 c_2) - a_2(b_1 c_3 - b_3 c_1) + a_3(b_1 c_2 - b_2 c_1)$$

Of the fourth order:

$$\begin{vmatrix} a_1 b_1 c_1 d_1 \\ a_2 b_2 c_2 d_2 \\ a_3 b_3 c_3 d_3 \\ a_4 b_4 c_4 d_4 \end{vmatrix} = a_1 \begin{vmatrix} b_2 c_2 d_2 \\ b_3 c_3 d_3 \\ b_4 c_4 d_4 \end{vmatrix} - a_2 \begin{vmatrix} b_1 c_1 d_1 \\ b_3 c_3 d_3 \\ b_4 c_4 d_4 \end{vmatrix} + a_3 \begin{vmatrix} b_1 c_1 d_1 \\ b_2 c_2 d_2 \\ b_4 c_4 d_4 \end{vmatrix} - a_4 \begin{vmatrix} b_1 c_1 d_1 \\ b_2 c_2 d_2 \\ b_3 c_3 d_3 \end{vmatrix}$$

etc. In general, to evaluate a determinant of the nth order, take the elements of the first column with signs alternately plus and minus, and form the sum of the products obtained by multiplying each of these elements by its corresponding minor. The minor corresponding to any element a_1 is the determinant (of next lower order) obtained by striking out from the given determinant the row and column containing a_1.

Properties of Determinants

1. The columns may be changed to rows and the rows to columns:

$$\begin{vmatrix} a_1 b_1 c_1 \\ a_2 b_2 c_2 \\ a_3 b_3 c_3 \end{vmatrix} = \begin{vmatrix} a_1 a_2 a_3 \\ b_1 b_2 b_3 \\ c_1 c_2 c_3 \end{vmatrix}$$

2. Interchanging two adjacent columns changes the sign of the result.

3. If two columns are equal, the determinant is zero.

4. If the elements of one column are m times the elements of another column, the determinant is zero.

5. To multiply a determinant by any number m, multiply all the elements of any one column by m.

6.

$$\begin{vmatrix} a_1 + p_1 + q_1, b_1 c_1 \\ a_2 + p_2 + q_2, b_2 c_2 \\ a_3 + p_3 + q_3, b_3 c_3 \end{vmatrix} = \begin{vmatrix} a_1 b_1 c_1 \\ a_2 b_2 c_2 \\ a_3 b_3 c_3 \end{vmatrix} + \begin{vmatrix} p_1 b_1 c_1 \\ p_2 b_2 c_2 \\ p_3 b_3 c_3 \end{vmatrix} + \begin{vmatrix} q_1 b_1 c_1 \\ q_2 b_2 c_2 \\ q_3 b_3 c_3 \end{vmatrix}$$

7.

$$\begin{vmatrix} a_1 b_1 c_1 \\ a_2 b_2 c_2 \\ a_3 b_3 c_3 \end{vmatrix} = \begin{vmatrix} a_1 + mb_1, b_1 c_1 \\ a_2 + mb_2, b_2 c_2 \\ a_3 + mb_3, b_3 c_3 \end{vmatrix}$$

Solution of Simultaneous Equations by Determinants.   If

$$a_1 x + b_1 y + c_1 z = p_1$$

$$a_2 x + b_2 y + c_2 z = p_2$$

$$a_3 x + b_3 y + c_3 z = p_3$$

where

$$D = \begin{vmatrix} a_1 b_1 c_1 \\ a_2 b_2 c_2 \\ a_3 b_3 c_3 \end{vmatrix} \neq 0$$

then

$$x = D_1 / D$$

$$y = D_2 / D$$

$$z = D_3 / D$$
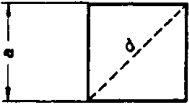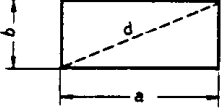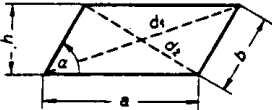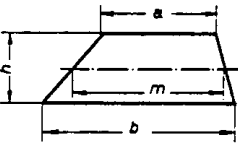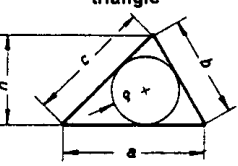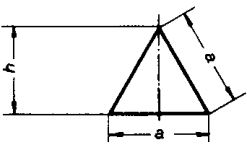
where

$$D_1 = \begin{vmatrix} p_1 b_1 c_1 \\ p_2 b_2 c_2 \\ p_3 b_3 c_3 \end{vmatrix}, D_2 = \begin{vmatrix} a_1 p_1 c_1 \\ a_2 p_2 c_2 \\ a_3 p_3 c_3 \end{vmatrix}, D_3 = \begin{vmatrix} a_1 b_1 p_1 \\ a_2 b_2 p_2 \\ a_3 b_3 p_3 \end{vmatrix}$$

Similarly for a larger (or smaller) number of equations.

*GEOMETRY**

Areas of Plane Figures

square

$A = a^2$

$a = \sqrt{A}$

$d = a\sqrt{2}$

rectangle

$A = a\,b$

$d = \sqrt{a^2 + b^2}$

parallelogram

$A = a\,h = a\,b\,\sin\alpha$

$d_1 = \sqrt{(a + h\,\cot\alpha)^2 + h^2}$

$d_2 = \sqrt{(a - h\,\cot\alpha)^2 + h^2}$

trapezoid

$A = \dfrac{a + b}{2}\,h = m\,h$

$m = \dfrac{a + b}{2}$

triangle

$A = \dfrac{a \cdot h}{2} = \rho\,s$

$= \sqrt{s(s-a)(s-b)(s-c)}$

$s = \dfrac{a + b + c}{2}$

equilateral triangle

$A = \dfrac{a^2}{4}\sqrt{3}$

$h = \dfrac{a}{2}\sqrt{3}$

$A = \frac{5}{8} r^2 \sqrt{10 + 2\sqrt{5}}$

$a = \frac{1}{2} r \sqrt{10 - 2\sqrt{5}}$

$\varrho = \frac{1}{4} r \sqrt{6 + 2\sqrt{5}}$

construction:
$\overline{AB} = 0.5\,r,\ \overline{BC} = \overline{BD},\ \overline{CD} = \overline{CE}$

pentagon
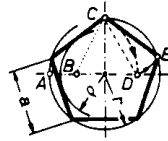
$A = \frac{3}{2} a^2 \sqrt{3}$

$d = 2\,a$

$\quad = \frac{2}{\sqrt{3}} s \quad \approx \quad 1.155\,s$

$s = \frac{\sqrt{3}}{2} d \quad \approx \quad 0.866\,d$

hexagon
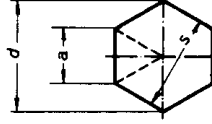
$A = 2\,a\,s \quad \approx \quad 0.83\,s^2$

$\quad = 2\,s\,\sqrt{d^2 - s^2}$

$a = s\,\tan 22.5° \approx 0.415\,s$

$s = d\,\cos 22.5° \approx 0.924\,d$

$d = \dfrac{s}{\cos 22.5°} \approx 1.083\,s$

octagon
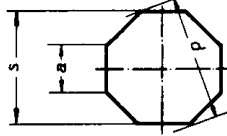
$A = A_1 + A_2 + A_3$

$\quad = \dfrac{a\,h_1 + b\,h_2 + b\,h_3}{2}$

polygon
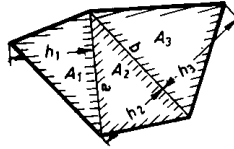
$A = \dfrac{\pi}{4} d^2 = \pi\,r^2$

$\quad \approx 0.785\,d^2$

$U = 2\,\pi\,r = \pi\,d$

circle
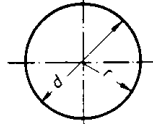
$A = \dfrac{\pi}{4} (D^2 - d^2)$

$\quad = \pi\,(d + b)b$

$b = \dfrac{D - d}{2}$

annulus
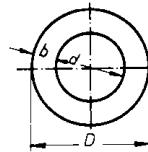
$A = \dfrac{\pi}{360°} r^2 \alpha = \dfrac{\hat{\alpha}}{2} r^2$

$\quad = \dfrac{b\,r}{2}$

$b = \dfrac{\pi}{180°} r\,\alpha$

$\hat{\alpha} = \dfrac{\pi}{180°} \alpha \quad (\hat{\alpha} = \alpha \ \text{in circular measure})$

sector of a circle
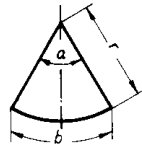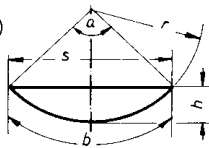
segment of a circle

$$s = 2r \sin\frac{a}{2}$$

$$A = \frac{h}{6s}(3h^2 + 4s^2) = \frac{r^2}{2}(\hat{a} - \sin a)$$

$$r = \frac{h}{2} + \frac{s^2}{8h}$$

$$h = r(1 - \cos\frac{a}{2}) = \frac{s}{2}\tan\frac{a}{4}$$
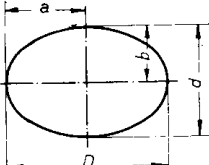
$\hat{a}$ see formula b 39

ellipse

$$A = \frac{\pi}{4}Dd = \pi a b$$

$$U \approx \pi\frac{D+d}{2}$$
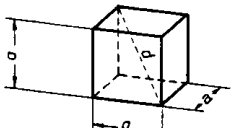
$$= \pi(a+b)\left[1 + \frac{1}{4}\lambda^2 + \frac{1}{64}\lambda^4 + \frac{1}{256}\lambda^6 + \frac{25}{16384}\lambda^8 + \dots\right].\text{ where } \lambda = \frac{a-b}{a+b}$$

Volumes and areas of solid bodies. For these, the following symbols are used: V = volume, A_o = surface area, A_m = generated surface.

cube

$$V = a^3$$

$$A_0 = 6a^2$$

$$d = \sqrt{3}\,a$$

cuboid
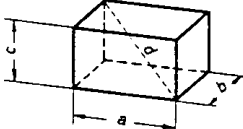
$$V = abc$$
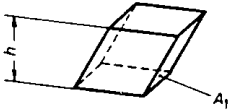
$$A_0 = 2(ab + ac + bc)$$

$$d = \sqrt{a^2 + b^2 + c^2}$$
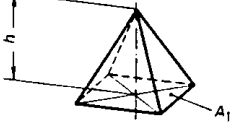
parallelepiped

$$V = A_1 h$$

(Cavalieri principle)
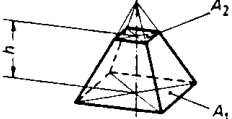
pyramid

$$V = \frac{A_1 h}{3}$$

frustum of pyramid

$$V = \frac{h}{3}(A_1 + A_2 + \sqrt{A_1 A_2})$$

$$\approx h\frac{A_1 + A_2}{2} \quad (\text{for } A_1 \approx A_2)$$

cylinder

$$V = \frac{\pi}{4} d^2 h$$

$$A_m = 2\pi r h$$

$$A_o = 2\pi r (r + h)$$
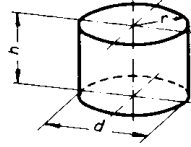
hollow cylinder

$$V = \frac{\pi}{4} h (D^2 - d^2)$$
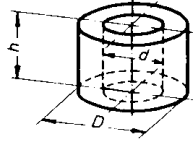
cone

$$V = \frac{\pi}{3} r^2 h$$
$$A_m = \pi r m$$
$$A_o = \pi r (r + m)$$
$$m = \sqrt{h^2 + r^2}$$
$$A_2 : A_1 = x^2 : h^2$$
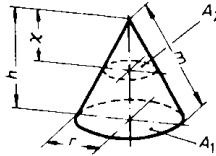
frustum of cone

$$V = \frac{\pi}{12} h (D^2 + Dd + d^2)$$

$$A_m = \frac{\pi}{2} m (D + d) = 2\pi p m$$

$$m = \sqrt{(\frac{D - d}{2})^2 + h^2}$$
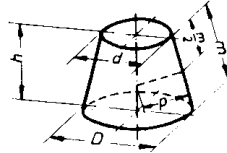
sphere

$$V = \frac{4}{3} \pi r^3 = \frac{1}{6} \pi d^3$$

$$\approx 4 \cdot 189 r^3$$

$$A_o = 4\pi r^2 = \pi d^2$$
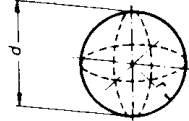
zone of a sphere

$$V = \frac{\pi}{6} h (3 a^2 + 3 b^2 + h^2)$$

$$A_m = 2\pi r h$$

$$A_o = \pi (2 r h + a^2 + b^2)$$
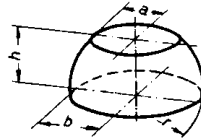
segment of a sphere

$$V = \frac{\pi}{6} h (\frac{3}{4} s^2 + h^2)$$

$$= \pi h^2 (r - \frac{h}{3})$$

$$A_m = 2\pi r h$$

$$= \frac{\pi}{4} (s^2 + 4 h^2)$$
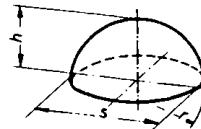
sector of a sphere

$$V = \frac{2}{3} \pi r^2 h$$

$$A_o = \frac{\pi}{2} r (4 h + s)$$
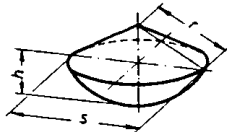
sphere with cylindrical boring

$$V = \frac{\pi}{6} h^3$$

$$A_o = 2 \pi h (R + r)$$
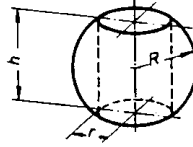
sphere with conical boring

$$V = \frac{2}{3} \pi r^2 h$$

$$A_o = 2 \pi r \left(h + \sqrt{r^2 - \frac{h^2}{4}} \right)$$
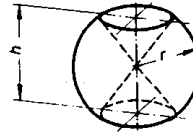
torus

$$V = \frac{\pi^2}{4} D d^2$$

$$A_o = \pi^2 D d$$
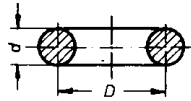
sliced cylinder

$$V = \frac{\pi}{4} d^2 h$$
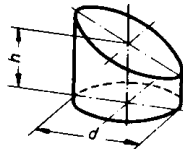
ungula

$$V = \frac{2}{3} r^2 h$$

$$A_m = 2 r h$$

$$A_o = A_m + \frac{\pi}{2} r^2 + \frac{\pi}{2} r \sqrt{r^2 + h^2}$$
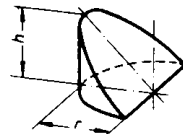
barrel

$$V \approx \frac{\pi}{12} h (2 D^2 + d^2)$$
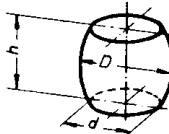
prismoid

$$V = \frac{h}{6} (A_1 + A_2 + 4 A)$$

This formula may be used for calculations involving solids shown in fig. C1...C3 and thus spheres and parts of spheres.
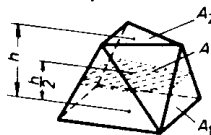
ANALYTIC GEOMETRY

1. Rectangular Coordinate Systems. The rectangular coordinate system in space is defined by three mutually perpendicular coordinate axes which intersect at the origin O as shown in Fig. 3.1. The position of a point $P(x, y, z)$ is given by the distances x, y, z from the coordinate planes *ZOY, XOZ,* and *XOY,* respectively.

2. Cylindrical Coordinate Systems. The position of any point P (r, θ, z) is given by the polar coordinates r and θ, the projection of P on the *XY* plane, and by z, the distance from the *XY* plane to the point (Fig. 3.2).

3. Spherical Coordinate Systems. The position of any point P (r, θ, ϕ) (Fig. 3.3) is given by the distance r $(= OP)$, the angle θ which is formed by the intersection of the X coordinate and the projection of OP on the *XY* plane, and the angle ϕ which is formed by OP and the coordinate z.
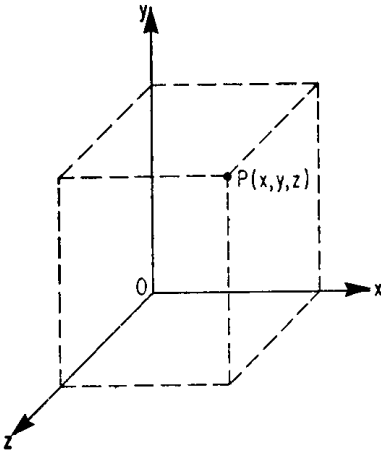
FIGURE 3.1 Rectangular coordinate system. (*From Rothbart.*[3])
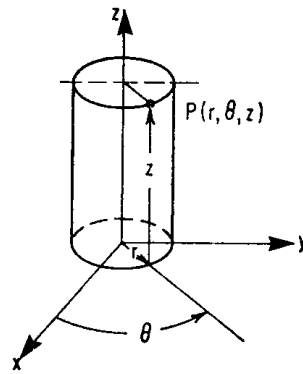
FIGURE 3.2 Cylindrical coordinate system. (*From Rothbart.*[3])
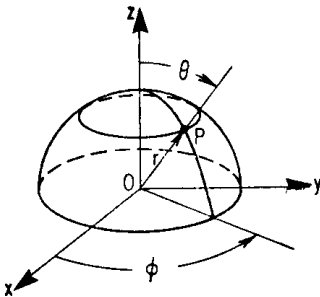
FIGURE 3.3 Spherical coordinate system. (*From Rothbart.*[3])
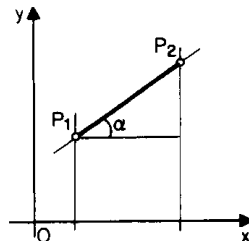
FIGURE 3.4

Relations between Coordinate Systems. Rectangular and cylindrical:

$$x = r \cos \theta \qquad y = r \sin \theta \qquad z = z \qquad \overline{OP} = r^2 + z^2$$

Rectangular and spherical:

$$x = r \sin \theta \cos \phi \qquad y = r \sin \theta \sin \phi \qquad z = r \cos \theta$$

4. Equations of a Straight Line*. Straight line distance (Fig. 3.4): Let $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$. Then, distance $P_1P_2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$; slope of $P_1P_2 = m = \tan \alpha = (y_2 - y_1)/(x_2 - x_1)$; coordinates of the midpoint are $x = 1/2(x_1 + x_2)$, $y = 1/2(y_1 + y_2)$.

Let m_1, m_2 be the slopes of two lines; then, if the lines are parallel, $m_1 = m_2$; if the lines are perpendicular to each other, $m_1 = -1/m_2$.

Various forms of a straight line equation:

Intercept form (Fig. 3.5)

$$\frac{x}{a} + \frac{y}{b} = 1$$

a, b = intercepts of the line on the axes.

Slope form (Fig. 3.6)

$$y = mx + b$$

$m = \tan \alpha$ = slope; b = intercept on the y axis.

Normal form (Fig. 3.7)

$$x \cos \beta + y \sin \beta = p$$

p = perpendicular from origin to line; β = angle from the x axis to p.
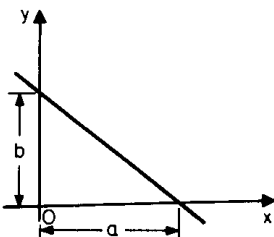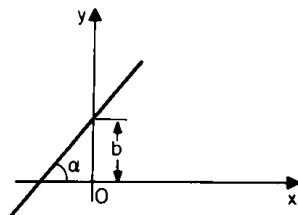
FIGURE 3.5   FIGURE 3.6

---

*This section is taken in part from Ref. I, *Marks' Standard Handbook for Mechanical Engineers,* 9th ed., by E. A. Avallone and T. Baumeister III (eds.). Copyright © 1987. Used by permission of McGraw-Hill. Inc. All rights reserved.
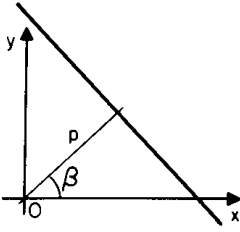
FIGURE 3.7
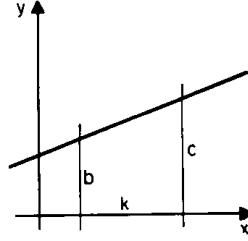
FIGURE 3.8

Parallel-intercept form (Fig. 3.8)

$$\frac{y - b}{c - b} = \frac{x}{k}$$

b, c = intercepts on two parallels at distance k apart.

General form

$$Ax + By + C = 0$$

Here $a = -C/A$, $b = -C/B$, $m = -A/B$, $\cos \beta = A/R$, $\sin \beta = B/R$, $p = -C/R$, where $R = \pm A^2 + B^2$ (sign to be so chosen that p is positive).

Line through (x_1, y_1) with slope m

$$y - y_1 = m(x - x_1)$$

Line through (x_1, y_1) and (x_2, y_2)

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1} (x - x_1)$$

Line parallel to x axis and to y axis, respectively

$$y = a \qquad x = b$$

Angles: If α = angle formed by the line with slope m_1 and the line with slope m_2, then

$$\tan \alpha = \frac{m_2 - m_1}{1 + m_2 m_1}$$

If parallel, $m_1 = m_2$. If perpendicular, $m_1 m_2 = -1$.

5. Circle.   Sec Fig. 3.9. Equations of a circle:

Center at the origin:

$$x^2 + y^2 = r^2$$
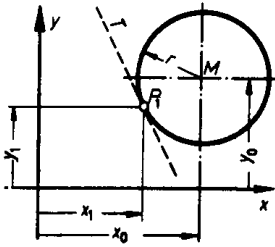
FIGURE 3.9 (*From Gieck.*[2])
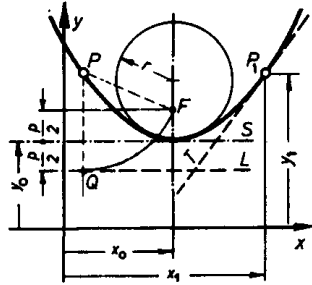
FIGURE 3.10 (*From Gieck.*[2])

Center at (x_o, y_o):

$$(x - x_o)^2 + (y - y_o) = r^2$$

General equation of a circle:

$$x^2 + y^2 + ax + by + c = 0$$

Radius of a circle:

$$r = (x_o^2 + y_o^2 - c)^{1/2}$$

Coordinates of the center M:

$$x_o = -\frac{a}{2} \quad y_o = -\frac{b}{2}$$

Tangent T at point P_1 (x_1, y_1):

$$y = r^2 - \frac{(x - x_o)(x_1 - x_o)}{(y_1 - y_2) + y_o}$$

6. Parabola. Sec Fig. 3.10. Equations of a parabola:

Vertex at the origin:

$$x^2 = 2py$$

Vertex at (x_o, y_o):

$$(x - x_o)^2 = 2p(y - y_o)$$

General equation of a parabola:

$$y = ax^2 + bx + c$$

Vertex radius:

$$r = p$$

Basic property:

$$\overline{PF} = \overline{PQ}$$

Tangent at point P_1 (x_1, y_1):

$$y = \frac{2(y_1 - y_o)(x - x_1)}{(x_1 - y_o)} + y_1$$

(F = focus, L = directrix, S = tangent at the vertex.)

7. **Hyperbola.** Sec Fig. 3.11. Equations of a hyperbola:

Point of intersection of asymptotes at the origin:

$$\frac{x^2}{a_2} - \frac{y^2}{b^2} - 1 = 0$$

Point of intersection of asymptotes at (x_o, y_o):

$$\frac{(x - x_o)^2}{a^2} - \frac{(y - y_o)^2}{b^2} - 1 = 0$$

General equation of a hyperbola:

$$ax^2 + by^2 + cx + dy + e = 0$$

Basic property:

$$\overline{F_2P} = \overline{F_1P} = 2a$$

Eccentricity:

$$e = (a^2 + b^2)^{1/2}$$

Gradient of asymptote:

$$\tan \alpha = m = \pm \frac{b}{a}$$
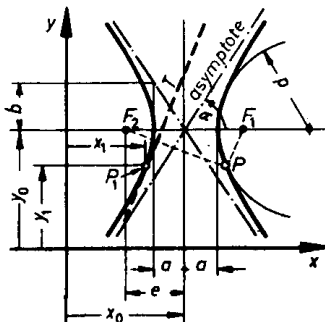
FIGURE 3.11 (*From Gieck.*[2])

Vertex radius:

$$p = \frac{b^2}{a}$$

Tangent T at point P_1 (x_1, y_1):

$$y = \frac{b_2}{a^2} \frac{(x_1 - x_o)(x - x_1)}{(y_1 - y_o)} + y_1$$

8. Ellipse.   See Fig. 3.12. Equation of an ellipse:

Point of interception of axes at the origin:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$$

Point of interception of axes at (x_o, y_o):

$$\frac{(x - x_o)^2}{a^2} + \frac{(y - y_o)^2}{b^2} - 1 = 0$$

Vertex radii:

$$r_N = \frac{b^2}{a} \qquad r_H = \frac{a^2}{b}$$

Eccentricity:

$$e = (a^2 - b^2)^{1/2}$$

Basic property:

$$\overline{F_1 P} + \overline{F_2 P} = 2a$$

Tangent T at $P_1(x_1, y_1)$:

$$y = -\frac{b^2}{a^2} \frac{(x_1 - x_o)(x - x_1)}{(y_1 - y_o)} + y_1$$

(F_1 and F_2 are focal points.)
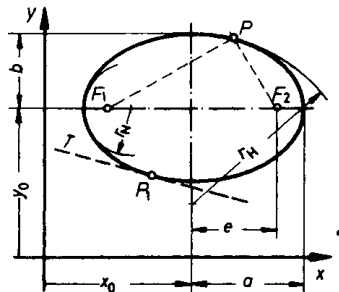
FIGURE 3.12   (*From Gieck.*[2])

9. Exponential curve. See Fig. 3.13.

Basic equation of an exponential curve: $y = a^x$. Here a is a positive constant, and x is a number.

All exponential curves pass through the point ($x = 0$, $y = 1$).

The derivative of the curve passing through this point with a gradient of 45° ($\tan \alpha = 1$) is equal to the curve itself. The constant a now becomes e ($e =$ base of the natural logarithm).

TRIGONOMETRY

1. Circular and Angular Measure of a Plane Angle. Circular measure is the ratio of the distance d measured along the arc to the radius r (Fig. 3.14). It is given in *radians* (rad), which have no dimensions:

$$\alpha = \frac{d}{r} \quad \text{rad}$$

Angular measure is obtained by dividing the angle α subtended at the center of a circle into 360 equal divisions known as *degrees* (°). A degree is divided into 60 minutes (unit: ′); a minute is divided into 60 seconds (unit: ″). Relationships between the more important circular and angular measures are shown in Table 3.1.
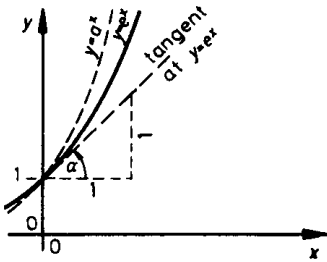
FIGURE 3.13   (*From Gieck.*[2])
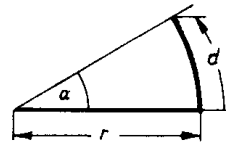
FIGURE 3.14   (*From Gieck.*[2])

TABLE 3.1   Relationships between Circular and Angular Measures

					$360° = 2\pi$ radians					
	or				$1 \text{ rad} = 57 \cdot 2958°$					
degrees	0°	30°	45°	60°	75°	90°	180°	270°	360°
radians	0	$\dfrac{\pi}{6}$	$\dfrac{\pi}{4}$	$\dfrac{\pi}{3}$	$\dfrac{5}{12}\pi$	$\dfrac{\pi}{2}$	π	$\dfrac{3}{2}\pi$	2π
	0	$0 \cdot 52$	$0 \cdot 79$	$1 \cdot 05$	$1 \cdot 31$	$1 \cdot 57$	$3 \cdot 14$	$4 \cdot 71$	$6 \cdot 28$

Source: From Gieck.[2]

2. Basic Relations between Trigonometric Functions. The principal trigonometric functions of the angle α (see right-angle triangle in Fig. 3.15) are:

$$\sin \alpha = \frac{\text{opposite}}{\text{hypotenuse}} = \frac{a}{c} \qquad \tan \alpha = \frac{\text{opposite}}{\text{adjacent}} = \frac{a}{b}$$

$$\cos \alpha = \frac{\text{adjacent}}{\text{hypotenuse}} = \frac{b}{c} \qquad \cot \alpha = \frac{\text{adjacent}}{\text{opposite}} = \frac{b}{a}$$

Values of the trigonometric functions for the more important angles are shown in Table 3.2. Trigonometric conversions are shown in Table 3.3. The relationships between sine and cosine functions are shown in Fig. 3.16, and those between tangent, sine, cotangent, and cosine functions are shown in Fig. 3.17.

3. Hyperbolic Functions. The hyperbolic functions are certain combinations of exponentials e^x and e^{-x}:

$$\cosh x = \frac{e^x + e^{-x}}{2}$$

$$\sinh x = \frac{e^x - e^{-x}}{2}$$

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$\coth x = \frac{e^x + e^{-x}}{e^x - e^{-x}} = \frac{1}{\tanh x} = \frac{\cosh x}{\sinh x}$$

$$\operatorname{sech} x = \frac{1}{\cosh x} = \frac{2}{e^x + e^{-x}}$$

$$\operatorname{csch} x = \frac{1}{\sinh x} = \frac{2}{e^x - e^{-x}}$$
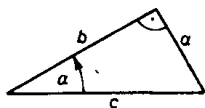
FIGURE 3.15 (*From Gieck.*[2])

TABLE 3.2 Values of Functions for the More Important Angles

angle α	0°	30°	45°	60°	75°	90°	180°	270°	360°
$\sin \alpha$	0	0,500	0,707	0,866	0,966	1	0	−1	0
$\cos \alpha$	1	0,866	0,707	0,505	0,259	0	−1	0	1
$\tan \alpha$	0	0,577	1,000	1,732	3,732	∞	0	∞	0
$\cot \alpha$	∞	1,732	1,000	0,577	0,268	0	∞	0	∞

Source: From Gieck.[2]

TABLE 3.3 Trigonometric Conversions

$$\sin^2 \alpha + \cos^2 \alpha = 1$$

$$1 + \tan^2\alpha = \frac{1}{\cos^2 \alpha}$$

$$\tan \alpha \cdot \cot \alpha = 1$$

$$1 + \cot^2 \alpha = \frac{1}{\sin^2 \alpha}$$

$$\sin(\alpha \pm \beta) = \sin \alpha \cos \beta \pm \cos \alpha \sin \beta$$

$$\cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta$$

$$\tan(\alpha \pm \beta) = \frac{\tan \alpha \pm \tan \beta}{1 \mp \tan \alpha \tan \beta};$$

$$\cot(\alpha \pm \beta) = \frac{\cot \alpha \cot \beta \mp 1}{\pm \cot \alpha + \cot \beta}$$

$$\sin \alpha + \sin \beta = 2 \sin \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}$$

$$\sin \alpha - \sin \beta = 2 \cos \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}$$

$$\cos \alpha + \cos \beta = 2 \cos \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}$$

$$\cos \alpha - \cos \beta = -2 \sin \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2};$$

$$\tan \alpha \pm \tan \beta = \frac{\sin(\alpha \pm \beta)}{\cos \alpha \cos \beta}$$

$$\cot \alpha \pm \cot \beta = \frac{\sin(\beta \pm \alpha)}{\sin \alpha \sin \beta}$$

$$\sin \alpha \cdot \cos \beta = \frac{1}{2} \sin(\alpha + \beta) + \frac{1}{2} \sin(\alpha - \beta)$$

$$\cos \alpha \cdot \cos \beta = \frac{1}{2} \cos(\alpha + \beta) + \frac{1}{2} \cos(\alpha - \beta)$$

$$\sin \alpha \cdot \sin \beta = \frac{1}{2} \cos(\alpha - \beta) - \frac{1}{2} \cos(\alpha + \beta)$$

$$\tan \alpha \cdot \tan \beta = \frac{\tan \alpha + \tan \beta}{\cot \alpha + \cot \beta} = -\frac{\tan \alpha - \tan \beta}{\cot \alpha - \cot \beta}$$

$$\cot \alpha \cdot \cot \beta = \frac{\cot \alpha + \cot \beta}{\tan \alpha + \tan \beta} = -\frac{\cot \alpha - \cot \beta}{\tan \alpha - \tan \beta}$$

$$\cot \alpha \cdot \tan \beta = \frac{\cot \alpha + \tan \beta}{\tan \alpha + \cot \beta} = -\frac{\cot \alpha - \tan \beta}{\tan \alpha - \cot \beta}$$

Source: From Gieck.[2]

Basic equations

Sine function $\qquad y = A \sin(ka - \varphi)$

Cosine function $\qquad y = A \cos(ka - \varphi)$

	with an amplitude of	$A = 1$	and $k = 1$
sine curve		$A = 1 \cdot 5$	and $k = 2$
sine curve		$A = 1$	and $k = 1$
cosine curve			
or sine curve with a phase shift of		$\varphi = -\dfrac{\pi}{2}$	
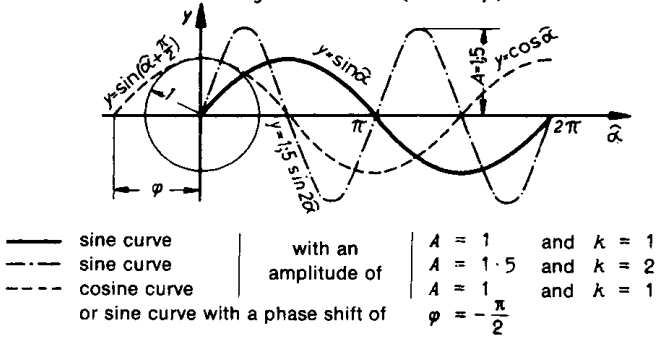
FIGURE 3.16 Relations between sine and cosine functions. (*From Gieck.*[2])
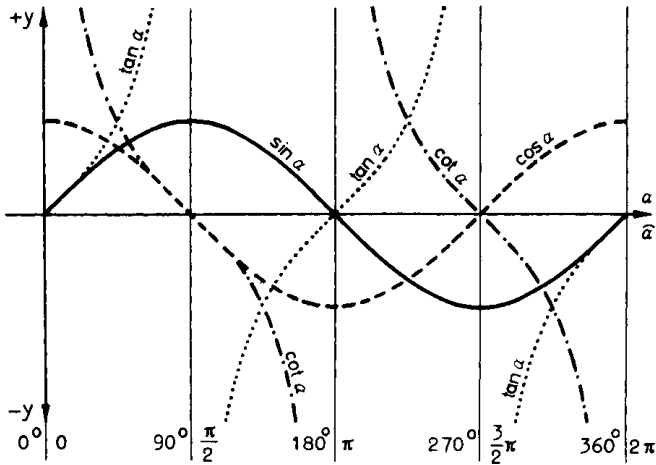
FIGURE 3.17 Relations between trigonometric functions. (*From Gieck.*[2])

In regard to inverse hyperbolic functions: If $x = \sinh y$, then y is the inverse hyperbolic sine of x written

$$y = \sinh^{-1}x \quad \text{or} \quad \text{arcsinh } x \quad \text{etc.}$$

DIFFERENTIAL AND INTEGRAL CALCULUS

1. Differential Calculus*

Derivative. The derivative of a function $y = f(x)$ of a single variable x is

---

*This section taken partly from Ref. 3, *Mechanical Design and Systems Handbook,* 2d ed., by H. A. Rothbart. Copyright © 1985. Used by permission of McGraw-Hill. Inc. All rights reserved.

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = f'(x)$$

i.e., the derivative of the function y is the limit of the ratio of the increment of the function y to the increment of the independent variable x as the increment of x varies and approaches zero as a limit. The derivative at a point can also be shown to equal the slope of the tangent line to the curve at the same point; i.e., $dy/dx = \tan \alpha$, as shown in Fig. 3.18. The derivative of $f(x)$ is a function of x and may also be differentiated with respect to x. The first differentiation of the first derivative yields the second derivative of the function d^2y/dx^2 or $f''(x)$. Similarly, the third derivative d^3y/dx^3 or $f'''(x)$ of the function is the first derivative of d^2y/dx^2, and so on.

Conditions of Maxima and Minima. From Fig. 3.18 it is seen that the function $f(x)$ possesses a maximum value where the derivative is zero and the concavity is downward, and the function possesses a minimum value where the slope is zero and the curve has an upward concavity. If the function $f(x)$ is concave upward, the second derivative will have a positive value; if negative, the curve will be concave downward. If the second derivative equals zero at a point, that point is a point of inflection. More particularly, where the nature of the curve is not well known, Table 3.4 may be used to determine the significance of the derivatives.

The *partial derivative* of a function $u = u(x, y)$ of two variables, taken with respect to the variable x, is defined by
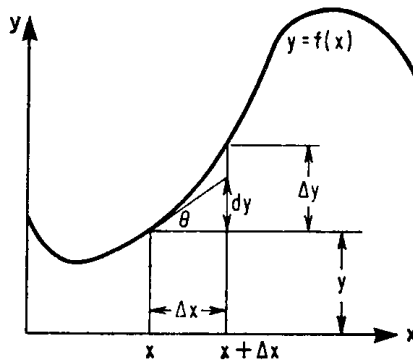
FIGURE 3.18 (*From Rothbart.*[3])

TABLE 3.4 Significance of the Derivative of $f(x)$

$f'(x)\big]_{x_0}$	$f''(x)\big]_{x_0}$	$f'''(x)\big]_{x_0}$	$f''''(x)\big]_{x_0}$	Comment
0	<0			x_0 a maximum point
0	>0			x_0 a minimum point
0	0	≠0		x_0 a point of inflection
0	0	0	<0	x_0 a maximum point
0	0	0	>0	x_0 a minimum point

Source: From Rothbart.[3]

$$\frac{\partial u}{\partial x} = \lim_{\Delta x \to 0} \frac{u(x + \Delta x, y) - u(x,y)}{\Delta x} = u_x$$

The partial derivatives are taken by differentiating with respect to one of the variables only, regarding the remaining variables as momentarily constant. Thus the partial derivative of u with respect to x of the function $u = 2xy^2$ is equal to $2y^2$; similarly $\partial u / \partial y = 4xy$. Higher derivatives are similarly formed. Thus

$$\frac{\partial^2 u}{\partial y^2} = u_{yy} = 4x \qquad \frac{\partial^2 u}{\partial x^2} = u_{xx} = 0$$

$$\frac{\partial^2 u}{\partial x \partial y} = u_{xy} = 4y \qquad \frac{\partial^2 u}{\partial y \partial x} = 4y$$

The order of differentiation in obtaining the mixed derivatives is immaterial if the derivatives are continuous.

Table 3.5 gives the conditions required to determine maxima, minima, and saddle points for $u = u(x,y)$.

Implicit Functions. If y is an implicit function of x, as, for example, in

$$xy - 5x^3y^2 = 3$$

and if it is difficult to solve the equation for y (or x), differentiate the terms as given, treating y as a function of x and solving for dy/dx. Thus, taking the derivatives of the above expression,

$$\frac{d}{dx} xy - \frac{d}{dx} 5x^3y^2 = \frac{d}{dx} 3$$

$$x \frac{dy}{dx} + y - 10x^3y \frac{dy}{dx} - y^2 15x^2 = 0$$

$$\frac{dy}{dx} = \frac{15^2x^2y^2 - y}{x - 10x^3y}$$

Another approach utilizes the relationship $dy/dx = -f_x/f_y$. Thus, in the foregoing example,

$$f_x = y - y^2 15x^2 \qquad f_y = x - 10x^3y$$

TABLE 3.5 Significance of the Derivative of $u(x, y)$

$\dfrac{\partial u}{\partial x}\Big]_{x_0,y_0}$	$\dfrac{\partial u}{\partial y}\Big]_{x_0,y_0}$	$\dfrac{\partial u^2}{\partial x^2}\Big]_{x_0,y_0}$	$\dfrac{\partial^2 u}{\partial y^2}\Big]_{x_0,y_0}$	Comments
0	0	<0	<0	$u(x,y)$ maximum at x_0, y_0 if $(u_{xx})(u_{yy}) - (u_{zy})^2 > 0$
0	0	>0	>0	$u(x,y)$ minimum at x_0, y_0 if $(u_{xx})(u_{yy}) - (u_{xy})^2 > 0$
0	0			Saddle point if u_{xx} and u_{yy} are of different sign

Source: From Rothbart.[3]
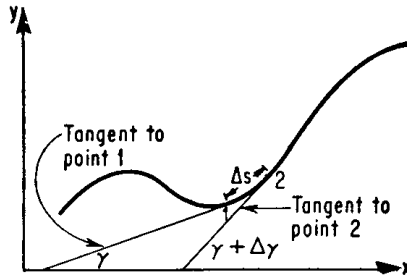
FIGURE 3.19 (*From Rothbart.*[3])

Curvature. The curvature K at a point P of a curve $y = f(x)$ (Fig. 3.19) is

$$K = \left| \lim_{\Delta s \to 0} \frac{\Delta \gamma}{\Delta s} \right| = \frac{d\gamma}{ds}$$

A working expression for the curvature is

$$K = \frac{y''}{[1 + (y'^2)]^{3/2}}$$

where the derivatives are evaluated at the point P. For a curve described in polar coordinates, the corresponding expression for the curvature is

$$K = \frac{\rho^2 + 2\pi'^2 - \rho\rho''}{(\rho^2 + \rho'^2)^{3/2}}$$

where ρ' and ρ'' represent the first and second derivatives of ρ with respect to θ.

The circle of curvature tangent, on its concave side, is curve $y = f(x)$ at P. The circle of curvature and the given curve have equal curvature at P. The circle of curvature is described by a radius of curvature R located at the center of curvature:

$$R = \frac{1}{K} = \frac{(1 + y'^2)^{3/2}}{y''}$$

The center of curvature is located at $(\overline{\alpha}, \overline{\beta})$ given by the following expressions:

$$\overline{\alpha} = x - \frac{y'(1 + y'^2)}{y''} \qquad \overline{\beta} = y + \frac{(1 + y'^2)}{y''}$$

where the expressions are evaluated at the point P. Table 3.6 shows derivatives of functions.

Differentials. The differential of a function is equal to the derivative of the function multiplied by the differential of the independent variable. Thus,

$$dy = \frac{dy}{dx} dx = f'(x)dx$$

TABLE 3.6 List of Derivatives

(Note: u, v, and w are functions of x)

$$\frac{d}{dx}(a) = 0 \qquad a = \text{constant}$$

$$\frac{d}{dx}(x) = 1$$

$$\frac{dy}{dx} = \frac{dy}{dv}\frac{dv}{dx} \qquad y = y(v)$$

$$\frac{d}{dx}(av) = a\frac{du}{dx}$$

$$\frac{dy}{dx} = \frac{1}{dx/dy} \qquad \text{if } \frac{dx}{dy} \neq 0$$

$$\frac{d}{dx}(\pm u \pm v \pm \cdots) = \pm\frac{du}{dx} \pm \frac{dv}{dx} + \cdots$$

$$\frac{d}{dx}(u^n) = nu^{n-1}\frac{du}{dx}$$

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx}$$

$$\frac{d}{dx}\frac{u}{v} = \frac{v\,du/dx - u\,dv/dx}{v^2}$$

$$\frac{d}{dx}(u^v) = vu^{v-1}\frac{du}{dx} + u^v \ln u\frac{dv}{dx}$$

$$\frac{d}{dx}(a^u) = a^u \ln a\frac{du}{dx}$$

$$\frac{d}{dx}(e^u) = e^u\frac{du}{dx}$$

$$\frac{d}{dx}(\ln u) = \frac{1}{u}\frac{du}{dx}$$

$$\frac{d}{dx}(\log_a u) = \frac{\log_a}{u}\frac{du}{dx}$$

$$\frac{d}{dx}(\sin u) = \cos u\frac{du}{dx}$$

$$\frac{d}{dx}(\cos u) = -\sin u\frac{du}{dx}$$

$$\frac{d}{dx}(\tan u) = \sec^2 u\frac{du}{dx}$$

TABLE 3.6 List of Derivatives (*Continued*)

$$\frac{d}{dx}(\csc u) = -\csc u \cot u \frac{du}{dx}$$

$$\frac{d}{dx}(\sec u) = \sec u \tan u \frac{du}{dx}$$

$$\frac{d}{dx}(\cot u) = -\csc^2 u \frac{du}{dx}$$

$$\frac{d}{dx}(\text{vers } u) = \sin u \frac{du}{dx}$$

$$\frac{d}{dx}\sin^{-1} u = \frac{1}{\sqrt{1 - u^2}} \frac{du}{dx} \qquad \frac{-\pi}{2} \leq \sin^{-1} u \leq \frac{\pi}{2}$$

$$\frac{d}{dx}\cos^{-1} u = \frac{1}{\sqrt{1 - u^2}} \frac{du}{dx} \qquad 0 \leq \cos^{-1} u \leq \pi$$

$$\frac{d}{dx}\tan^{-1} u = \frac{1}{1 + u^2} \frac{du}{dx}$$

$$\frac{d}{dx}\sinh^{-1} u = \frac{1}{\sqrt{u^2 + 1}} \frac{du}{dx}$$

$$\frac{d}{dx}\cosh^{-1} u = \frac{1}{\sqrt{u^2 - 1}} \frac{du}{dx} \qquad u > 1$$

$$\frac{d}{dx}\tanh^{-1} u = \frac{1}{1 - u^2} \frac{du}{dx}$$

$$\frac{d}{dx}\text{csch}^{-1} u = -\frac{1}{u \sqrt{u^2 + 1}} \frac{du}{dx}$$

$$\frac{d}{dx}\text{sech}^{-1} u = -\frac{1}{\sqrt{1 - u^2}} \frac{du}{dx} \qquad u > 0$$

$$\frac{d}{dx}\coth^{-1} u = \frac{1}{1 - u^2} \frac{du}{dx}$$

$$\frac{d}{dx}\csc^{-1} u = -\frac{1}{u \sqrt{u^2 - 1}} \frac{du}{dx} \qquad -\pi < \csc^{-1} u \leq -\frac{\pi}{2}, 0 < \csc^{-1} u \leq \frac{\pi}{2}$$

$$\frac{d}{dx}\sec^{-1} u = \frac{1}{u \sqrt{u^2 - 1}} \frac{du}{dx} \qquad -\pi \leq \sec^{-1} u < -\frac{\pi}{2}, 0 \leq \sec^{-1} u < \frac{\pi}{2}$$

$$\frac{d}{dx}\cot^{-1} u = \frac{-1}{1 + u^2} \frac{du}{dx}$$

$$\frac{d}{dx}\text{vers}^{-1} u = \frac{1}{\sqrt{2u - u^2}} \frac{du}{dx} \qquad 0 \leq \text{vers}^{-1} u \leq \pi$$

TABLE 3.6   List of Derivatives (*Continued*)

$$\frac{d}{dx} \sinh u = \cosh u \, \frac{du}{dx}$$

$$\frac{d}{dx} \cosh u = \sinh u \, \frac{du}{dx}$$

$$\frac{d}{dx} \tanh u = \operatorname{sech}^2 u \, \frac{du}{dx}$$

$$\frac{d}{dx} \operatorname{csch} u = -\operatorname{csch} u \coth u \, \frac{du}{dx}$$

$$\frac{d}{dx} \operatorname{sech} u = -\operatorname{sech} u \tanh u \, \frac{du}{dx}$$

$$\frac{d}{dx} \coth u = -\operatorname{csch}^2 u \, \frac{du}{dx}$$

Source: From Rothbart.[3]

The total differential dz of a function of two variables $z = z(x,y)$ is

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$$

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}$$

For a function of three variables $u = u(x,y,z)$,

$$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy + \frac{\partial u}{\partial z} dz$$

$$\frac{du}{dt} = \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial y} \frac{dy}{dt} + \frac{\partial u}{\partial z} \frac{dz}{dt}$$

x, y, z being functions of the independent variable t.

In the following relationships for differential of arc in rectangular coordinates, ds represents the differential of arc and β is the angle of the tangent drawn at the point in question; i.e., $\tan\beta = $ slope:

$$ds^2 = dx^2 + dy^2$$

$$ds = \left[1 + \left(\frac{dy}{dx}\right)^2\right]^{1/2} dx = \left[1 + \left(\frac{dx}{dy}\right)^2\right]^{1/2} dy$$

$$\frac{dx}{ds} = \cos\beta = \frac{1}{(1 + y'^2)^{1/2}} \qquad \frac{dy}{ds} = \sin\beta = \frac{y'}{(1 + y'^2)^{1/2}}$$

In the following relationships for differential of arc in polar coordinates for the function $\rho = \rho(\theta)$, ds represents the differential of arc:

$$ds = \sqrt{d\rho^2 + \rho^2 d\theta^2} \quad ds = \left[\rho^2 + \left(\frac{d\rho}{d\theta}\right)^2\right]^{1/2} d\theta$$

Indeterminate Forms. The function $f(x) = u(x)/v(x)$ has an indeterminate form $0/0$ at $x = a$ if $u(x)$ and $v(x)$ each approach zero as x approaches a through values greater than $a(x \to a +\ )$. The function $f(x)$ is not defined at $x = a$, and therefore it is often useful to assign a value to $f(a)$. L'Hôpital's rule is readily applied to indeterminacies of the form $0/0$:

$$\lim_{x \to a^+} \frac{u(x)}{v(x)} = \lim_{x \to a^+} \frac{u'(x)}{v'(x)}$$

L'Hôpital's rule may be reapplied as often as necessary, but it is important to remember to differentiate numerator and denominator separately. The above discussion is equally valid if $x \to a -$.

Other indeterminate forms such as ∞/∞, $0 \cdot \infty$, $\infty - \infty$, 0^0, ∞^0, and 1^∞ may also be evaluated by L'Hôpital's rule by changing their forms. For example, in order to evaluate the indeterminate form $0 \cdot \infty$, the function $u(x)v(x)$ may be written $u(x)/[1/v(x)]$ and the same technique employed as before.

2. Expansion in Series*. The range of values of x for which each of the series is convergent is stated at the right of the series. Arithmetic and geometrical series and the binomial theorem were given earlier, near the beginning of the chapter.

Exponential and Logarithmic Series

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \qquad [-\infty < x < +\infty]$$

$$a^x = e^{mx} = 1 + \frac{m}{1!} x + \frac{m^2}{2!} x^2 + \frac{m^2}{3!} x^3 + \cdots \qquad [a > 0,\ -\infty < x < +\infty]$$

where $m = \ln a = (2.302\ 6 \times \log_{10} a)$

$$\ln (1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} + \cdots \qquad [-1 < x < +1]$$

$$\ln (1 - x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \frac{x^5}{5} - \cdots \qquad [-1 < x < +1]$$

$$\ln \left(\frac{1 + x}{1 - x}\right) = 2 \left(x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \cdots\right) \qquad [-1 < x < +1]$$

$$\ln \left(\frac{x + 1}{x - 1}\right) = 2 \left(\frac{1}{x} + \frac{1}{3x^3} + \frac{1}{5x^5} + \frac{1}{7x^7} + \cdots\right) \qquad [x < -1 \text{ or } +1 < x]$$

$$\ln x = 2\left[\frac{x-1}{x+1} + \frac{1}{3}\left(\frac{x-1}{x+1}\right)^3 + \frac{1}{5}\left(\frac{x-1}{x+1}\right)^5 + \cdots\right] \qquad (0 < x < \infty]$$

$$\ln(a+x) = \ln a + 2\left[\frac{x}{2a+x} + \frac{1}{3}\left(\frac{x}{2a+x}\right)^3 + \frac{1}{5}\left(\frac{x}{2a+x}\right)^5 + \cdots\right]$$

$$[0 < a < +\infty, \ -a < x < +\infty]$$

Series for the Trigonometric Functions

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \qquad [-\infty < x < +\infty]$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots \qquad [-\infty < x < +\infty]$$

$$\tan x = x + \frac{x^3}{3} + \frac{2x^3}{15} + \frac{17x^7}{315} + \frac{62x^9}{2835} + \cdots \qquad [-\pi/2 < x < +\pi/2]$$

$$\cot x = \frac{1}{x} - \frac{x}{3} - \frac{x^3}{45} - \frac{2x^5}{945} - \frac{x^7}{4725} - \cdots \qquad [-\pi < x < +\pi]$$

$$\sin^{-1} y = y + \frac{y^3}{6} + \frac{3y^5}{40} + \frac{5y^7}{112} + \cdots \qquad [-1 \le y \le +1]$$

$$\tan^{-1} y = y - \frac{y^3}{3} + \frac{y^5}{5} - \frac{y^7}{7} + \cdots \qquad [-1 \le y \le +1]$$

$$\cos^{-1} y = \tfrac{1}{2}\pi - \sin^{-1} y \qquad \cot^{-1} y = \tfrac{1}{2}\pi - \tan^{-1} y$$

In these formulas, *all angles must be expressed in radians.* If D = the number of degrees in the angle, and x = its radian measure, then $x = 0.017\,453D$.

Series for the Hyperbolic Functions

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots \qquad [-\infty < x < \infty]$$

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots \qquad [-\infty < x + \infty]$$

$$\sinh^{-1} y = y - \frac{y^3}{6} + \frac{3y^5}{40} - \frac{5y^7}{112} + \cdots \qquad [-1 < y + 1]$$

$$\tanh^{-1} y = y + \frac{y^3}{3} + \frac{y^5}{5} + \frac{y^7}{7} + \cdots \qquad [-1 < y < +1]$$

General Formulas of Maclaurin and Taylor. If $f(x)$ and all its derivatives are continuous in the neighborhood of the point $x = 0$ (or $x = a$), then, for any value of x in this neighborhood, the function $f(x)$ may be expressed as a power series arranged according to ascending powers of x (or of $x - a$), as follows:

$$f(x) = f(0) + \frac{f'(0)}{1!} x + \frac{f''(0)}{2!} x^2$$

$$+ \frac{f'''(0)}{3!} x^3 + \cdots + \frac{f^{(n-1)}(0)}{(n-1)!} x^{n-1} + (P_n)x^n \text{ (Maclaurin)}$$

$$f(x) = f(a) + \frac{f'(a)}{1!} (x - a) + \frac{f''(a)}{2!} (x - a)^2 + \frac{f'''(a)}{3!} (x - a)^3 + \cdots$$

$$+ \frac{f^{(n-1)}(a)}{(n-1)!} (x - a)^{n-1} + (Q_n)(x - a)^n \text{ (Taylor)}$$

Here $(P_n)x^n$, or $(Q_n)(x - a)^n$, is called the *remainder term*; the values of the coefficients P_n and Q_n may be expressed as follows:

$$P_n = \frac{[f^{(n)}(sx)]}{n!} = \frac{(1 - t)^{n-1} f^{(n)}(tx)}{(n - 1)!}$$

$$Q_n = \frac{f^{(n)}[a + s(x - a)]}{n!} = \frac{(1 - t)^{n-1} f^{(n)}[a + t(x - a)]}{(n - 1)!}$$

where s and t are certain unknown numbers between 0 and 1; the s form is due to Lagrange, the t form to Cauchy.

The error due to neglecting the remainder term is less than $(\overline{P}_n)x^n$, or $(\overline{Q}_n)(x - a)^n$, where $\overline{P}_n$, or $\overline{Q}_n$, is the largest value taken on by P_n, or Q_n, when s or t ranges from 0 to 1. If this error, which depends on both n and x, approaches 0 as n increases (for any given value of x), then the general expression with remainder becomes (for that value of x) a convergent infinite series.

The sum of the first few terms of Maclaurin's series gives a good approximation to $f(x)$ for values of x near $x = 0$; Taylor's series gives a similar approximation for values near $x = a$.

Fourier's Series. Let $f(x)$ be a function which is finite in the interval from $x = -c$ to $x = +c$ and whose graph has finite arc length in that interval (see note below). Then, for any value of x between $-c$ and c,

$$f(x) = \tfrac{1}{2}a_0 + a_1 \cos \frac{\pi x}{c} + a_2 \cos \frac{2\pi x}{c} + a_3 \cos \frac{3\pi x}{c} + \cdots + b_1 \sin \frac{\pi x}{c}$$

$$+ b_2 \sin \frac{2\pi x}{c} + b_3 \sin \frac{3\pi x}{c} + \cdots$$

where the constant coefficients are determined as follows:

$$a_n = \frac{1}{c} \int_{-c}^{c} f(t) \cos \frac{n\pi t}{c} \, dt \qquad b_n = \frac{1}{c} \int_{-c}^{c} f(t) \sin \frac{n\pi t}{c} \, dt$$

In case the curve $y = f(x)$ is symmetrical with respect to the origin, the a's are all zero, and the series is a sine series. In case the curve is symmetrical with respect to the axis, the b's are all zero, and a cosine series results. (In this case, the series will be valid not only for values of x between $-c$ and c, but also for $x = -c$ and

$x = c$.) A Fourier series can always be integrated term by term; but the result of differentiating term by term may not be a convergent series.

   Note.   If $x = x_0$ is a point of discontinuity, $f(x_0)$ is to be defined as $\frac{1}{2}[f_1(x_0) + f_2(x_0)]$, where $f_1(x_0)$ is the limit of $f(x)$ when x approaches x_0 from below, and $f_2(x_0)$ is the limit of $f(x)$ when x approaches x_0 from above.

3. Integral Calculus

Indefinite Integrals.   An integral of $f(x)\ dx$ is any function whose differential is $f(x)\ dx$, and is denoted by $\int f(x)\ dx$. All the integrals of $f(x)\ dx$ are included in the expression $\int f(x)\ dx + C$, where $\int f(x)\ dx$ is any particular integral and C is an arbitrary constant. The process of finding (when possible) an integral of a given function consists in recognizing by inspection a function which, when differentiated, will produce the given function, or in transforming the given function into a form in which such recognition is easy. The most common integrable forms are collected in Table 3.7.

Definite Integrals.   The definite integral of $f(x)\ dx$ from $x = a$ to $x = b$ is denoted by

$$\int_a^b f(x)\ dx$$

   The fundamental theorem for the evaluation of a definite integral is the following:

$$\int_a^b f(x)\ dx = \int f(x)\ dx\big|_{x=b} - \int f(x)\ dx\big|_{x=a}$$

i.e., the definite integral is equal to the difference between two values of any one of the indefinite integrals of the function in question. In other words, the limit of a sum can be found whenever the function can be integrated.

   Properties of definite integrals:

$$\int_a^b = -\int_b^a$$

$$\int_a^c + \int_c^b = \int_a^b$$

The mean value of $f(x)$ between a and b:

$$\overline{f} = \frac{1}{b-a}\int_a^b f(x)\ dx$$

If b is variable, then

$$\frac{d}{db}\int_a^b f(x)\ dx = f(b)$$

TABLE 3.7   List of Most Common Integrals

$$\int a\,du = a\int du = au + C.$$

$$\int (u + v)\,dx = \int u\,dx + \int v\,dx.$$

$$\int u\,dv = uv - \int v\,du. \quad \text{(integration by parts)}$$

$$\int dy \int f(x, y)\,dx = \int dx \int f(x, y)\,dy.$$

$$\int x^n\,dx = \frac{x^{n-1}}{n + 1} + C, \text{ when } n \neq -1.$$

$$\int \frac{dx}{x} = \ln x + C = \ln cx.$$

$$\int e^x\,dx = e^x + C.$$

$$\int \sin x\,dx = -\cos x + C.$$

$$\int \cos x\,dx = \sin x + C.$$

$$\int \frac{dx}{\sin^2 x} = -\cot x + C.$$

$$\int \frac{dx}{\cos^2 x} = \tan x + C.$$

$$\int \frac{dx}{\sqrt{1 - x^2}} = \sin^{-1} x + C = -\cos^{-1} x + c.$$

$$\int \frac{dx}{1 + x^2} = \tan^{-1} x + C = -\cot^{-1} x + c.$$

$$\int (a + bx)^n\,dx = \frac{(a + bx)^{n+1}}{(n + 1)b} + C.$$

$$\int \frac{dx}{a + bx} = \frac{1}{b} \ln (a + bx) + C = \frac{1}{b} \ln c(a + bx).$$

$$\int \frac{ax}{(a + bx)^2} = -\frac{1}{b(a + bx)} + C.$$

$$\int \frac{dx}{1 - x^2} = \tfrac{1}{2} \ln \frac{1 + x}{1 - x} + C = \tanh^{-1} x + C, \text{ when } x < 1.$$

TABLE 3.7 List of Most Common Integrals (*Continued*)

$$\int \frac{dx}{x^2 - 1} = \frac{1}{2} \ln \frac{x - 1}{x + 1} + C = -\coth^{-1} x + C, \text{ when } x > 1.$$

23. $\int \dfrac{dx}{a + 2bx + cx^2} =$

$$\left. \begin{array}{l} \dfrac{1}{\sqrt{ac - b^2}} \tan^{-1} \dfrac{b + cx}{\sqrt{ac - b^2}} + C \end{array} \right\} [ac - b^2 > 0].$$

$$= \dfrac{1}{2\sqrt{b^2 - ac}} \ln \dfrac{\sqrt{b^2 - ac} - b - cx}{\sqrt{b^2 - ac} + b + cx} + C$$

$$\left. \vphantom{\begin{array}{c} a \\ b \\ c \end{array}} \right\} [b^2 - ac > 0].$$

$$= -\dfrac{1}{\sqrt{b^2 - ac}} \tanh^{-1} \dfrac{b + cx}{\sqrt{b^2 - ac}} + C$$

$$\int \sqrt{a + bx} \, dx = \frac{2}{3b} (\sqrt{a + bx})^3 + C.$$

$$\int \frac{dx}{\sqrt{a + bx}} = \frac{2}{b} \sqrt{a + bx} + C.$$

$$\int \frac{(m + nx) \, dx}{\sqrt{a + bx}} = \frac{2}{3b^2} (3mb - 2an + nbx) \sqrt{a + bx} + C.$$

$$\int \sqrt{a + 2bx + cx^2} \, dx = \frac{b + cx}{2c} \sqrt{a + 2bx + cx^2} + \frac{ac - b^2}{2c} \int \frac{dx}{\sqrt{a + 2bx + cx^2}} + C.$$

$$\int a^x \, dx = \frac{a^x}{\ln a} + C.$$

$$\int x^n e^{ax} \, dx = \frac{x^n e^{ax}}{a} \left[1 - \frac{n}{ax} + \frac{n(n - 1)}{a^2 x^2} - \cdots \pm \frac{n!}{a^n x^n} \right] + C.$$

$$\int \ln x \, dx = x \ln x - x + C.$$

$$\int \frac{\ln x}{x^2} \, dx = -\frac{\ln x}{x} - \frac{1}{x} + C.$$

$$\int \frac{(\ln x)^n}{x} \, dx = \frac{1}{n + 1} (\ln x)^{n+1} + C.$$

$$\int \sin^2 x \, dx = -\frac{1}{4} \sin 2x + \frac{1}{2}x + C = -\frac{1}{2} \sin x \cos x + \frac{1}{2}x + C.$$

$$\int \cos^2 x \, dx = \frac{1}{4} \sin 2x + \frac{1}{2}x + C = \frac{1}{2} \sin x \cos x + \frac{1}{2}x + C.$$

$$\int \sin mx \, dx = -\frac{\cos mx}{m} + C.$$

TABLE 3.7 List of Most Common Integrals (*Continued*)

$$\int \cos mx \, dx = \frac{\sin mx}{m} + C.$$

$$\int \sin mx \cos nx \, dx = -\frac{\cos (m + n)x}{2(m + n)} - \frac{\cos (m - n)x}{2(m - n)} + C.$$

$$\int \sin mx \sin nx \, dx = \frac{\sin (m - n)x}{2(m - n)} - \frac{\sin (m + n)x}{2(m + n)} + C.$$

$$\int \cos mx \cos nx \, dx = \frac{\sin (m - n)x}{2(m - n)} + \frac{\sin (m + n)x}{2(m + n)} + C.$$

$$\int \tan x \, dx = -\ln \cos x + C.$$

$$\int \cot x \, dx = \ln \sin x + C.$$

$$\int \frac{dx}{\sin x} = \ln \tan \frac{x}{2} + C.$$

$$\int \frac{dx}{\cos x} = \ln \tan \left(\frac{\pi}{4} + \frac{x}{2} \right) + C.$$

$$\int \frac{dx}{1 + \cos x} = \tan \frac{x}{2} + C.$$

$$\int \frac{dx}{1 - \cos x} = -\cot \frac{x}{2} + C.$$

$$\int \sin x \cos x \, dx = \tfrac{1}{2} \sin^2 x + C.$$

$$\int \frac{dx}{\sin x \cos x} = \ln \tan x + C.$$

$$\int \frac{\cos x \, dx}{a + b \cos x} = \frac{x}{b} - \frac{a}{b} \int \frac{dx}{a + b \cos x} + C.$$

$$\int \frac{\sin x \, dx}{a + b \cos x} = -\frac{1}{b} \ln (a + b \cos x) + C.$$

$$\int \frac{A + B \cos x + C \sin x}{a + b \cos x + c \sin x} \, dx = A \int \frac{dy}{a + p \cos y} + (B \cos u + C \sin u) \int \frac{\cos y \, dy}{a + p \cos y}$$

$$- (B \sin u - C \cos u) \int \frac{\sin y \, dy}{a + p \cos y}, \text{ where } b - p \cos u, c = p \sin u \text{ and } x - u = y.$$

$$\int e^{ax} \sin bx \, dx = \frac{a \sin bx - b \cos bx}{a^2 + b^2} e^{ax} + C.$$

TABLE 3.7   List of Most Common Integrals (*Continued*)

$$\int e^{ax} \cos bx \, dx = \frac{a \cos bx + b \sin bx}{a^2 + b^2} e^{ax} + C.$$

$$\int \sinh x \, dx = \cosh x + C.$$

$$\int \tanh x \, dx = \ln \cosh x + C.$$

$$\int \cosh x \, dx = \sinh x + C.$$

$$\int \coth x \, dx = \ln \sinh x + C.$$

If c is a parameter, then

$$\frac{\partial}{\partial c} \int_a^b f(x, c) \, dx = \int_a^b \frac{\partial f(x, c)}{\partial c} \, dx$$

The following definite integrals have received special names, and their values are tabulated in standard references:

1. Elliptic integral of the first kind:

$$F(u, k) = \int_0^u \frac{dx}{\sqrt{1 - k^2 \sin^2 x}} \qquad \text{when } k^2 < 1$$

2. Elliptic integral of the second kind:

$$E(u, k) = \int_0^u \sqrt{1 - k^2 \sin^2 x} \, dx \qquad \text{when } k^2 < 1$$

3, 4. Complete elliptic integrals of the first and second kinds: put $u = \pi/2$ in the two equations just given.

5. The probability integral

$$\frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} \, dx$$

6. The gamma function

$$\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} \, dx$$

Multiple Integrals.   Multiple integrals are of the form

$$\iint f(x, y) \quad \iiint f(x, y, z) \quad \text{etc.}$$

Two successive integrations, for example, an integration with respect to y holding

x constant, and an integration with respect to x between constant limits, will yield the value for the double integral

$$\int_a^b \int_{y_1(x)}^{y_2(x)} f(x, y)dy\, dx$$

Similarly a triple integral is evaluated by three successive single integrations. The order of integration can be reversed if the function $f(x, y, \ldots)$ is continuous.

DIFFERENTIAL EQUATIONS

1. General. An ordinary differential equation is one which contains a single independent variable, or argument, and a single dependent variable, or function, with its derivatives of various orders. A partial differential equation is one which contains a function of several independent variables, and its partial derivatives of various orders. The order of a differential equation is the order of the highest derivative which occurs in it. A solution of a differential equation is any relation between the variables, which, when substituted in the given equation, will satisfy it. The general solution of an ordinary differential equation of the nth order will contain n arbitrary constants. A differential equation is usually said to be solved when the problem is reduced to simple quadratures, i.e., intergrations of the form $y = \int f(x)\, dx$.

2. Methods of Solving Ordinary Differential Equations

*Differential Equations of the First Order**

1. If possible, separate the variables; i.e., collect all the x's and dx on one side, and all the y's and dy on the other side; then integrate both sides, and add the constant of integration.
2. If the equation is homogeneous in x and y, the value of dy/dx in terms of x and y will be of the form $dy/dx = f(y/x)$. Substituting $y = xt$ will enable the variables to be separated.

Solution:

$$\ln x = \int \frac{dt}{f(t) - t} + C$$

3. The expression $f(x, y)\, dx + F(x, y)\, dy$ is an *exact differential* if

$$\frac{\partial f(x, y)}{\partial y} = \frac{\partial F(x, y)}{\partial x} \quad (= P,\ \text{say}).$$

In this case the solution of $f(x, y)\, dx + F(x, y)\, dy = 0$ is

---

*This section taken in part from Ref. 1, *Marks' Standard Handbook for Mechanical Engineers,* 9th ed., by E. A. Avallone and T. Baumeister III (eds.). Copyright © 1987. Used by permission of McGraw-Hill, Inc. All rights reserved.

$$\int f(x, y) \, dx + \int [F(x, y) - \int P \, dx] \, dy = C$$

or

$$\int F(x, y) \, dy + \int [f(x, y) - \int P \, dy] \, dx = C$$

4. Linear differential equation of the first order:

$$\frac{dy}{dx} + f(x) \cdot y = F(x)$$

Solution: $y = e^{-P}[\int e^P F(x) \, dx + C]$, where $P = \int f(x) \, dx$.

5. Bernoulli's equation:

$$\frac{dy}{dx} + f(x) \cdot y = F(x) \cdot y^n$$

Substituting $y^{1-n} = v$ gives $(dv/dx) + (1 - n)f(x) \cdot v = (1 = n)F(x)$, which is linear in v and x.

6. Clairaut's equation: $y = xp + f(p)$, where $p = dy/dx$. The solution consists of the family of lines given by $y = Cx + f(C)$, where C is any constant, together with the curve obtained by eliminating p between the equations $y = xp + f(p)$ and $x + f'(p) = 0$, where $f'(p)$ is the derivative of $f(p)$.

Differential Equations of the Second Order

7.

$$\frac{d^2y}{dx^2} = -n^2 y$$

Solution: $y = C_1 \sin (nx + C_2)$ or $y = C_3 \sin nx + C_4 \cos nx$.

8.

$$\frac{d^2y}{dx^2} = +n^2 y$$

Solution: $y = C_1 \sinh (nx + C_2)$ or $y = C_3 e^{nx} + C_4 e^{-nx}$

9.

$$\frac{d^2y}{dx^2} = f(y)$$

Solution:

$$x = \int \frac{dy}{\sqrt{C_1 + 2P}} + C_2$$

where $P = \int f(y) \, dy$

10.

$$\frac{d^2y}{dx^2} = f(x)$$

Solution:

$$y = \int P \, dx + C_1 x + C_2,$$

where $P = \int f(x)dx$, or

$$y = xP - \int xf(x) \, dx + C_1 x + C_2.$$

11.

$$\frac{d^2y}{dx^2} = f\left(\frac{dy}{dx}\right)$$

Putting

$$\frac{dy}{dx} = z \qquad \frac{d^2y}{dx^2} = \frac{dz}{dx} \qquad x = \int \frac{dz}{f(z)} + C_1 \qquad y = \int \frac{zdz}{f(z)} + C_2$$

Then eliminate z from these two equations.

12. The equation for damped vibration:

$$\frac{d^2y}{dx^2} + 2b\frac{dy}{dx} + a^2y = 0$$

CASE 1: If $a^2 - b^2 > 0$, let $m = \sqrt{a^2 - b^2}$. *Solution:* $y = C_1 e^{-bx} \sin(mx + C_2)$ or $y = e^{-bx}[C_3 \sin(mx) + C_4 \cos(mx)]$.
CASE 2: If $a^2 - b^2 = 0$, solution is $y = e^{-bx}(C_1 + C_2x)$.
CASE 3: If $a^2 - b^2 < 0$, let $n = \sqrt{b^2 - a^2}$. *Solution:* $y = C_1 e^{-bx} \sinh(nx + C_3)$ or $y = C_3 e^{-(b+n)x} + C_4 e^{-(b-n)x}$.

13.

$$\frac{d^2y}{dx^2} + 2b\frac{dy}{dx} + a^2y = c$$

Solution: $y = c/a^2 + y_1$, where y_1 = the solution of the corresponding equation with second member zero (see item 12 above).

14.

$$\frac{d^2y}{dx^2} + 2b\frac{dy}{dx} + a^2y = c \sin(kx)$$

Solution:

$$y = R \sin(kx - S) + y_1$$

where $R = c/\sqrt{(a^2 - k^2)^2 + 4b^2k^2}$, $\tan S = 2bk/(a^2 - k^2)$, and y_1 = the solution of the corresponding equation with second member zero (see item 12 above).

15.

$$\frac{d^2y}{dx^2} + 2b\frac{dy}{dx} + a^2y = f(x)$$

Solution: $y = y_0 + y_1$, where $y_0 = $ any particular solution of the given equation, and $y_1 = $ the general solution of the corresponding equation with second member zero (see item 12 above).

If $b^2 < a^2$,

$$y_0 = \frac{1}{2\sqrt{b^2 - a^2}} [e^{m1x} \int e^{-m1x} f(x)\, dx - e^{m2x} \int e^{-m2x} f(x)\, dx]$$

where $m_1 = -b + \sqrt{b^2 - a^2}$ and $m_2 = -b - \sqrt{b^2 - a^2}$.

If $b^2 < a^2$, let $m = \sqrt{a^2 - b^2}$; then

$$y_0 = \frac{1}{m} e^{-bx} [\sin (mx) \int e^{bx} \cos (mx) \cdot f(x)\, dx$$

$$- \cos (mx) \int e^{bx} \sin (mx) \cdot f(x)\, dx]$$

If $b^2 = a^2$, $y_0 = e^{-bx}[x \int e^{bx} f(x)\, dx - \int x \cdot e^{bx} f(x)\, dx]$.

Types 12 to 15 are examples of linear differential equations with constant coefficients. The solutions of such equations are often found most simply by the use of Laplace transforms.

Linear Equations (Constant Coefficients). For the linear equation of the nth order

$$\frac{A_n(x)d^n y}{dx^n} + \frac{A_{n01}(x)d^{n-1}y}{dx^{n-1}} + \cdots + \frac{A_1(x)dy}{dx} + A_0(x)y = E(x)$$

the general solution is $y = u + c_1 u_1 + c_2 u_2 + \cdots + c_n u_n$. Here u, the particular integral, is any solution of the given equation, and $u_1, u_2, \ldots, u_n$ form a fundamental system of solutions of the homogeneous equation obtained by replacing $E(x)$ by zero.

To solve the homogeneous equation of the nth order $A_n d^n y/dx^n + A_{n-1}d^{n-1}y/dx^{n-1} + \cdots + A_1 dy/dx + A_0 y = 0$, $A_n \neq 0$, where $A_n, A_{n-1}, \ldots, A_0$ are constants, find the roots of the auxiliary equation

$$A_n p^n + A_{n-1}p^{n-1} + \cdots + A_1 p + A_0 = 0$$

For each simple real root r, there is a term ce^{rx} in the solution. The terms of the solution are to be added together. When r occurs twice among the n roots of the auxiliary equation, the corresponding term is $e^{rx}(c_1 + c_2 x)$. When r occurs three times, the corresponding term is $e^{rx}(c_1 + c_2 x + c_3 x^2)$, and so forth. When there is a pair of conjugate complex roots $a + bi$ and $a - bi$, the real form of the terms in the solution is $e^{ax}(c_1 \cos bx + d_1 \sin bx)$. When the same pair occurs twice, the corresponding term is $e^{ax}[(c_1 + c_2 x) \cos bx + (d_1 + d_2 x) \sin bx]$, and so forth.

The general nonhomogeneous linear differential equation of order n, with constant coefficients, or

$$\frac{A_n d^n y}{dx^n} + \frac{A_{n-1}d^{n-1}y}{dx^{n-1}} + \cdots + \frac{A_1 dy}{dx} + A_0 y = E(x)$$

may be solved by adding any particular integral to the complementary function, or

general solution, of the homogeneous equation obtained by replacing $E(x)$ by zero. The complementary function may be found from the rules just given. And the particular integral may be found assuming that $E(x)$ is a single term or a sum of terms each of which is of the type k, $k \cos bx$, $k \sin bx$, ke^{ax}, where a and b are any real numbers. (See Ref. 1 for more details.)

Solutions of Partial Differential Equations. The only means of solution that will be discussed is the separation-of-variables method. For the following equation,

$$f_1 \frac{\partial^2 z}{\partial x^2} + f_2 \frac{\partial z}{\partial x} + f_3 z + g_1 \frac{\partial^2 z}{\partial y^2} + g_2 \frac{\partial x}{\partial y} + g_3 z = 0$$

the solution of $z = z(x, y)$ will be assumed to have the form of a product of two functions $X(x)$ and $Y(y)$, which are functions of x and y only, respectively:

$$z = z(x,y) = X(x)Y(y)$$

Substituting $z = XY$ into the original differential equation, we obtain

$$-\frac{1}{X}\left[f_1 \frac{\partial^2 X}{\partial x^2} + f_2 \frac{\partial X}{\partial x} + f_3 X\right] = \frac{1}{Y}\left[g_1 \frac{\partial^2 Y}{\partial y^2} + g_2 \frac{\partial Y}{\partial y} + g_3 Y\right]$$

Note that the left-hand side contains the function of x only and that the right-hand side contains functions of y only. Since the right- and left-hand sides are independent of x and y, respectively, they must be equal to a common constant, called a separation constant α; thus,

$$f_1 \frac{d^2 X}{dx^2} + f_2 \frac{dX}{dx} + [f_3 + \alpha]X = 0$$

and

$$g_1 \frac{d^2 Y}{dy^2} + g_2 \frac{dY}{dy} + [g_3 - \alpha]Y = 0$$

Once the solutions to the above have been obtained, the product solution XY is obtained. The method outlined may be extended to additional variables and it is applicable also when f_1, f_2, f_3 are functions of x and g_1, g_2, and g_3 are functions of y.

LAPLACE TRANSFORMATION

The Laplace transformation of a function $f(t)$ is

$$F(s) = \mathcal{L}[f(t)] = \int_0^\infty e^{-st} f(t) \, dt$$

where $f(t)$ = a function of a real variable (usually t = time)
      s = a complex variable of the form $(\sigma + j\omega)$

$F(s)$ = an equation expressed in the transform variable s, resulting from operating on a function of time with the Laplace integral

$\mathcal{L}$ = an operational symbol indicating that the quantity which it prefixes is to be transformed into the frequency domain

Example

$$f(t) = A \qquad f(t) = \int_0^\infty Ae^{-st} \, dt = \frac{A}{s}$$

Table 3.8 lists the transforms of common functions.

An inverse transformation is represented symbolically as

$$\mathcal{L}^{-1}F(s) = f(t)$$

For any $f(t)$ there is only one direct transform, $F(s)$. For any given $F(s)$ there is only one inverse transform $f(t)$. Therefore, tables are generally used for determining inverse transforms.

COMPLEX VARIABLES

1. Complex Numbers. A complex number z consists of a real part x and an imaginary part y and is represented as

$$z = x + iy$$

where $i = \sqrt{-1}$ $\qquad\qquad (i^2 = -1)$

The conjugate $\bar{z}$ of a complex number is defined as

$$\bar{z} = x - iy$$

Two complex numbers are equal only if their real parts are equal and their imaginary parts are equal; i.e.,

$$x_1 + iy_1 = x_2 + iy_2$$

only if

$$x_1 = x_2 \qquad \text{and} \qquad y_1 = y_2$$

Also

$$x + iy = 0$$

only if

$$x = 0 \qquad \text{and} \qquad y = 0$$

Complex numbers satisfy the distributive, associative, and commutative laws of algebra.

Complex numbers may be graphically represented on the $z(x - y)$ plane or in polar (r, θ) coordinates. The polar coordinates of a complex number are

TABLE 3.8 Laplace Transforms

$f(t)$	$F(s) = [f(t)]$
A	A/s
$f_1(t) + f_2(t)$	$F_1(s) + F_2(s)$
$e^{-\alpha t}$	$\dfrac{1}{s + \alpha}$
$\dfrac{1}{\tau} e^{-t/\tau}$	$\dfrac{1}{rs + 1}$
$Ae^{-\alpha t}$	$\dfrac{A}{s + \alpha}$
$\sin \beta t$	$\dfrac{\beta}{s^2 + \beta^2}$
$\cos \beta t$	$\dfrac{s}{s^2 + \beta^2}$
$\dfrac{1}{\beta} e^{-\alpha t} \sin \beta t$	$\dfrac{1}{s^2 + 2\alpha s + \alpha^2 + \beta^2}$
$\dfrac{e^{-\alpha t}}{\beta - \alpha} - \dfrac{e^{-\beta t}}{\beta - \alpha}$	$\dfrac{1}{(s + \alpha)(s + \beta)}$
$\dfrac{Ae^{-\alpha t} - Be^{-\beta t}}{C}$	$\dfrac{s + a}{(s + \alpha)(s + \beta)}$
where $A = a - \alpha$, $B = a - \beta$, $C = \beta - \alpha$	
$\dfrac{e^{-\alpha t}}{A} + \dfrac{e^{-\beta t}}{B} + \dfrac{e^{-\delta t}}{C}$	$\dfrac{1}{(s + \alpha)(s + \beta)(s + \delta)}$
where $A = (\beta - \alpha)(\delta - \alpha)$ $\quad\quad B = (\alpha - \beta)(\delta - \beta)$ $\quad\quad C = (\alpha - \delta)(\beta - \delta)$	
t	$\dfrac{1}{s^2}$
t^2	$\dfrac{2}{s^3}$
t^n	$\dfrac{n!}{s^{n+1}}$
$d/dt[f(t)]$	$sF(s) - f(0^+)^*$
$d^2/dt^2[f(t)]$	$s^2F(s) - sf(0^+) - \dfrac{df}{dt}(0^+)$
$d^3/dt^3[f(t)]$	$s^2F(s) - s^2f(0^+) - s\dfrac{df(0^+)}{dt} - \dfrac{d^2f(0^+)}{dt^2}$
$\int f(t)dt$	$\dfrac{1}{s}[F(s) + \int f(t)dt|0^+|$
$\dfrac{1}{\alpha} \sinh \alpha t$	$\dfrac{1}{s^2 - \alpha^2}$
$\cosh \alpha t$	$s/(s^2 - \alpha^2)$

*$f(0^+)$ = initial value of $f(t)$, evaluated as t approaches zero from positive values.
Source: From Avallone and Baumeister.[1]

$$r = \sqrt{x^2 + y^2} = |z| = \text{mod } z, \, r \geq 0$$

where mod = modulus, and

$$\theta = \tan^{-1} \frac{y}{x} = \arg z = \text{amp } z$$

where arg = argument and amp = amplitude. arg z is multiple-valued, but for an angular interval of range 2π there is only one value of θ for a given z.

2. Elementary Complex Functions

Polynomials. A polynomial in z, $a_n z^n + a_{n-1} z^{n^{1-}} + \cdots + a_0$, where n is a positive integer, is simply a sum of complex numbers times integral powers of z which have already been defined. Every polynomial of degree n has precisely n complex roots provided each multiple root of multiplicity m is counted m times.

Exponential Functions. The exponential function e^z is defined by the equation $e^z = e^{x+iy} = e^x \cdot e^{iy} = e^x (\cos y + i \sin y)$. Properties: $e^0 = 1$; $e^{z1} \cdot e^{z2} = e^{z1+z2}$; $e^{z1}/e^{z2} = e^{z1-z2}$; $e^{z+2k\pi i} = e^z$.

Trigonometric Functions. $\sin z = (e^{iz} - e^{-iz})/2i$; $\cos z = (e^{iz} + e^{-iz})/2$; $\tan z = \sin z/\cos z$; $\cot z = \cos z/\sin z$; $\sec z = 1/\cos z$; $\cos z = 1/\sin z$. Fundamental identities for these functions are the same as their real counterparts. Thus $\cos^2 z + \sin^2 z = 1$, $\cos(z_1 \pm z_2) = \cos z_1 \cos z_2 \pm \sin z_1 \sin z_2$, $\sin(z_1 \pm z_2) = \sin z_1 \cos z_2 \pm \cos z_1 \sin z_2$. The sine and cosine of z are periodic functions of period 2π; thus $\sin(z + 2\pi) = \sin z$. For computation purposes, $\sin z = \sin (x + iy) = \sin x \cosh y + i \cos x \sinh y$, where $\sin x$, $\cosh y$, etc., are the real trigonometric and hyperbolic fucntions. Similarly, $\cos z = \cos x \cosh y = i \sin x \sinh y$. If $x = 0$ in the results given, $\cos iy = \cosh y$, $\sin iy = i \sinh y$.

VECTORS

1. Representation.* A vector quantity has magnitude and direction; a scalar quantity has magnitude only. Common vector quantities are acceleration, alternating currents, voltages, force, and velocity. A vector can be represented graphically by a straight line with an arrowhead, as in Fig. 3.20. (Length represents magnitude; direction is determined from the position of the line and the arrowhead.)

Vectors are usually indicated by boldface type (**A**), or by an arrow over the symbol (*A*), or by a bar (**A̅**).
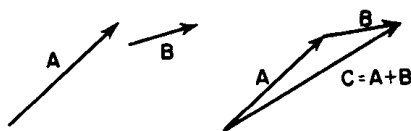
FIGURE 3.20 (*From Perry.*[4])

---

*This section and the next one on algebra are partly taken from Ref. 4, *Engineering Manual,* 3d ed., by R. H. Perry (ed.). Copyright © 1976. Used by permission of McGraw-Hill, Inc. All rights reserved.

A vector **V** in three dimensions can be represented by its projections along three mutually perpendicular lines, the x, y, and z axes. Vectors of unit magnitude, directed in the positive sense along these three axes, are denoted by **i, j,** and **k,** respectively. If a, b, and c represent the lengths of the projections of **V** along these axes, we may represent **V** as $\mathbf{V} = a\mathbf{j} + b\mathbf{j} + c\mathbf{k}$ (Fig. 3.20). The length (magnitude) of **V** is

$$\mathbf{V} = (a^2 + b^2 + c^2)^{1/2}$$

2. Algebra

Equality.   **A** = **B** if and only if both have the same magnitude and the same direction.

Addition and Subtraction. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ (commutative law); $\mathbf{A} + \mathbf{B} + \mathbf{C} = (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$ (associative law). If $\mathbf{A} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{B} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, $\mathbf{A} \pm \mathbf{B} = (a_1 \pm b_1)\mathbf{i} + (a_2 \pm b_2)\mathbf{j} + (a_3 \pm b_3)\mathbf{k}$.

Product of Vector V and Scalar s. $s\mathbf{V} = \mathbf{V}s = (sa)\mathbf{i} + (sb)\mathbf{j} + (sc)\mathbf{k}$.

Scalar Product of Two Vectors $\mathbf{V_1}$, $\mathbf{V_2}$.   The scalar (dot or inner) product, indicated by $\mathbf{V_1} \cdot \mathbf{V_2}$, is a scalar defined by $\mathbf{V_1} \cdot \mathbf{V_2} = |\mathbf{V_1}||\mathbf{V_2}| \cos \theta$, where θ = angle between the vectors. $\mathbf{V_1} \cdot \mathbf{V_2} = a_1a_2 + b_1b_2 + c_1c_2$; $(\mathbf{V_1} + \mathbf{V_2}) \cdot \mathbf{V_3} = \mathbf{V_1} \cdot \mathbf{V_3} + \mathbf{V_2} \cdot \mathbf{V_3}$; $\mathbf{V_1} \cdot (\mathbf{V_2} + \mathbf{V_3}) = \mathbf{V_1} \cdot \mathbf{V_2} + \mathbf{V_1} \cdot \mathbf{V_3} = (\mathbf{V_2} + \mathbf{V_3}) \cdot \mathbf{V_1}$ (commutative); $\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1$; $\mathbf{i} \cdot \mathbf{j} = \mathbf{i} \cdot \mathbf{k} = \mathbf{j} \cdot \mathbf{k} = 0$.

Vector Product.   With reference to Fig. 3.21, the vector (outer) product of $\mathbf{V_1}$ and $\mathbf{V_2}$ is defined as the vector $\mathbf{V} = \mathbf{V_1} \times \mathbf{V_2}$, $\mathbf{V_1} \times \mathbf{V_2} = (b_1c_2 - b_2c_1)\mathbf{i} + (c_1a_2 - c_2a_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}$. $|\mathbf{V}| = |\mathbf{V_1}||\mathbf{V_2}| \sin \theta$; $\mathbf{V_1} \times \mathbf{V_2} = -\mathbf{V_2} \times \mathbf{V_1}$; $\mathbf{V_1} \times (\mathbf{V_2} + \mathbf{V_3}) = \mathbf{V_1} \times \mathbf{V_2} + \mathbf{V_1} \times \mathbf{V_3}$; $(\mathbf{V_1} + \mathbf{V_2}) \times \mathbf{V_3} = \mathbf{V_1} \times \mathbf{V_3} + \mathbf{V_2} \times \mathbf{V_3}$; $\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = 0$; $\mathbf{i} = \mathbf{j} \times \mathbf{k} = -\mathbf{k} \times \mathbf{j}$; $\mathbf{j} = \mathbf{k} \times \mathbf{i} = -\mathbf{i} \times \mathbf{k}$; $\mathbf{k} = \mathbf{i} \times \mathbf{j} = -\mathbf{j} \times \mathbf{i}$; $\mathbf{V} \times \mathbf{V} = 0$.

Multiple Products.   (1) $\mathbf{A}(\mathbf{B} \cdot \mathbf{C})$; here $\mathbf{B} \cdot \mathbf{C}$ is a scalar, so that $\mathbf{A}(\mathbf{B} \cdot \mathbf{C})$ is a vector parallel to **A**. Clearly, $\mathbf{A}(\mathbf{B} \cdot \mathbf{C}) \neq (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$. (2) $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B})$. (3) $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$.

Vector Function. $\mathbf{r} = \mathbf{F}(x,y,z)$ gives a vector **r** as a function of scalars x, y, and z.
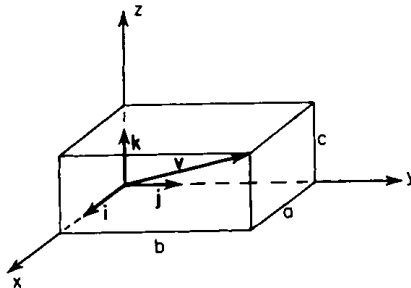
FIGURE 3.21   (*From Perry.*[4])

3. Differential Operations.   By definition,

$$\nabla = \text{del} = \mathbf{i}\,\frac{\partial}{\partial x} + \mathbf{j}\,\frac{\partial}{\partial y} + \mathbf{k}\,\frac{\partial}{\partial z}$$

and

$$\nabla^2 = \text{Laplacian } \nabla^2 = \text{Laplacian} = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

$$\nabla s = \text{grad } s = \frac{\partial s}{\partial x}\,\mathbf{i} + \frac{\partial s}{\partial y}\,\mathbf{j} + \frac{\partial s}{\partial z}\,\mathbf{k},$$

the gradient of a scalar function $S(x,y,z)$. For a vector function $\mathbf{V}(x,y,z) = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$, the divergence of $\mathbf{V}$ is

$$\nabla \cdot \mathbf{V} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$$

And the curl of $\mathbf{V}$ is

$$\nabla \times \mathbf{V} = \text{curl } \mathbf{V} = \text{rot } \mathbf{V} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}$$

The divergence theorem states that if $\mathbf{F}$ is a vector function and V is a volume bounded by a surface S, then

$$\iiint_V \text{div } \mathbf{F}\, dv = \iiint_V \nabla \cdot \mathbf{F}\, dv = \iint_S \mathbf{F} \cdot d\mathbf{S}$$

The integrations are to be carried out over the volume V and the surface S. And if $\mathbf{n}$ is the unit outward normal and $dS = |d\mathbf{S}|$ is the scalar element of surface, $d\mathbf{S} = \mathbf{n}\, dS$, $\mathbf{F} \cdot d\mathbf{S} = \mathbf{F} \cdot \mathbf{n}\, dS$.

   Stokes' theorem states that if $\mathbf{F}$ is a vector function and S is a surface bounded by a simple closed curve $\mathbf{C}$, then

$$\int_C \mathbf{F} \cdot d\mathbf{r} \iint_S \text{curl } \mathbf{F} \cdot d\mathbf{S} = \iint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S}$$

Here, $d\mathbf{r} = dx\,\mathbf{i} + dy\,\mathbf{j} + dz\,\mathbf{k}$. The theorem implies that the surface integral of $(\nabla \times \mathbf{F})$ over any surface S which is bounded by C is equal to the line integral of $\mathbf{F}$ over the contour C, taken in the direction related to that of $\mathbf{n}$ (in $d\mathbf{S} = \mathbf{n}\, dS$) by the right-hand rule.

STATISTICS AND PROBABILITY

This section includes topics related to measurements and errors.

1. Error of Observation. The error of an observation is $e_i = m_i - m$, $i = 1, 2,$ $\ldots, n$, where m_i are the observed values, e_i the errors, and m the mean value, that is, the arithmetic mean of a very large number (theoretically infinite) of observations.

In a large number of measurements, random errors are as often negative as positive and have little effect on the arithmetic mean. All other errors are classed as systematic. If due to the same cause, they affect the mean in the same sense and give it a definite bias.

Best Estimate and Measured Value. If all systematic errors have been eliminated, it is possible to consider the sample of individual repeated measurements of a quantity with a view to securing the "best" estimate of the mean value m and assessing the degree of reproducibility which has been obtained. The final result will then be expressed in the form $E \pm L$, where E is the best estimate of m, and L the characteristic limit of variation associated with a certain risk. The value measured is not merely E, but the entire result $E \pm L$.

The Arithmetic Mean. If a large number of measurements have been made to determine directly the mean m of a cerain quantity, all measurements having been made with equal skill and care, the best estimate of m from a sample of n is the arithmetic mean $\overline{m}$ of the measurements in the sample,

$$\overline{m} = \frac{1}{n} \sum_{i=1}^{n} m_i$$

2. Standard Deviation. Standard deviation is the root-mean-square of the deviations e_i of a set of observations from the mean,

$$\sigma = \left(\frac{1}{n} \sum_{i=1}^{n} e_i^2 \right)^{1/2}$$

Since neither the mean m nor the errors of observation e_i are ordinarily known, the deviations from the arithmetic mean, or the residuals, $x_i = m_i - \overline{m}$, $i = 1, 2, \ldots,$ n, will be referred to as errors. Likewise, for the unbiased value

$$\sigma = (n - 1)^{-1/2} \left[\sum_{i=1}^{n} (m_i - \overline{m})^2 \right]^{1/2} = (n - 1)^{-1/2} \left(\sum_{i=1}^{n} e_i^2 \right)^{1/2}$$

will be used, in which n is replaced by $n - 1$ since one degree of freedom is lost by using $\overline{m}$ instead of m, $\overline{m}$ being related to the m_i.

3. Normal Distribution

Relative Frequency of Errors. The Gauss-Laplace, or normal, distribution of frequency of errors is (Fig. 3.22)

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$$

or
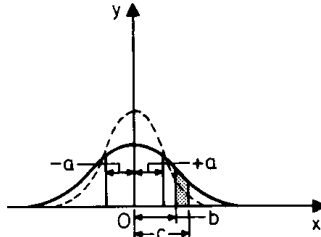
FIGURE 3.22

$$y = \frac{1}{\sqrt{\pi}} h e^{-h^2 x^2}$$

where $2h^2\sigma^2 = 1$, or $h = 1/(\sqrt{2}\alpha)$, and y represents the proportionate number of errors of value x. The area under the curve in Fig. 3.22 is unity. The dotted curve is also an error distribution curve with a greater value of the precision index h which measures the concentration of observations about their mean.

Probability. The fraction of the total number of errors whose values lie between $x = -a$ and $x = a$ is

$$P = \frac{h}{\sqrt{\pi}} \int_{-a}^{+a} e^{-h^2 x^2}\, dx = \frac{2}{\sqrt{\pi}} \int_{0}^{ha} e^{-h^2 x^2}\, d(hx) \qquad (3.1)$$

that is, P is the probability of an error x having a value between $-a$ and a. Similarly, the shaded area in Fig. 3.22 represents the probability of errors between b and c.

Probable Error. Results of measurements are sometimes expressed in the form $E \pm r$, where r is the probable error of a single observation and is defined as the number which the actual error may with equal probability be greater or less than. From Eq. (3.1) and

$$\frac{2}{\sqrt{\pi}} \int_{0}^{hr} e^{-h^2 x^2}\, d(hx) = 0.50$$

$$hr = 0.47694$$

or

$$r = 0.4769x\, \sqrt{2}\sigma = 0.6745\sigma$$

NUMERICAL METHODS

Numerical techniques do not always yield exact results. The goal of approximate and numerical methods is to provide convenient techniques for obtaining useful information from mathematical formulations of physical problems. Often these mathematical statements are not solvable by analytical means. Or perhaps analytic

solutions are available but in a form that is inconvenient for direct interpretation numerically.

1. Approximation Identities. For the following relationships the sign $\cong$ means approximately equal to, when X is small:

$$\frac{1}{1 \pm X} \cong \mp X \qquad \frac{1 + Y}{1 \mp X} \cong 1 + Y \pm X$$

$$(1 \pm X)^n \cong 1 \pm nX \qquad (a \pm X)^2 \cong a^2 \pm 2aX$$

$$\sin X \cong X(X \text{ rad})$$

$$\sqrt{Y(Y + X)} \cong \frac{2Y + X}{2} \qquad \sqrt{1 \pm X} \cong 1 \pm \frac{X}{2}$$

$$(1 \pm X)^{-n} \cong 1 \mp nX \qquad (1 \pm X)^{-1/2} \cong 1 \mp \frac{X}{2}$$

$$e^2 \cong 1 + X \qquad \tan X \cong X$$

$$\sqrt{Y^2 + X^2} \cong Y + \frac{X^2}{2Y} \left(\frac{X}{Y} \text{ small}\right)$$

$$\sum_1^m \sqrt{n} \cong \tfrac{2}{3}M^{3/2} + \frac{\sqrt{m}}{2} - 0.245$$

$$n! \cong e^{-n}n^n\sqrt{2\pi n} \qquad n! \cong \sqrt{2\pi} \left\{\frac{\sqrt{n^2 + n + \frac{1}{4}}}{e}\right\}^{n+1/2}$$

2. Interpolation and Finite Differences. Let $y_0, y_1, y_2, \ldots, y_n$ denote a set of values of a function $y = f(x)$. Then $y_1 - y_0, y_2 - y_1$, etc., are called the *first differences* of y and are written with the notation

$$\Delta y_n = y_{n+1} - y_n \qquad (3.1)$$

The differences of these first differences are called *second differences,* so that

$$\Delta^2 y_n = \Delta y_{n+1} - \Delta y_n - y_{n+2} - 2y_{n+1} + y_n$$

Likewise,

$$\Delta^3 y_n = \Delta^2 y_{n+1} - \Delta^2 y_n = y_{n+3} - 3y_{n+2} + 3y_{n+1} - y_n$$

etc.

The relation between differences and derivatives is

$$\Delta^n f(x) = (\Delta x)^n f^{(n)} f^{(n)}(x + \theta_n \Delta x) \qquad 0 < \theta < 1 \qquad (3.2)$$

Consequently,

$$\lim_{\Delta x \to 0} \frac{\Delta^n f(x)}{(\Delta x)^n} = f^{(n)}(x)$$

The process of interpolation is used not only for finding in-between values of a function from a numerical table, but also for replacing complicated functions by simpler ones which can be made to coincide with them as closely as desired over a given interval.

Since polynomials are the simplest functions, interpolating functions are usually polynomials, the coefficients of whose terms are expressed in differences of ascending order.

The most important and useful of the polynomial interpolation formulas are Newton's, Stirling's, and Bessel's (see Refs. 5 and 6). Newton's formula for forward interpolation reads:

$$y = y_0 + u\Delta y_0 + \frac{u(u-1)}{2!}\Delta^2 y_0 + \frac{u(u-1)(u-2)}{3!}\Delta^3 y_0$$

$$+ \frac{u(u-1)(u-2)(u-3)}{4!}\Delta^4 y_0 + \cdots$$

$$+ \frac{u(u-1)(u-2)\cdots(u-n+1)}{n!}\Delta^n y_0 \tag{3.3}$$

where $u = (x - x_0)/h$, and it is called by that name because it employs y's from y_0 forward to the right.

y_0 may, of course, be any tabular value, but the formula will contain only values of y which come after the value chosen as the starting point.

3. Numerical Differentiation. The derivative of a tabulated function can be found at any point by representing the function by an appropriate interpolation formula and then differentiating the formula with respect to u. Then the derivative is given by the relation

$$\frac{dy}{dx} = \frac{1}{h}\frac{dy}{du}$$

where dy/du can be obtained, for example, from Eq. (3.3) by differentiation.

4. Numerical Integration. The numerical value of a definite integral can be found, to any desired degree of accuracy, by means of any of several quadrature formulas which express the integral as a linear function of given values of the integrand. In any problem, the definite integral should be set up before a quadrature formula is applied.

When the values of the integrand are given at equidistant intervals of width h of the independent variable, various quadrature formulas can be derived by integrating any of the standard interpolation formulas between given limits. By integrating Eq. (3.3), for example, and retaining differences of successively higher orders, the following formulas can be derived:

Trapezoidal Rule

$$\int_a^b y\, dx = \int_a^b f(x)\, dx = h(\tfrac{1}{2}y_0 + y_1 + y_2 + \cdots$$

$$+ y_{n-2} + y_{n-1} + \tfrac{1}{2}y_n) \tag{3.4}$$

where $h = (b - a)/n$, and n is either even or odd:

$$\text{Correction term} = \text{error} = -\frac{nh^3}{12} f''(\xi) \qquad a \le \xi \le b$$

Equation (3.4) is the simplest of the quadrature formulas, but is also the least accurate. The accuracy can be increased by decreasing the interval h.

Simpson's Rule

$$\int_a^b y\, dx = \int_a^b f(x)\, dx = \frac{h}{3}\,(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots$$

$$+ 2y_{n-2} + 4y_{n-1} + y_n) \tag{3.5}$$

$$\text{Error} = -\frac{nh^5}{180} f^{iv}(\xi) \quad a \le \xi \le b$$

$$= -\frac{(b - a)}{180} h^4 f^{iv}(\xi) \approx -\frac{h}{90} \Sigma\, \Delta^4 y$$

In (3.5), n must be *even*.

Because of its simplicity, flexibility, and relatively high accuracy, Simpson's rule is the most useful of all quadrature formulas.

5. Numerical Solution of Ordinary Differential Equations.* A variety of methods have been devised to solve ordinary differential equations numerically.

Equations of the First Order. Let the first-order differential equation be

$$\frac{dy}{dx} = f(x, y)$$

with the initial condition (x_0, y_0), that is, $y = y_0$ when $x = x_0$. (Note that any differential equation of the first order in the variables x and y can be written in this form.)

The procedure, called the *modified Euler method,* is as follows:

Step 1. From the given initial conditions (x_0, y_0) compute $y_0' = f(x_0, y_0)$ and

$$y_0'' = \frac{\partial f(x_0, y_0)}{\partial x} + \frac{\partial f(x_0, y_0)}{\partial y} y_0'$$

Then determine $y_0 = y_0 + hy_0' + (h^2/2)y_0''$, where h = subdivision of the independent variable.

Step 2. Determine $y_1' = f(x_1, y_1)$. $(x_1 = x_0 + h)$. These prepare us for:

Predictor Steps

Step 3. For $n \geq 1$ calculate $(y_{n+1})_1 = y_{n-1} + 2hy_n'$.

Step 4. Calculate $(y_{n+1}')_1 = f[x_{n+1}, (y_{n+1})_1]$.

Corrector Steps

Step 5. Calculate $(y_{n+1})_2 = y_n + (h/2)[(y_{n+1}')_1 + y_n']$, where y_n, y_n' without the subscripts are the previous values obtained by this process (or by steps 1 and 2).

Step 6. $(y_{n+1}')_2 = f[x_{n+1}, (y_{n+1})_2]$.

Step 7. Repeat the corrector steps 5 and 6 if necessary until the desired accuracy is produced in y_{n+1}, y_{n+1}'.

Example. Consider the equation $y' = 2y^2 + x$ with the initial conditions $y_0 = 1$ when $x_0 = 0$. Let $h = 0.1$. A few steps of the computation are illustrated.

Step 1

$$y_0' = 2y_0{}^2 + x_0 = 2$$

$$y_0' = 1 + 4y_0' = 1 + 8 = 9$$

$$y_1 = 1 + (0.1)(2) + \left[\frac{(0.1)^2}{2}\right]9 = 1.245$$

Step 2

$$y_1' = 2y_1^2 + x_1 = 3.100 + 0.1 = 3.210$$

Step 3

$$(y_2)_1 = y_0 + 2hy_1' = 1 + 2(0.1)3.210 = 1.642$$

Step 4

$$(y_2')_1 = 2(y_2)_1^2 + x_2 = 5.592$$

Step 5

$$(y_2)_2 = y_1 + \frac{0.1}{2}[(y_2')_1 + y_1'] = 1.685$$

Step 6

$$(y_2')_2 = 2(y_2)_2^2 + x_2 = 5.878$$

Step 5 (repeat)

$$(y_2)_3 = y_1 + (0.05)[(y_2')_2 + y_1'] = 1.699$$

Step 6 (repeat)

$$(y_2')_3 = 2(y_2)_3^2 + x_3 = 5.974$$

and so forth. This procedure may be programmed for a computer. A discussion of the truncation error of this process may be found in Refs. 6 and 7.

Equations of the Second and Higher Orders. The substitution of $dy/dx = y'$ reduces a second-order equation of the form

$$\frac{d^2y}{dx^2} + P\frac{dy}{dx} + Qy = f(x)$$

to the two first-order equations

$$\frac{dy}{dx} = y' \qquad y'' = f(x) - Py' - Qy$$

A similar procedure will reduce an equation of any order to a system of first-order equations, and these can be solved numerically by the method already given or by those given in Refs. 5 and 6.

6. Numerical Solution of Partial Differential Equations.
The techniques will be introduced by an example. Consider the typical linear-diffusion problem with given initial and boundary conditions:

$$\frac{\partial z}{\partial t} = \frac{\partial^2 z}{\partial x^2} \qquad 0 < x < 1, 0 < t \le T$$

$$z(x,0) = f(x) \qquad 0 < x < 1$$

$$z(0,t) = g(t) \qquad 0 < t \le T$$

$$z(1,t) = h(t) \qquad 0 < t \le T$$

A finite-difference analog for this problem is developed by introducing a net whose mesh points are denoted by $x_i = ih$, $t_j = jk$, where $i = 0, 1, 2, \ldots, M$; $j = 0, 1, \ldots, N$; $h = \Delta x = 1/M$; and $k = \Delta t = T/N$. The boundaries are specified by $i = 0$ and $i = M$ and any "false" boundaries by $i = -1, -2, \ldots, i = M + 1, M + 2, \ldots$, and so forth. The initial is denoted by $j = 0$, and the discrete approximation at $x_i = \pm h$, $t_j = jk$ is designated $Z_{i,j}$.

If an approximate solution $Z_{i,j}$ is assumed to be known at all the mesh points up to the t_j, a method must be specified to advance the solution to time t_{j+1}.

The value of $Z_{i,j+1}$ at $x = 0$ and $x = 1$ should be selected as those boundary conditions specified above, that is,

$$Z_{0,j+1} = g(t_{j+1}) \qquad Z_{M,j+1} = h(t_{j+1})$$

At other points $0 < i < M$, the partial differential equation will be replaced by some difference equation. The simplest replacement consists in approximating the space derivative by a centered second difference and the time derivative by a forward difference at (x_i, t_j).

Based on formulas given in Sec. 2, Interpolation and Finite Differences, earlier, the partial derivatives may be approximated by

$$\frac{\partial z}{\partial t} = \frac{1}{k}(Z_{i,j+1} - Z_{i,j})$$

$$\frac{\partial^2 z}{\partial x^2} = \frac{1}{h^2}(Z_{i+1,j} + Z_{i-1,j} - 2Z_{i,j})$$

The resulting difference equation is now

$$\frac{1}{k}\,(Z_{i,j+1} - Z_{i,j}) = \frac{1}{h^2}\,(Z_{i+1,j} - 2Z_{i,j} + Z_{i-1,j}) \qquad i = 1, \ldots, M - 1$$

Upon solving for $Z_{i,j+1}$, we obtain the explicit equation for "marching" ahead in time:

$$Z_{i,j+1} = rZ_{i-1,j} + (1 - 2r)\,Z_{i,j} + rZ_{i+1,j} \qquad i = 1, \ldots, M - 1$$

where $r = k/h^2 = \Delta t/(\Delta x)^2$.

For a discussion of error and convergence and consistency of this computational procedure see Refs. 7 and 8.

NOTE ON SETS AND BOOLEAN ALGEBRA

Sets and Elements

The concept of a set appears throughout modern mathematics. A set is a well-defined list or collection of objects and is generally denoted by capital letters, A, B, C, . . . The objects composing the set are called elements and are denoted by lower case letters, a, b, x, y, . . . The notation

$$x \in A$$

is read "x is an element of A" and means that x is one of the objects composing the set A.

There are two basic ways to describe a set. The first way is to list the elements of the set.

$$A = \{2, 4, 6, 8, 10\}$$

This often is not practical for very large sets.

The second way is to describe properties that determine the elements of the set.

$$A = \{\text{even numbers from 2 to 10}\}$$

This method is sometimes awkward since a single set may sometimes be described in several different ways.

In describing sets, the symbol : is read "such that." The expression

$$B = \{x : x \text{ is an even integer, } x > 1, x < 11\}$$

is read "B equals the set of all x such that x is an even integer, x is greater than 1, and x is less than 11."

Two sets, A and B, are equal, written $A = B$, if they contain exactly the same elements. The sets A and B above are equal. If two sets, X and Y, are not equal, it is written $X \neq Y$.

Subsets. A set C is a subset of a set A, written $C \subseteq A$, if each element in C is also an element in A. It is also said that C is contained in A. Any set is a subset of itself. That is, $A \subseteq A$ always. A is said to be an "improper subset of itself." Otherwise, if $C \subseteq A$ and $C \neq A$, then C is a proper subset of A. Two theorems are important about subsets:

$$\text{If } X \subseteq Y \text{ and } Y \subseteq X, \text{ then } X = Y \qquad (2.1.1)$$

(Transitivity)

$$\text{If } X \subseteq Y \text{ and } Y \subseteq Z, \text{ then } X \subseteq Z \qquad (2.1.2)$$

Universe and Empty Set. In an application of set theory, it often happens that all sets being considered are subsets of some fixed set, say integers or vectors. This fixed set is called the universe and is sometimes denoted U.

It is possible that a set contains no elements at all. The set with no elements is called the empty set or the null set and is denoted ϕ.

Set Operations. New sets may be built from given sets in several ways. The union of two sets, denoted $A \cup B$, is the set of all elements belonging to A or to B, or to both.

$$A \cup B = \{x : x \in A \text{ or } x \in B\}$$

The union has the properties:

$$A \subseteq A \cup B \text{ and } B \subseteq A \cup B \qquad (2.1.3)$$

The intersection is denoted $A \cap B$ and consists of all elements, each of which belongs to both A and B.

$$A \cap B = \{x : x \in A \text{ and } x \in B\}$$

The intersection has the properties

$$A \cap B \subseteq A \text{ and } A \cap B \subseteq B \qquad (2.1.4)$$

If $A \cap B = \emptyset$, then A and B are called disjoint.

In general, a union makes a larger set and an intersection makes a smaller set.

The complement of a set A is the set of all elements in the universe set that are not in A. This is written

$$A = \{x : x \in U, x \notin A\}$$

The difference of two sets, denoted $A - B$, is the set of all elements that belong to A but do not belong to B.

Algebra on Sets. The operations of union, intersection, and complement obey certain laws known as Boolean algebra. Using these laws, it is possible to convert an expression involving sets into other equivalent expressions. The laws of Boolean algebra are given in Table 3.9.

Venn Diagrams. To give a pictorial representation of a set, Venn diagrams are often used. Regions in the plane are used to correspond to sets, and areas are shaded to indicate unions, intersections, and complements. Examples of Venn diagrams are given in Fig. 3.23.

In some engineering reference books, different notation related to Boolean algebra has been adopted especially when used to design combination logic circuit and control system logic simplification.

TABLE 3.9   Laws of Boolean Algebra

---

1. Idempotency
   $A \cup A = A$ $A \cap A = A$

2. Associativity
   $(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$

3. Commutativity
   $A \cup B = B \cup A$ $A \cap B = B \cap A$

4. Distributivity
   $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

5. Identity
   $A \cup \emptyset = A$ $A \cap U = A$
   $A \cup U = U$ $A \cap \emptyset = \emptyset$

6. Complement
   $A \cup \sim A = U$ $A \cap \sim A = \emptyset$
   $\sim(\sim A) = A$
   $\sim U = \emptyset$
   $\sim \emptyset = U$

7. DeMorgan's laws
   $\sim(A \cup B) = \sim A \cap \sim B$ $\sim(A \cap B) = \sim A \cup \sim B$
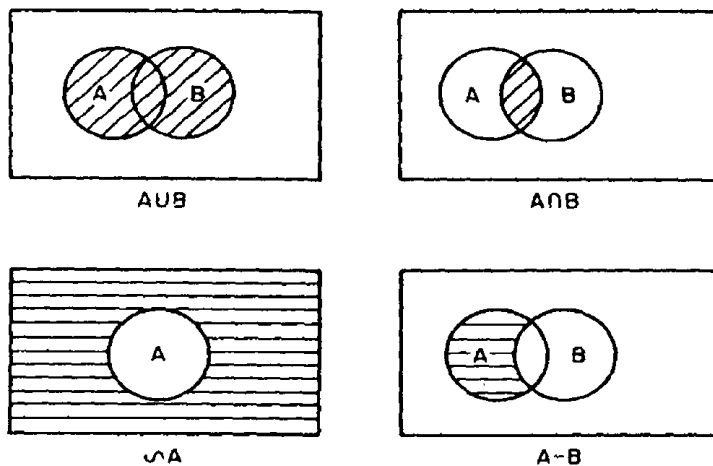
FIGURE 3.23   Venn diagrams.*

The "AND" function is represented by a dot (.), the "OR" function by a + symbol. The inverse (NOT) function is represented by a bar above the signal, so NOT A is represented by A.

Using the Boolean notation the circuit shown on Figure 3.24 can be represented by:

---

*Source: Ref. 1.
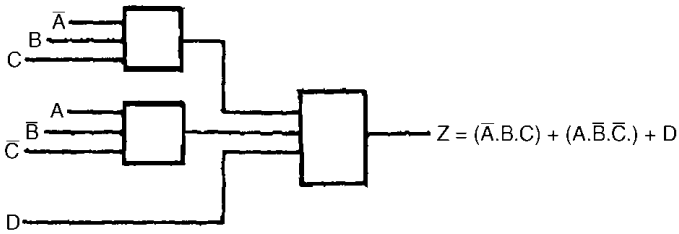
FIGURE 3.24 Circuit represented by Boolean expression.

$$Z = (\overline{A}.B.C) + (A.\overline{B}.\overline{C}.) + D$$

Boolean algebra allows complex expressions for circuits to be written and presented in a simple and concise form.

Boolean algebra can be used to minimize expressions, but it relies on intuition and there is no logical procedure. It is easy to make errors on double or triple inversions, and also in the swapping of the 'dot' and 'plus' symbols.

DIGITAL COMPUTERS*

Computers are machines used for automatically processing information represented by mechanical or electrical means. They may be classified as analog or digital according to the techniques used to represent and process the information. Information in an analog computer is represented as continuous quantities that are physically measurable. This information is processed by components that are interconnected to form an analogous model of the problem to be solved. Digital computers, on the other hand, represent information as discrete physical states that are encoded into symbolic formats. Digital information is processed by sequences of operational steps that are preplanned to solve the given problem.

Binary Notation

Information in a digital computer is represented by strings of digits that assume one of the two values 0 or 1. These units of information are called bits. The term 'bit' is a contraction for binary digit. Bits are used internally in the computer to represent both numerical and non-numerical information.

In order to achieve efficiency in dealing with information, a fixed number of bits are grouped together and referenced as a discrete unit. These units are used to encode and format the information that can be processed by the computer. Units of 8 bits are common and are called bytes (Figure 3.25). Bytes are used to encode the basic symbolic characters that provide the computer with input-output information such as the alphabet, decimal digits, punctuation marks, and special characters. The byte size is not universally adopted on all machines but is perhaps the most popular basic unit of information.

Larger bit groups are organized into units called words (and sometimes into smaller units called half words or larger units called double words). These units are used to encode the basic instruction repertoire of the machine and format the
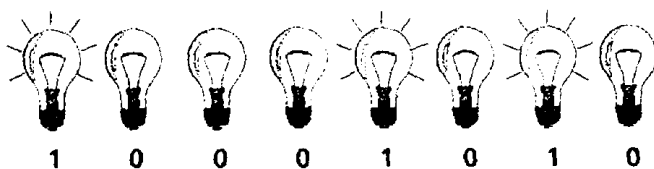
'1' indicates when a bulb is 'on' – '0' indicates when a bulb is 'off'
Code for this part of 'the message' is 10001010

FIGURE 3.25   Illustration of principle of operation.

numerical data. Common word sizes used on various commercially available computers are 16, 24, 32, 36, 48, and 60 bits.

Numerical information processed by a computer is represented in the binary numbering system. The binary numbering system uses a position notation analogous to that used in decimal numbers. For example, the decimal number 596.37 represents the value $5 \times 10^2 + 9 \times 10^1 + 6 \times 10^0 + 3 \times 10^{-1} + 7 \times 10^{-2}$. The value assigned to any of the 10 possible digits 0 through 9 depends on its position relative to the decimal point; i.e., zero and positive exponents of 10 are assigned to digits appearing to the left of the decimal point and negative exponents of 10 to the right. In a similar manner, the binary number 1011.011 represents $1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 0 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3}$. The radix (base) of this system is 2. Each position of a binary number is occupied by one of the two possible digits 0 or 1. The radix point in this system serves a purpose similar to that of the decimal point in the decimal system.

The functional operators available in the computer for setting up the solution of a problem are encoded into the words of the machine. To handle numerical calculations, the instruction repertoire usually includes the four rules of arithmetic, i.e., $+$, $-$, $\times$, and $\div$. These instructions operate on data encoded in the binary system. This, however, is not a serious operational problem, since the user specifies information for the problem in the decimal system, or through mnemonics, and the computer is used to convert these formats into its own internal representations.

Formats for Numerical Data

Three different formats are used to represent numerical information internal to the computer: fixed-point, floating-point and encoded-decimal. These formats are described in this section.

a) A word in **fixed-point** format is given as a string of 0's and 1's representing a binary number. The radix point is not explicitly given but is implied by the program processing the data to be at a fixed position in the word, e.g., immediately to the right of the word so that the number represented would be an integer. To represent algebraic numbers, several alternate forms are used. Most often, 1's or 2's complement forms are adopted because they lead to a reduction and simplification of the electronics needed to perform the arithmetic operations.

b) **Floating-point** format is a mechanized version of the scientific notation, i.e., $\pm M \times 10^{\pm E}$, where $\pm M$ and $\pm E$ represent the signed mantissa and signed exponent of the number, respectively. Using a machine word to represent numbers in this format, a large range of numbers is possible. The signed mantissa

and signed exponent are explicitly given in each word. The exponent, however, is implied as a power of 2 or 16 rather than 10. The radix point is implied to the left of the mantissa. After each operation the exponent is adjusted so that the most significant digit is other than zero; i.e., the mantissa M is maintained so that its value is in the range $0.1 \leq M < 1$, where these bounds are in the radix of the system. This is termed normalizing. Zero is treated as a special case in this notation, e.g., zero mantissa and zero exponent.

Often greater precision is needed in the calculations than is possible with one word. In this case, two words are used to represent the number. Since the exponent need not be defined again, the added word is appended to the mantissa to extend the precision of the number. Such a representation is called double precision, although extended precision would be a more accurate term, since the precision may be more than doubled.

The range of possible numbers in floating-point notation depends on the number of bits designated to the exponent and implied base of the system. For example, if 7 bits are used for the signed exponent and if 16 is implied as the base, then $16^{-64} \leq$ range $\leq 16^{63}$.

The precision of a floating-point number depends on the number of bits used for the unsigned mantissa and also depends on the implied base. If m is the number of bits in the unsigned mantissa, the precision expressed as equivalent number of decimal digits d is

$$0.301(m - 1) \leq d \leq 0.301m \text{ when implied base is 2}$$

$$0.301(m - 4) \leq d \leq 0.301m \text{ when implied base is 16}$$

For example, if 24 bits of a 32-bit word are used for the unsigned mantissa, 6.02 to 7.22 decimal digits can be represented when the implied base of the system is 16. The fractional component of 6.02 indicates that a fraction of 7-digit decimal numbers cannot be represented with a 24-bit mantissa. In a similar manner, 7.55 indicates that some 8-digit decimal numbers cannot be represented. A double-precision number in this system, on the other hand, can represent numbers that can be expressed with 15.65 to 16.85 equivalent decimal digits.

c) **Encoded-decimal** representation is usually not used in scientific calculations. However, it is supplied in many computers as a convenience in commercial applications. Table 3.10 gives some typical schemes used to encode the decimal digits in which each decade is represented with 4 bits.

Formats for Non-numerical Data

A large variety of codes is used for representing the alphabet, digits, punctuation marks, and special symbols internal to the computer. The most popular ones are the 7-bit ASCII code (see Table 3.11) and the 8-bit EBCDIC code. These are available in the literature in many different sources and are not presented here.

The simplest form of representing data values is the logical type, also called Boolean. Logical data have one of two values, either true or false. These values may be conveniently encoded by a single bit.

TABLE 3.10 Schemes for Encoding Decimal Digits

Decimal digit	BCD	Excess-3	4221 code
0	0000	0011	0000
1	0001	0100	0001
2	0010	0101	0010
3	0011	0110	0011
4	0100	0111	0110
5	0101	1000	1001
6	0110	1001	1100
7	0111	1010	1101
8	1000	1011	1110
9	1001	1100	1111

TABLE 3.11 Partial List of ASCII Code, in which each Bit Pattern has been Extended with a 0 on its Left to Produce the Eight-bit Pattern Commonly Used Today

Symbol	ASCII	Symbol	ASCII	Symbol	ASCII
(space)	00100000	?	00111111	∧	01011110
!	00100001	@	01000000	—	01011111
"	00100010	A	01000001	a	01100001
#	00100011	B	01000010	b	01100010
$	00100100	C	01000011	c	01100011
%	00100101	D	01000100	d	01100100
&	00100110	E	01000101	e	01100101
'	00100111	F	01000110	f	01100110
(	00101000	G	01000111	g	01100111
)	00101001	H	01001000	h	01101000
*	00101010	I	01001001	i	01101001
+	00101011	J	01001010	j	01101010
,	00101100	K	01001011	k	01101011
-	00101101	L	01001100	l	01101100
.	00101110	M	01001101	m	01101101
/	00101111	N	01001110	n	01101110
0	00110000	O	01001111	o	01101111
1	00110001	P	01010000	p	01110000
2	00110010	Q	01010001	q	01110001
3	00110011	R	01010010	r	01110010
4	00110100	S	01010011	s	01110011
5	00110101	T	01010100	t	01110100
6	00110110	U	01010101	u	01110101
7	00110111	V	01010110	v	01110110
8	00111000	W	01010111	w	01110111
9	00111001	X	01011000	x	01111000
:	00111010	Y	01011001	y	01111001
;	00111011	Z	01011010	z	01111010
<	00111100	[	01011011	{	01111011
=	00111101	\	01011100	}	01111101
>	00111110	]	01011101		

Structured Data Types

The above types of numerical and non-numerical data formats are recognized and manipulated by the hardware operations of the computer (Figure 3.26). Other more complex data structures may be programmed into the computer by building upon these primitive data types. Such complex structures might include arrays, defined as ordered lists of elements of identical type; sets, defined as unordered lists of elements of identical type; records, defined as ordered lists of elements that need not be of the same type; files, defined as sequential collections of identical records; and databases, defined as organized collections of different record or file types.

Digital-Computer Components

Figure 3.27 shows a schematic that can be used to discuss the hardware configuration of a digital-computer system.

The principal part of the system is composed of the central processing unit and the memory of the system. The memory is organized into words (bytes, half words, and/or double words) that can be located by an address. The central processing

Direction of the flow is shown by the arrows
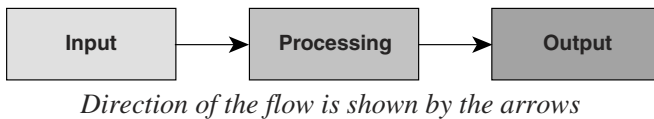
FIGURE 3.26   The most simple machine representation.
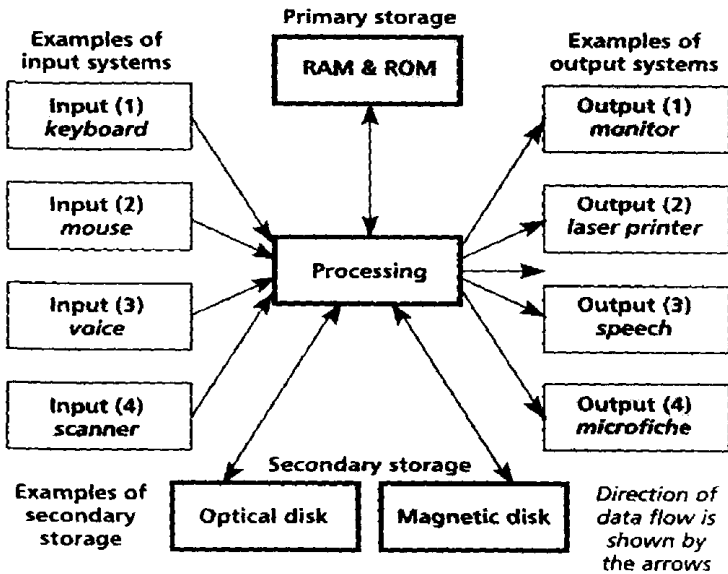
RAM (Random Access Memory); ROM (Read Only Memory)

FIGURE 3.27   Computer system organization—an expansion of the basic ideas.

unit makes available a repertoire of instructions that are used to set up the problem solutions. The general format of each instruction is as follows:

name: operator, operand(s)

The name designates an address that contains the operator and one or more operands. The operations permitted by the hardware of the central processing unit are encoded into the operator. The operand(s) refer to other named words that are used in the operation and may refer to either data or other instructions by their address.

The instructions and data are organized into sequences that, when executed in order, will produce the problem solution. These sequences are variously called programs, subprograms, routines, subroutines, functions, etc. They are inserted into the memory prior to machine execution.

Software Systems

Computer software is organized into systems that provide the operating and functional facilities for computer users. Three systems generally provided are (1) operating systems, which consist of components for the control and operation of the computer hardware and software; (2) program-preparation systems, which consist of components for preparing and modifying programs for computer execution; and (3) data-management systems, which consist of components for generating, storing, updating, accessing, retrieving, editing, revising, and maintaining the information on the computer files. Data-management systems are of major concern to data processing rather than scientific or engineering applications. The components that support file management and input-output control, which are important parts of data-management systems, are usually included in the design of operating and data-preparation systems.

Generations of Computers

Inside each computer there are tiny devices called transistors, which carry out these lightning-speed switching operations. Today, millions of these transistors can easily be packed into a silicon chip—but this has obviously not always been the case. In the early days of electronic computers, valves, as shown in Figure 3.28 (a), were

(a) Valve

(b) Transistor

(c) Silicon Chip
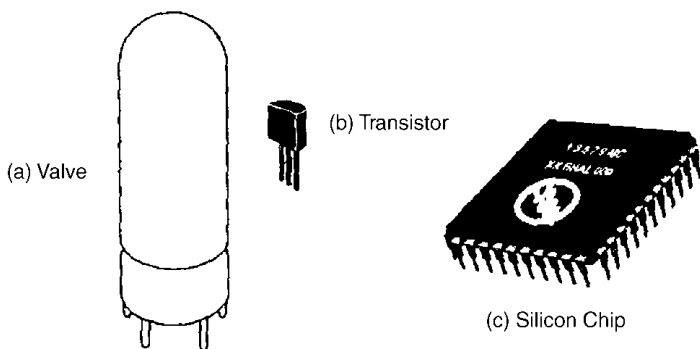
FIGURE 3.28   Generations of computer devices.

used as the switching device. They were quite large, consumed lots of energy, and thus the computers built up using these components took up the space of a very large room—they became known as the first generation of computers. Later on the transistor was invented, and this smaller and more energy-efficient device led to the second generation of computers. They were more powerful, consumed less energy, and took up less space than the first generation machines.

A real breakthrough came in the 1960s when several transistors were put into a single package called an integrated circuit or silicon chip. These became known as the third-generation machines and were smaller and more powerful again. In the 1970s and 1980s it became possible to pack more and more transistors into a single silicon chip, until at the end of the 1990s, it was possible to get well over 5,000,000 in a single-chip device. Computers making use of these latest chips are called fourth-generation computers. Fifth-generation computers have been on the drawing board for a number of years, but are discussed elsewhere [13].

The fundamental point about digital computers: these machines are very fast! A high-speed fiber-optic (see chapter 11) link could transmit about 100 million codes (see Table 3.11) each second. At this rate we would be able to send entire textual contents of this handbook in less than a few seconds.

CALCULATORS

Calculators, while not computers by our definition, use much of the same technology as computers in construction and are used in performing calculations in much the same manner. The important distinction between computers and calculators is that calculators involve human intervention between the computational steps. This distinction is lost with some calculators labeled as "programmable," that is, calculators that allow new functions to be programmed and the results of intermediate steps to be stored for later recall and use. Programs may also be stored on plastic magnetic strips that can be loaded in the machine when needed.

Calculators come in desktop or hand-held models, with or without paper-print capabilities for recording the computational results, and they have a wide variety of function keys. In addition to the four rules of arithmetic, i.e., $+$, $-$, $\times$, and $\div$, special function keys may include square root, square, power, factorial, reciprocal, conversion to or from SI (metric) units, conversion between degrees and radian, logarithm and exponential (both base 10 and e), elementary statistical regression, percentage, sign reversal, generation of SI constants (e.g., 2.54 cm/in), and trigonometric functions and inverse trigonometric functions (both in radian and degree). Calculators are classified as scientific or business machines depending on the special functions available on the keyboard.

Commonly Used Symbols for a Calculator

Ways of using a calculator to solve various types of problems are shown below.

---

* *Source:* Refs. 1, 12, 13, 15.

Symbols used

△ a	input data, i.e. input the value of a	◯	perform operation shown in the circle
⊡ M	store in the memory	M+	add to the data in the memory
⊡ MR	recall contents of the memory to the display	⊡ D	the value displayed

Examples*:

$$I = PlAN$$
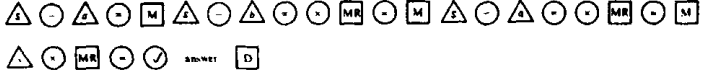
$$E = \frac{Wl}{Ax}$$

$$v = \sqrt{(2gh)}$$

$$v = \sqrt{(u^2 + 2as)}$$
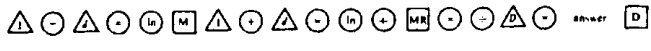
$$l = \frac{\pi h}{4a^2}(D^2 - d^2)$$

$$A = \sqrt{[s(s-a)(s-b)(s-c)]}$$

$$a = \frac{b \sin A}{\sin B}$$

$$a = \sqrt{(b^2 + c^2 - 2bc \cos A)}$$

$$A = \arctan(\theta + \varphi)$$
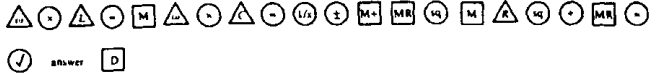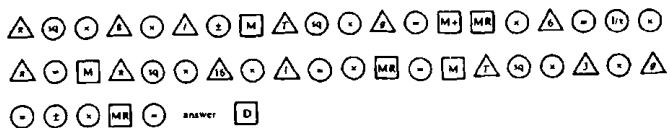
$$P = D + \frac{\ln(1+d)}{\ln(1-d)}$$

$$p = P(1 - e^{-t/\tau})$$

$$M = \pi(R^4 - r^4)$$

$$d = (12M/5b)^{1/3}$$

$$Z = \sqrt{[R^2 + (\omega L - 1/\omega C)]^2}$$

$$m = \frac{\pi(16\pi^2 l - 3gT^2)}{6(gT^2 - 8\pi^2 l)}$$

*Source: Ref. 14.

REFERENCES*

1. E. A. Avallone and T. Baumeister III (eds.), *Mark's Standard Handbook for Mechanical Engineers,* 9th ed., McGraw-Hill, New York, 1987, sec. 2.

2. K. Gieck, *Engineering Formulas,* 4th ed., McGraw-Hill, New York, 1983.

3. H. A. Rothbart, *Mechanical Design and Systems Handbook,* 2d ed., McGraw-Hill, New York, 1985, part 1.

4. R. H. Perry (ed.), *Engineering Manual,* 3d ed., McGraw-Hill, New York, 1976.

5. W. Flugge (ed.), *Handbook of Engineering Mechanics,* McGraw-Hill, New York, 1962, part 1.

6. G. Dahlquist and A. Bjorck, *Numerical Methods,* Prentice-Hall, Englewood Cliffs, N.J., 1974.

7. R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Perry's Chemical Engineer's Handbook,* 6th ed., McGraw-Hill, New York, 1984, sec. 2.

8. S. V. Patankar, *Numerical Heat Transfer and Fluid Flow,* Hemisphere / McGraw-Hill, New York, 1980.

9. O. W. Eshbach and M. Souders, *Handbook of Engineering Fundamentals,* John Wiley & Sons, New York, 1975.

10. W. M. Rohsenow, J. P. Hartnett, and E. N. Ganić (eds.), *Handbook of Heat Transfer Fundamentals,* 2d ed., McGraw-Hill, New York, 1985, chap. 2.

11. O. C. Zienkiewicz, *The Finite Element Method in Structural and Continuum Mechanics,* McGraw-Hill, New York, 1967.

12. D. Christiansen, *Electronics Engineers' Handbook,* 4th ed., McGraw-Hill, 1997.

13. R. Bradley, *Understanding Computer Science,* 4th ed., Stanley Thornes Pub.Ltd, 1999.

14. R. Timings and T. May, *Newnes Mechanical Engineer's Pocket Book,* 2d ed., Newnes, 2001.

15. J. G. Brookshear, *Computer Science—An Overview,* 6th ed., Addison-Wesley, 2000.

---

*Those references listed above but not cited in the text were used for comparison between different data sources, clarification, clarity of presentation, and, most important, reader's convenience when further interest in the subject exists.

CHAPTER 4
APPLIED CHEMISTRY

COMMON DEFINITIONS

Elements. There are a very large number of different substances in existence, each substance containing one or more of a number of basic materials called elements. An ''element'' is a substance which cannot be separated into anything simpler by chemical means. There are 92 naturally ocurring elements and 13 others, which have been artificially produced.

Some examples of common elements with their symbols are: Hydrogen (H), Helium (He), Carbon (C), Nitrogen (N), Oxygen (O), Sodium (Na), Magnesium (Mg), Aluminium (Al), Silicon (Si), Phosphorus (P), Sulphur (S), Potassium (K), Calcium (Ca), Iron (Fe), Nickel (Ni), Copper (Cu), Zinc (Zn), Silver (Ag), Tin (Sn), Gold (Au), Mercury (Hg), Lead (Pb) and Uranium (U).

Elements are made up of very small parts called atoms. An ''atom'' is the smallest part of an element that can take part in a chemical change and retain the properties of the element. Each of the elements has a unique type of atom.

Atoms. In atomic theory, a model of an atom can be regarded as a miniature solar system. Atoms contain a dense nucleus of positively charged *protons* and uncharged *neutrons*, surrounded by negatively charged orbital *electrons* that occupy discrete energy levels and orbital configurations (see Fig. 4.1). In a stable atom, the positively charged protons and the negatively charged electrons are equal. Each type of atom, or *element,* is assigned a *mass number* based on the number of protons and neutrons contained in its nucleus. Mathematically these relationships are expressed as:

A = mass number

Z = number of positively charged protons

$A - Z$ = number of uncharged neutrons

Each element is also assigned an *atomic number* equal to Z. Gain or loss of electrons results in ionized atoms with a net charge. Chemical reactions are concerned only with changes in electrical structure, while nuclear reactions involve changes in the constitution of the nucleus. *Isotopes* are atoms having the same atomic number, the same number of electrons, and the same chemical reactions, but a different mass number because their nuclei contain different numbers of neutrons.

The arrangement of the elements in order of their atomic number is known as the periodic table (Chapter 2, Table 2.2).
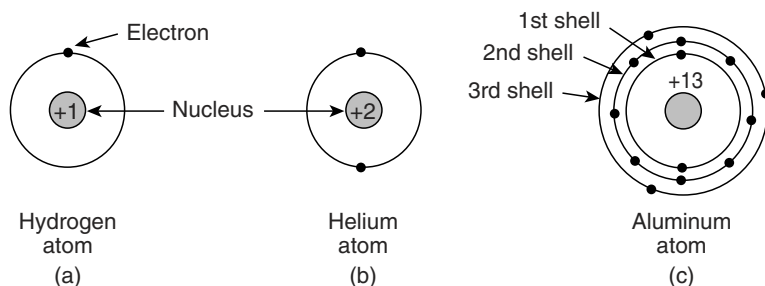
FIGURE 4.1   Atom structure.

The simplest atom is hydrogen, which has 1 electron orbiting the nucleus and 1 proton in the nucleus. The atomic number of hydrogen is thus 1. The hydrogen atom is shown diagrammatically in Figure 4.1(a). Helium has 2 electrons orbiting the nucleus, both of them occupying the same shell at the same distance from the nucleus as shown in Figure 4.1(b).

The first shell of an atom can have up to 2 electrons only, the second shell can have up to 8 electrons only and the third shell up to 18 electrons only. Thus an aluminium atom, which has 13 electrons orbiting the nucleus, is arranged as shown in Figure 4.1(c).

Molecules.   A molecule consists of a group of two or more atoms held together through chemical bonds (represented diagrammatically as geometric arrangements). Chemical bonds in turn are based on the electrical configurations of the individual atoms involved, and occur when electrons in one atom occupy vacant orbital sites in another atom. Chemical bonds between atoms result in a joint electrical configuration of greater stability, or lower energy, than that possessed by the individual atoms. The total negative charge of a molecule's combined electrons equals the total positive charge of its combined protons.

Compounds.   When elements combine chemically (chemical reaction) their atoms interlink to form molecules of a new substance called a compound. A compound is a new substance containing two or more elements chemically combined so that their properties are changed.

For example, the elements hydrogen and oxygen are quite unlike water, which is the compound they produce when chemically combined.

The components of a compound are in fixed proportion and are difficult to separate.

Avogadro's Number.   Avogadro's number is the number of molecules (6.023×10^{23}) in 1 g · mol of a substance.

Gram Equivalent Weight.   This involves two kinds of reactions:

- *Nonredox reaction:* the mass in grams of a substance equivalent to 1 g · atom of hydrogen, 0.5 g · atom of oxygen, or 1 g · ion of the hydroxyl ion. It can be determined by dividing the molecular weight of the substance by the number of hydrogen or oxygen atoms or hydroxyl ions (or their equivalent) supplied or required by the molecule in a given reaction.

- *Redox reaction:* the molecular weight in grams divided by the change in oxidation state.

Molality.   Expressed mathematically by the symbol m, molality is defined as:

$$m = \frac{\text{number of gram moles of solute}}{\text{number of kilograms of solvent}}$$

Molarity.   Expressed mathematically by the symbol M, molarity is defined as:

$$M = \frac{\text{number of gram moles of solute}}{\text{number of liters of solution}}$$

Normality.   Expressed mathematically as N, normality is defined as:

$$N = \frac{\text{number of gram equivalents of solute}}{\text{number of liters of solution}}$$

Oxidation.   Oxidation is the loss of electrons by an atom or group of atoms.

Reduction.   Reduction is the gain of electrons by an atom or group of atoms.

Ion Product of Water.   Ion product of water (K_w) is the product of the hydrogen ion (H^+) and hydroxyl ion (OH^-) concentrations in gram-ions per liter:

$$K_w = [H^+][OH^-]$$

pH Values.   pH value (see Fig. 4.1*a*) is expressed as the negative logarithm (base 10) of the hydrogen ion concentration in gram-ions per liter:

$$pH = -\log [H^+]$$

See Table 4.1 for establishing pH values by using suitable indicators.

Generally the pH scale (pH meaning 'the potency of hydrogen') represents, on a scale from 0 to 14, degrees of acidity and alkalinity. 0 is strongly acidic, 7 is neutral and 14 is strongly alkaline. Some average pH values include: concentrated hydrochloric acid, HCl 1.0, lemon juice 3.0, milk 6.6, pure water 7.0, sea water 8.2, concentrated sodium hydroxide, NaOH 13.0.

Acids have the following properties:

• Almost all acids react with carbonates and bicarbonates (a carbonate being a compound containing carbon and oxygen—an example being sodium carbonate, i.e., washing soda).

• Dilute acids have a sour taste; examples include citric acid (lemons), acetic acid (vinegar) and lactic acid (sour milk).

• Acid solutions turn litmus paper red, methyl orange red and phenolphthalein colorless, as mentioned above.

$[H^+]$	1	10^{-1}	10^{-2}	•••	10^{-7}	•••	10^{-12}	10^{-13}	10^{-14}
pH value	0	1	2		7		12	13	14
		Acid			Neutral		Alkaline		

FIGURE 4.1*a*   pH values.

TABLE 4.1 List of Acid-Base Indicators

Indicator	pH-Range	Color change	
		From	To
thymol blue (benz.)	1·2...2·8	red	yellow
p-dimethylamino-	2·9...4·0	red	orange-yellow
azobromophenolblue	3·0...4·6	yellow	red-violet
congo red	3·0...4·2	blue-violet	red-orange
methyl orange	3·1...4·4	red	yellow-(orange)
brom cresol green	3·8...5·4	yellow	blue
methyl red	4·4...6·2	red	(orange)-yellow
litmus	5·0...8·0	red	blue
bromocresol purple	5·2...6·8	yellow	purple
bromophenol red	5·2...6·8	orange yellow	purple
bromothymol blue	6·0...7·6	yellow	blue
phenol red	6·4...8·2	yellow	red
neutral red	6·4...8·0	(blue)-red	orange-yellow
cresol red	7·0...8·8	yellow	purple
meta cresol purple	7·4...9·0	yellow	purple
thymol blue	8·0...9·6	yellow	blue
phenolphthalein	8·2...9·8	colorless	red-violet
alizarin yellow 66	10·0...12·1	light-yellow	light brown-yellow

Source: From Gieck.[1] Data shown in this table can also be found in Refs. 2 through 5.

Alkalis have the following properties:

- Alkalis neutralize acids to form a salt and water only.
- Alkalis have little effect on metals.
- Alkalis turn litmus paper blue, methyl orange yellow and phenolphthalein pink, as mentioned above.
- Alkalis are slippery when handled; strong alkalis are good solvents for certain oils and greases.

Note. A list of common chemicals is given in Table 4.2. A list of chemical reactions for the preparation of some chemicals is given in Table 4.3. A procedure for the preparation of freezing mixtures in a certain temperature range is given in Tables 4.4, 4.4a and 4.4b. A list of common laboratory solutions and reagents is given in Table 4.5.

STOICHIOMETRY

Stoichiometry is the theory of the proportions in which chemical species combine with one another. The stoichiometric equation of a chemical reaction is a statement of the relative number of molecules or moles of reactants and products that partic-

TABLE 4.2 List of Common Chemicals

Trade Name	Chemical	Chemical formula
acetone	acetone	$(CH_3)_2 \cdot CO$
acetylene	acetylene	C_2H_2
ammonia	ammonia	NH_3
ammonium (hydroxide of)	ammonium hydroxide	NH_4OH
aniline	aniline	$C_6H_5 \cdot NH_2$
bauxite	hydrated aluminum oxides	$Al_2O_3 \cdot 2\ H_2O$
bleaching powder	calcium hypochlorite	$CaCl(OCl)$
blue vitriol	copper sulfate	$CuSO_4 \cdot 5\ H_2O$
borax	sodium tetraborate	$Na_2B_4O_7 \cdot 10\ H_2O$
butter of zinc	zinc chloride	$ZnCl_2 \cdot 3\ H_2O$
cadmium sulfate	cadmium sulfate	$CdSO_4$
calcium chloride	calcium chloride	$CaCl_2$
carbide	calcium carbide	CaC_2
carbolic acid	phenol	C_6H_5OH
carbon dioxide	carbon dioxide	CO_2
carborundum	silicon carbide	SiC
caustic potash	potassium hydroxide	KOH
caustic soda	sodium hydroxide	$NaOH$
chalk	calcium carbonate	$CaCO_3$
cinnabar	mercuric sulfide	HgS
ether	di-ethyl ether	$(C_2H_5)(_2O)$
fixing salt or hypo	sodium thiosulfate	$Na_2S_2O_3 \cdot 5\ H_2O$
glauber's salt	sodium sulfate	$Na_2SO_4 \cdot 10\ H_2O$
glycerine or glycerol	glycerine	$C_3H_5(OH)_3$
graphite	crystalline carbon	C
green vitriol	ferrous sulfate	$FeSO_4 \cdot 7\ H_2O$
gypsum	calcium sulfate	$CaSO_4 \cdot 2\ H_2O$
heating gas	propane	C_3H_8
hydrochloric acid	hydrochloric acid	HCl
hydrofluoric acid	hydrofluoric acid	HF
hydrogen sulfide	hydrogen sulfide	H_2S
iron chloride	ferrous chloride	$FeCl_2 \cdot 4\ H_2O$
iron sulfide	ferrous sulfide	FeS
laughing gas	nitrous oxide	N_2O
lead sulfide	lead sulfide	PbS
limestone	calcium carbonate	$CaCO_3$
magnesia	magnesium oxide	MgO
marsh gas	methane	CH_4
minimum or red lead	plumbate	$2\ PbO \cdot PbO_2$
nitric acid	nitric acid	HNO_3
phosphoric acid	ortho phosphoric acid	H_3PO_4
potash	potassium carbonate	K_2CO_3
potassium bromide	potassium bromide	KBr
potassium chlorate	potassium chlorate	$KClO_3$
potassium chloride	potassium chloride	KCl
potassium chromate	potassium chromate	K_2CrO_4
potassium cyanide	potassium cyanide	KCN
potassium dichromate	potassium dichromate	$K_2Cr_2O_7$

TABLE 4.2 List of Common Chemicals (*Continued*)

Trade Name	Chemical	Chemical formula
potassium iodide	potassium iodide	KI
prussic acid	hydrogen cyanide	HCN
pyrolusite	manganese dioxide	MnO_2
quicklime	calcium monoxide	CaO
red prussiate of potassium	potassium ferrocyan.	$K_3Fe(CN)_6$
salammoniac	ammonium chloride	NH_4Cl
silver bromide	silver bromide	$AgBr$
silver nitrate	silver nitrate	$AgNO_3$
slaked lime	calcium hydroxide	$Ca(OH)_2$
soda ash	hydrated sodium carb.	$Na_2CO_3 \cdot 10\ H_2O$
sodium monoxide	sodium oxide	Na_2O
soot	amorphous carbon	C
stannous chloride	stannous chloride	$SnCl_2 \cdot 2\ H_2O$
sulfuric acid	sulffuric acid	H_2SO_4
table salt	sodium chloride	$NaCl$
tinstone, tin putty	stannic oxide	SnO_2
trilene	trichlorethylene	C_2HCl_3
urea	urea	$CO(NH_2)_2$
white lead	basic lead carbonate	$2\ PbCO_3 \cdot Pb(OH)_2$
white vitriol	zinc sulphate	$ZnSO_4 \cdot 7\ H_2O$
yellow prussiate of potass.	potass. ferrocyanide	$K_4Fe(CN)_6 \cdot 3\ H_2O$
zinc blende	zinc sulfide	ZnS
zinc or chinese white	zinc oxide	ZnO

Source: From Gieck.[1] Data shown in this table can also be found in Refs. 2 through 5.

TABLE 4.3 List of Chemical Reactions for Preparation of Some Chemicals

To prepare	Use reaction
ammonia	$CO(NH_2)_2 + H_2O \rightarrow 2\ NH_3 + CO_2$
ammonium chloride	$NH_4OH + HCl \rightarrow NH_4Cl + H_2O$
ammonium hydroxide	$NH_3 + H_2O \rightarrow NH_4OH$
cadmium sulfide	$CdSO_4 + H_2S \rightarrow Cds + H_2SO_4$
carbon dioxide	$CaCO_3 + 2\ HCl \rightarrow CO_2 + CaCl_2 + H_2O$
chlorine	$CaOCl_2 + 2\ HCl \rightarrow Cl_2 + CaCl_2 + H_2O$
hydrogen	$H_2SO_4 + Zn \rightarrow H_2 + ZnSO_4$
hydrogen sulfide	$FeS + 2\ HCl \rightarrow H_2S + FeCl_2$
lead sulfide	$Pb(NO_3)_2 + H_2S \rightarrow PbS + 2\ HNO_3$
oxygen	$2\ KClO_3 \rightarrow 3\ O_2 + 2\ KCl$
sodium hydroxide	$Na_2O + H_2O \rightarrow 2\ NaOH$
zinc sulfide	$ZnSO_4 + H_2S \rightarrow ZnS + H_2SO_4$

Source: From Gieck.[1] Data shown in this table can also be found in Refs. 2 through 5.

TABLE 4.4 Preparation of Freezing Mixtures

Drop in temperature from °C	to °C	Mixture (The figures stand for proportions by mass)
+ 10	− 12	4 H_2O + 1 KCl
+ 10	− 15	1 H_2O + 1 NH_4NO_3
+ 8	− 24	1 H_2O + 1 $NaNO_3$ + 1 NH_4Cl
0	− 21	3-0 ice (crushed) + 1 NaCl
0	− 39	1-2 ice (crushed) + 2 $CaCl_2 \cdot 6 H_2O$
0	− 55	1-4 ice (crushed) + 2 $CaCl_2$ + 6 H_2O
+ 15	− 78	1 methyl alcohol + 1 CO_2 solid

Source: From Gieck.[1] Data shown in this table can also be found in Refs. 2 through 6.

TABLE 4.4a Freezing Mixtures Prepared with Ice or Snow

Ammonium nitrate (parts)	Crushed ice or snow in water (parts)	Temperature (°C)
1	0.94	−4
1	1.20	−14
1	1.31	−17

Calcium chloride (parts)	Crushed ice or snow in water (parts)	Temperature (°C)
1	0.49	−20
1	0.61	−39
1	0.70	−55
1	1.23	−22
1	4.92	−4

Solid carbon dioxide with alcohol	−72

Source: Ref. 6

TABLE 4.4b Anti-freeze Mixtures

	Freezing point (°C)			
Concentration (% vol.)	10	20	30	40
Ethanol (ethyl alcohol)	−3.3	−7.8	−14.4	−22.2
Methanol (methyl alcohol)	−5.0	−12.1	−21.1	−32.2
Ethylene glycol	−4.0	−8.9	−15.6	−24.4
Glycerine	−1.7	−5.0	−9.4	−15.6

Source: Ref. 6

TABLE 4.5 List of Laboratory Solutions, Reagents, and Indicators

Unless otherwise stated, the term g per liter *signifies grams of the formula indicated dissolved in water and made up to a liter of solution.*

Acetic acid, $HC_2H_3O_2$—$6N$: 350 mL glacial acetic acid per liter.

Alcohol, amyl, $C_5H_{11}OH$: use as purchased.

Alcohol, ethyl, C_2H_5OH; 95% alcohol, as purchased.

Alizarin, dihydroxyanthraquinone (indicator): dissolve 0.1 g in 100 mL alcohol; pH range yellow 5.5—6.8 red.

Alizarin yellow R, sodium *p*-nitrobenzeneazosalicylate (indicator): dissolve 0.1 g in 100 mL water; pH range yellow 10.1—violet 12.1.

Alizarin yellow GG, salicyl yellow, sodium *m*-nitrobenzeneazosalicylate (indicator); dissolve 0.1 g in 100 mL 50% alcohol; pH range yellow 10.0—12.0 lilac.

Alizarin S, alizarin carmine, sodium alizarin sulfonate (indicator): dissolve 0.1 g in 100 mL water; pH range yellow 3.7—5.2 violet.

Aluminon (qualitative test for aluminum). The reagent consists of 0.1% solution of the ammonium salt of aurin tricarboxylic acid. A bright red precipitate, persisting in alkaline solution, indicates aluminum.

Aluminum chloride, $AlCl_3$—$0.5N$: 22 g per liter.

Aluminum nitrate, $Al(NO_3)_3 \cdot 7.5H_2O$—$0.5N$: 58 g per liter.

Aluminum sulfate, $Al_2(SO_4)_3 \cdot 18H_2O$—$0.5N$: 55 g per liter.

Ammonium acetate, $NH_4C_2H_3O_2$—$3N$: 231 g per liter.

Ammonium carbonate, $(NH_4)_2CO_3 \cdot H_2O$—$3N$: 171 g per liter; for the anhydrous salt: 144 g per liter.

Ammonium chloride, NH_4Cl—$3N$: 161 g per liter.

Ammonium hydroxide, NH_4OH—$15N$: the concentrated solution which contains 28% NH_3; for $6N$: 400 mL per liter.

Ammonium molybdate, $(NH_4)_2MoO_4$—N: dissolve 88.3 g of solid $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ in 100 mL $6N$ NH_4OH. Add 240 g of solid NH_4NO_3 and dilute to one liter. Another method is to take 72 g of MoO_3, add 130 mL of water and 75 mL of $15N$ NH_4OH; stir mechanically until nearly all has dissolved, then add it to a solution of 240 mL concentrated HNO_3 and 500 mL of water; stir continuously while solutions are being mixed; allow to stand 3 days, filter, and use the clear filtrate.

Ammonium nitrate, NH_4NO_3—N: 80 g per liter.

Ammonium oxalate, $(NH_4)_2C_2O_4 \cdot H_2O$—$0.5N$: 40 g per liter.

Ammonium polysulfide (yellow ammonium sulfide), $(NH_4)_2S_x$: allow the colorless $(NH_4)_2S$ to stand, or add sulfur.

Ammonium sulfate, $(NH_4)_2SO_4$—$0.5\ N$: 33 g per liter; saturated: dissolve 780 g of $(NH_4)_2SO_4$ in water and make up to a liter.

Ammonium sulfide (colorless), $(NH_4)_2S$—saturated: pass H_2S through 200 mL of concentrated NH_4OH in the cold until no more gas is dissolved, add 200 mL NH_4OH and dilute with water to a liter; the addition of 15 g of sulfur is sufficient to make the polysulfide.

Antimony pentachloride, $SbCl_5$—$0.5N$: 30 g per liter.

Antimony trichloride, $SbCl_3$—$0.5N$: 38 g per liter.

TABLE 4.5 List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

Aqua regia: mix 3 parts of concentrated HCl and 1 part of concentrated HNO_3 just before ready to use.

Arsenic acid, $H_3AsO_4 \cdot 0.5H_2O$—$0.5N$ ($= \frac{1}{2}H_3AsO_4 \div 5$): 15 g per liter.

Arsenous oxide, As_2O_3—$0.25N$: 8 g per liter for saturation.

Aurichloric acid, $HAuCl_4 \cdot 3H_2O$: dissolve in ten parts of water.

Aurin, see *rosolic acid.*

Azolitmin solution (indicator): make up a 1% solution of azolitmin by boiling in water for 5 minutes; it may be necessary to add a small amount of NaOH to make the solution neutral; pH range red 4.5—8.3 blue.

Bang's reagent (for glucose estimation): dissolve 100 g of K_2CO_3, 66 g of KCl, and 160 of $KHCO_3$ in the order given in about 700 mL of water at 30°C. Add 4.4 g of copper sulfate and dilute to 1 liter after the CO_2 is evolved. This solution should be shaken only in such a manner as not to allow the entry of air. After 24 hours 300 mL are diluted to a liter with saturated KCl solution, shaken gently and used after 24 hours; 50 mL $\approx$ 10 mg glucose.

Barfoed's reagent (test for glucose): dissolve 66 g of cupric acetate and 10 mL of glacial acetic acid in water and dilute to one liter.

Barium chloride, $BaCl_2 \cdot 2H_2O$—$0.5N$: 61 g per liter.

Barium hydroxide, $Ba(OH)_2 \cdot 8H_2O$—$0.2N$: 32 g per liter for saturation.

Barium nitrate, $Ba(NO_3)_2$—$0.5N$: 65 g per liter.

Baudisch's reagent: see *cupferron.*

Benedict's qualitative reagent (for glucose): dissolve 173 g of sodium citrate and 100 g of anhydrous sodium carbonate in about 600 mL of water, and dilute to 850 mL; dissolve 17.3 g of $CuSO_4 \cdot 5H_2O$ in 100 mL of water and dilute to 150 mL; this solution is added to the citrate-carbonate solution with constant stirring. See also the *quantitative reagent* below.

Benedict's quantitative reagent (sugar in urine): This solution contains 18 g copper sulfate, 100 g of anhydrous sodium carbonate, 200 g of potassium citrate, 125 g of potassium thiocyanate, and 0.25 g of potassium ferrocyanide per liter; 1 mL of this solution $\approx$ 0.002 sugar.

Benzidine hydrochloride solution (for sulfate determination): mix 6.7 g of benzidine $[C_{12}H_8(NH_2)_2]$ or 8.0 g of the hydrochloride $[C_{12}H_8(NH_2)_2 \cdot 2HCl]$ into a paste with 20 mL of water; add 20 mL of HCl (sp. gr. 1.12) and dilute the mixture to one liter with water; each mL of this solution is equivalent to 0.00357 g H_2SO_4.

Benzopurpurine 4B (indicator): dissolve 0.1 g in 100 mL water; pH range blueviolet 1.3—4.0 red.

Benzoyl auramine (indicator): dissolve 0.25 g in 100 mL methyl alcohol; pH range violet 5.0—5.6 pale yellow. Since this compound is not stable in aqueous solution, hydrolyzing slowly in neutral medium, more rapidly in alkaline and still more rapidly in acid solution, the indicator should not be added until one is ready to titrate. The acid quinoid form of the compound is dichroic, showing a red-violet in thick layers and blue in thin. At a pH of 5.4 the indicator appears a neutral gray color by daylight or a pale red under tungsten light. The change to yellow is easily recognized in either case. Cf. Scanlan and Reid, *Ind. Eng. Chem., Anal. Ed.* 7, 125 (1935).

Bertrand's reagents (glucose estimation): (*a*) 40 g of copper sulfate diluted to one liter; (*b*) rochelle salt 200 g, NaOH 150 g, and sufficient water to make one liter; (*c*) ferric sulfate 50 g, H_2SO_4 200 g, and sufficient water to make one liter; (*d*) $KMnO_4$ 5 g and sufficient water to make one liter.

TABLE 4.5 List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

Bial's reagent (for pentoses): dissolve 1 g of orcinol in 500 mL of 30% HCl to which 30 drops of a 10% ferric chloride solution have been added.

Bismuth chloride, $BiCl_3$—0.5N: 52 g per liter, using 1:5 HCl in place of water.

Bismuth nitrate, $Bi(NO_3)_3 \cdot 5H_2O$—0.25N: 40 g per liter, using 1:5 HNO_3 in place of water.

Bismuth standard solution (quantitative color test for Bi): dissolve 1 g of bismuth in a mixture of 3 mL of concentrated HNO_3 and 2.8 mL of H_2O and make up to 100 mL with glycerol. Also dissolve 5 g of KI in 5 mL of water and make up to 100 mL with glycerol. The two solutions are used together in the colorimetric estimation of Bi.

Boutron-Boudet solution: see *soap solution.*

Bromchlorophenol blue, dibromodichlorophenol-sulfonphthalein (indicator): dissolve 0.1 g in 8.6 mL 0.02 N NaOH and dilute with water to 250 mL; pH range yellow 3.2—4.8 blue.

Bromcresol green, tetrabromo-*m*-cresol-sulfonphthalein (indicator): dissolve 0.1 g in 7.15 mL 0.02 N NaOH and dilute with water to 250 mL; or, 0.1 g in 100 mL 20% alcohol; pH range yellow 4.0—5.6 blue.

Bromcresol purple, dibromo-*o*-cresol-sulfonphthalein (indicator): dissolve 0.1 g in 9.5 mL 0.02 N NaOH and dilute with water to 250 mL; or, 0.1 g in 100 mL 20% alcohol; pH range yellow 5.2—6.8 purple.

Bromine water, saturated solution: to 400 mL water add 20 mL of bromine; use a glass stopper coated with petrolatum.

Bromphenol blue, tetrabromophenol-sulfonphthalein (indicator): dissolve 0.1 g in 7.45 mL 0.02 N NaOH and dilute with water to 250 mL; or, 0.1 g in 100 mL 20% alcohol; pH range yellow 3.6—4.6 violet-blue.

Bromphenol red, dibromophenol-sulfonphthalein (indicator): dissolve 0.1 g in 9.75 mL 0.02 N NaOH and dilute with water to 250 mL; pH range yellow 5.2—7.0 red.

Bromthymol blue, dibromothymol-sulfonphthalein (indicator): dissolve 0.1 g in 8.0 mL 0.02 N NaOH and dilute with water to 250 mL; or, 0.1 g in 100 mL of 20% alcohol; pH range yellow 6.0—7.6 blue.

Brucke's reagent (protein precipitant): dissolve 50 g of KI in 500 mL of water, saturate with HgI_2 (about 120 g), and dilute to one liter.

Cadmium chloride, $CdCl_2$—0.5N: 46 g per liter.

Cadmium nitrate, $Cd(NO_3)_2 \cdot 4H_2O$—0.5N: 77 g per liter.

Cadmium sulfate, $CdSO_4 \cdot 4H_2O$—0.5N: 70 g per liter.

Calcium chloride, $CaCl_2 \cdot 6H_2O$—0.5N: 55 g per liter.

Calcium hydroxide, $Ca(OH)_2$—0.04N: 10 g per liter for saturation.

Calcium nitrate, $Ca(NO_3)_2 \cdot 4H_2O$—0.5N: 59 g per liter.

Calcium sulfate, $CaSO_4 \cdot 2H_2O$—0.03N: mechanically stir 10 g in a liter of water for 3 hours; decant and use the clear liquid.

Carbon disulfide, CS_2: commercial grade which is colorless.

Chloride reagent: dissolve 1.7 g of $AgNO_3$ and 25 g KNO_3 in water, add 17 mL of concentrated NH_4OH and make up to one liter with water.

Chlorine water, saturated solution: pass chlorine gas into small amounts of water as needed; solutions deteriorate on standing.

Chloroform, $CHCl_3$: commercial grade.

TABLE 4.5 List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

Chloroplatinic acid, $H_2PtCl_6 \cdot 6H_2O$—10% solution: dissolve 1 g in 9 mL of water; keep in a dropping bottle.

Chlorphenol red, dichlorophenol-sulfonphthalein (indicator): dissolve 0.1 g in 11.8 mL 0.02 N NaOH and dilute with water to 250 mL; or, 0.1 g in 100 mL 20% alcohol; pH range yellow 5.2—6.6 red.

Chromic chloride, $CrCl_3$—0.5N: 26 g per liter.

Chromic nitrate, $Cr(NO_3)_3$—0.5N: 40 g per liter.

Chromic sulfate, $Cr_2(SO_4)_3 \cdot 18H_2O$—0.5$N$: 60 g per liter.

Cobaltous nitrate, $Co(NO_3)_2 \cdot 6H_2O$—0.5N: 73 g per liter.

Cobaltous sulfate, $CoSO_4 \cdot 7H_2O$—0.5N: 70 g per liter.

Cochineal (indicator): triturate 1 g with 75 mL alcohol and 75 mL water, let stand for two days and filter; pH range red 4.8—6.2 violet.

Congo red, sodium tetrazodiphenyl-naphthionate (indicator): dissolve 0.1 g in 100 mL water; pH range blue 3.0—5.2 red.

Corallin (indicator): see *rosolic acid.*

Cresol red, o-cresol-sulfonphthalein (indicator): dissolve 0.1 g in 13.1 mL 0.02 N NaOH and dilute with water to 250 mL; or, 0.1 g in 100 mL 20% alcohol; pH range yellow 7.2—8.8 red.

o-Cresolphthalein (indicator): dissolve 0.1 g in 250 mL alcohol; pH range colorless 8.2—10.4 red.

Cupferron (iron analysis): dissolve 6 g of ammonium nitrosophenyl-hydroxylamine (cupferron) in water and dilute to 100 mL. This solution is stable for about one week if protected from light.

Cupric chloride, $CuCl_2 \cdot 2H_2O$—0.5N: 43 g per liter.

Cupric nitrate, $Cu(NO_3)_2 \cdot 6H_2O$—0.5N: 74 g per liter.

Cupric sulfate, $CuSO_4 \cdot 5H_2O$—0.5N: 62 g per liter.

Cuprous chloride, $CuCl$—0.5N: 50 g per liter, using 1:5 HCl in place of water.

Cuprous chloride, acid (for gas analysis, absorption of CO): cover the bottom of a two-liter bottle with a layer of copper oxide ⅜ inch deep, and place a bundle of copper wire an inch thick in the bottle so that it extends from the top to the bottom. Fill the bottle with HCl (sp. gr. 1.10). The bottle is shaken occasionally, and when the solution is colorless or nearly so, it is poured into half-liter bottles containing copper wire. The large bottle may be filled with hydrochloric acid, and by adding the oxide or wire when either is exhausted, a constant supply of the reagent is available.

Cuprous chloride, ammoniacal: this solution is used for the same purpose and is made in the same manner as the acid cuprous chloride above, except that the acid solution is treated with ammonia until a faint odor of ammonia is perceptible. Copper wire should be kept with the solution as in the acid reagent.

Curcumin (indicator): prepare a saturated aqueous solution; pH range yellow 6.0—8.0 brownish red.

Dibromophenol-tetrabromophenol-sulfonphthalein (indicator): dissolve 0.1 g in 1.21 mL 0.1N NaOH and dilute with water to 250 mL; pH range yellow 5.6—7.2 purple.

Dimethyl glyoxime, $(CH_3CNOH)_2$—0.01N: 6 g in 500 mL of 95% alcohol.

2,4-Dinitrophenol (indicator): dissolve 0.1 g in a few mL alcohol, then dilute with water to 100 mL; pH range colorless 2.6—4.0 yellow.

TABLE 4.5   List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

2,5-Dinitrophenol (indicator): dissolve 0.1 g in 20 mL alcohol, then dilute with water to 100 mL; pH range colorless 4—5.8 yellow.

2,6-Dinitrophenol (indicator): dissolve 0.1 g in a few mL alcohol, then dilute with water to 100 mL; pH range colorless 2.4—4.0 yellow.

Esbach's reagent (estimation of proteins): dissolve 10 g of picric acid and 20 g of citric acid in water and dilute to one liter.

Eschka's mixture (sulfur in coal): mix 2 parts of porous calcined MgO with 1 part of anhydrous Na_2CO_3; not a solution but a dry mixture.

Ether, $(C_2H_5)_2O$—use commercial grade.

***p*-Ethoxychrysoidine,** *p*-ethoxybenzeneazo-*m*-phenylenediamine (indicator): dissolve 0.1 g of the base in 100 mL 90% alcohol; or, 0.1 g of the hydrochloride salt in 100 mL water; pH range red 3.5—5.5 yellow.

Ethyl bis-(2,4-dinitrophenyl) acetate (indicator): the stock solution is prepared by saturating a solution containing equal volumes of alcohol and acetone with the indicator; pH range colorless 7.4.—9.1 deep blue. This compound is available commercially. The preparation of this compound is described by Fehnel and Amstutz, *Ind. Eng. Chem., Anal. Ed.* 16, 53 (1944), and by von Richter, *Ber.* 21, 2470 (1888), who recommended it for the titration of orange- and red-colored solutions or dark oils in which the end-point of phenolphthalein is not easily visible. The indicator is an orange solid which after crystallization from benzene gives pale yellow crystals melting at $150-153.5°C$, uncorrected.

Fehling's solution (sugar detection and estimation): (*a*) Copper sulfate solution: dissolve 34.639 g of $CuSO_4·5H_2O$ in water and dilute to 500 mL. (*b*) Alkaline tartrate solution: dissolve 173 g of rochelle salts ($KNaC_4H_4O_6 · 4H_2O$) and 125 g of KOH in water and dilute to 500 mL. Equal volumes of the two solutions are mixed just prior to use. The Methods of the Assoc. of Official Agricultural Chemists give 50 g of NaOH in place of the 125 g KOH.

Ferric chloride, $FeCl_3$—0.5*N*: 27 g per liter.

Ferric nitrate, $Fe(NO_3)_3 · 9H_2O$—0.5*N*: 67 g per liter.

Ferrous ammonium sulfate, Mohr's salt, $FeSO_4 · (NH_4)_2SO_4 · 6H_2O$—0.5*N*: 196 g per liter.

Ferrous sulfate, $FeSO_4 ·7 H_2O$—0.5*N*: 80 g per liter; add a few drops of H_2SO_4.

Folin's mixture (for uric acid): dissolve 500 g of ammonium sulfate, 5 g of uranium acetate, and 6 mL of glacial acetic acid, in 650 mL of water. The volume is about a liter.

Formal or Formalin: use the commercial 40% solution of formaldehyde.

Froehde's reagent (gives characteristic colorations with certain alkaloids and glycosides): dissolve 0.01 g of sodium molybdate in 1 mL of concentrated H_2SO_4; use only a freshly prepared solution.

Gallein (indicator): dissolve 0.1 g in 100 mL alcohol; pH range light brown-yellow 3.8—6.6 rose.

Glyoxylic acid solution (protein detection): cover 10 g of magnesium powder with water and add slowly 250 mL of a saturated oxalic acid solution, keeping the mixture cool; filter off the magnesium oxalate, acidify the filtrate with acetic acid and make up to a liter with water.

Guaiacum tincture: dissolve 1 g of guaiacum in 100 mL of alcohol.

Gunzberg's reagent (detection of HCl in gastric juice): dissolve 4 g of phloroglucinol and 2 g of vanillin in 100 mL of absolute alcohol; use only a freshly prepared solution.

TABLE 4.5 List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

Hager's reagent (for alkaloids): this reagent is a saturated solution of picric acid in water.

Hanus solution (for determination of iodine number): dissolve 13.2 g of iodine in a liter of glacial acetic acid that will not reduce chromic acid; add sufficient bromine to double the halogen content determined by titration (3 mL is about the right amount). The iodine may be dissolved with the aid of heat, but the solution must be cold when the bromine is added.

Hematoxylin (indicator): dissolve 0.5 g in 100 mL alcohol; pH range yellow 5.0—6.0.

Heptamethoxy red, 2,4,6,2′,4′,2″,4″-heptamethoxytriphenyl carbinol (indicator): dissolve 0.1 g in 100 mL alcohol; pH range red 5.0—7.0 colorless.

Hydriodic acid, HI—0.5N: 64 g per liter.

Hydrobromic acid, HBr—0.5N: 40 g per liter.

Hydrochloric acid, HCl—5N: 182 g per liter; sp. gr. 1.084.

Hydrofluoric acid, H_2F_2—48% solution: use as purchased, and keep in the special container.

Hydrogen peroxide, H_2O_2—3% solution: use as purchased.

Hydrogen sulfide, H_2S: prepare a saturated aqueous solution.

Indicator solutions: a number of indicator solutions are listed in this section under the names of the indicators; e.g., alizarin, aurin, azolitmin, et al., which follow alphabetically. See also various index entries.

Indigo carmine, sodium indigodisulfonate (indicator): dissolve 0.25 g in 100 mL 50% alcohol; pH range blue 11.6—14.0 yellow.

Indo-oxine, 5,8-quinolinequinone-8-hydroxy-5-quinoyl-5-imide (indicator): dissolve 0.05 g in 100 mL alcohol; pH range red 6.0—8.0 blue. Cf. Berg and Becker, *Z Anal. Chem.* 119, 81(1940).

Iodeosin, tetraiodofluorescein (indicator): dissolve 0.1 g in 100 mL ether saturated with water; pH range yellow 0—about 4 rose-red; see also under *methyl orange.*

Iodic acid, HIO_3—0.5N ($HIO_3/12$): 15 g per liter.

Iodine: see *tincture of iodine.*

Lacmoid (indicator): dissolve 0.5 g in 100 mL alcohol; pH range red 4.4—6.2 blue.

Litmus (indicator): powder the litmus and make up a 2% solution in water by boiling for 5 minutes; pH range red 4.5—8.3 blue.

Lead acetate, $Pb(C_2H_3O_2)_2 \cdot 3H_2O$—0.5$N$: 95 g per liter.

Lead chloride, $PbCl_2$—saturated solution is 1/7N.

Lead nitrate, $Pb(NO_3)_2$—0.5N: 83 g per liter.

Lime water: see *calcium hydroxide.*

Magnesia mixture: 100 g of $MgSO_4$, 200 g of NH_4Cl, 400 mL of NH_4OH, 800 mL of water; each cc $\approx$ 0.01 g phosphorus P).

Magnesium chloride, $MgCl_2 \cdot 6H_2O$—0.5N: 50 g per liter.

Magnesium nitrate, $Mg(NO_3)_2 \cdot 6H_2O$—0.5N: 64 g per liter.

Magnesium sulfate, epsom salts, $MgSO_4 \cdot 7H_2O$—0.5N: 62 g per liter; saturated solution dissolve 600 g of the salt in water and dilute to one liter.

Manganous chloride, $MnCl_2 \cdot 4H_2O$—0.5N: 50 g per liter.

Manganous nitrate, $Mn(NO_3)_2 \cdot 6H_2O$—0.5N: 72 g per liter.

Manganous sulfate, $MnSO_4 \cdot 7H_2O$—0.5N: 69 g per liter.

TABLE 4.5   List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

Marme's reagent (gives yellowish-white precipitate with salts of alkaloids): saturate a boiling solution of 4 parts of KI in 12 parts of water with CdI_2; then add an equal volume of cold saturated KI solution.

Marquis reagent (gives a purple-red coloration, then violet, then blue with morphine, codeine, dionine, and heroine): mix 3 mL of concentrated H_2SO_4 with 3 drops of a 35% formaldehyde solution.

Mayer's reagent (gives white precipitate with most alkaloids in a slightly acid solution): dissolve 13.55 g of $HgCl_2$ and 50 g of KI in a liter of water.

Mercuric chloride, $HgCl_2$—0.5*N*: 68 g per liter.

Mercuric nitrate, $Hg(NO_3)_2$—0.5*N*: 81 g per liter.

Mercuric sulfate, $HgSO_4$—0.5*N*: 74 g per liter.

Mercurous nitrate, $HgNO_3$: mix 1 part of $HgNO_3$, 20 parts of H_2O, and 1 part of HNO_3.

Metacresol purple, *m*-cresol-sulfonphthalein (indicator): dissolve 0.1 g in 13.6 mL 0.02*N* NaOH and dilute with water to 250 mL; acid pH range red 0.5—2.5 yellow, alkaline pH range yellow 7.4—9.0 purple.

Metanil yellow, diphenylaminoazo-*m*-benzene sulfonic acid (indicator): dissolve 0.25 g in 100 mL alcohol; pH range red 1.2—2.3 yellow.

Methyl green, hexamethylpararosaniline hydroxymethylate (component of mixed indicator): dissolve 0.1 g in 100 mL alcohol; when used with equal part of hexamethoxytriphenyl carbinol gives color change from violet to green at a titration exponent (pI) of 4.0.

Methyl orange, orange III, tropeolin D, sodium *p*-dimethylaminoazobenzenesulfonate (indicator): dissolve 0.1 g in 100 mL water; pH range red 3.0—4.4 orange-yellow. If during a titration where methyl yellow is being used a precipitate forms which tends to remove the indicator from the aqueous phase, methyl orange will be found to be a more suitable indicator. This occurs, for example, in titrations of soaps with acids. The fatty acids, liberated by the titration, extract the methyl yellow so that the end-point cannot be perceived. Likewise methyl orange is more suitable for titrations in the presence of immiscible organic solvents such as carbon tetrachloride or ether used in the extraction of alkaloids for analysis. Iodeosin (*q. v.*) has also been proposed as an indicator for such cases. Cf. Mylius and Foerster, *Ber.* 24, 1482 (1891); *Z Anal. Chem.* 31, 240 (1892).

Methyl red, *p*-dimethylaminoazobenzene-*o'*-carboxylic acid (indicator): dissolve 0.1 g in 18.6 mL 0.02 *N* NaOH and dilute with water to 250 mL; or, 0.1 g in 60% alcohol; pH range red 4.4—6.2 yellow.

Methyl violet (indicator): dissolve 0.25 g in 100 mL water; pH range blue 1.5—3.2 violet.

Methyl yellow, *p*-dimethylaminoazobenzene, benzeneazodimethylaniline (indicator): dissolve 0.1 g in 200 mL alcohol; pH range red 2.9—4.0 yellow. The color change from yellow to orange can be perceived somewhat more sharply than the change of methyl orange from orange to rose, so that methyl yellow seems to deserve preference in many cases. See also under *methyl orange.*

Methylene blue, *N,N,N',N'*-tetramethylthionine (component of mixed indicator): dissolve 0.1 g in 100 mL alcohol; when used with equal part of methyl yellow gives color change from blue-violet to green at a titration exponent (pI) of 3.25; when used with equal part of 0.2% methyl red in alcohol gives color change from red-violet to green at a titration exponent (pI) of 5.4; when used with an equal part of neutral red gives color change from violet-blue to green at a titration exponent (pI) of 7.0.

Millon's reagent (gives a red precipitate with certain proteins and with various phenols): dissolve 1 part of mercury in 1 part of HNO_3 (sp. gr. 1.40) with gentle heating, then add 2 parts of water; a few crystals of KNO_3 help to maintain the strength of the reagent.

TABLE 4.5 List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

Mohr's salt: see *ferrous ammonium sulfate.*

α-Naphthol solution: dissolve 144 g of α-naphthol in enough alcohol to make a liter of solution.

α-Naphtholbenzein (indicator): dissolve 0.1 g in 100 mL 70% alcohol; pH range colorless 9.0—11.0 blue.

α-Naphtholphthalein (indicator): dissolve 0.1 g in 50 mL alcohol and dilute with water to 100 mL; pH range pale yellow-red 7.3—8.7 green.

Nessler's reagent (for free ammonia): dissolve 50 g of KI in the least possible amount of cold water; add a saturated solution of $HgCl_2$ until a very slight excess is indicated; add 400 mL of a 50% solution of KOH; allow to settle, make up to a liter with water, and decant.

Neutral red, toluylene red, dimethyldiaminophenazine chloride, aminodimethylaminotoluphenazine hydrochloride (indicator): dissolve 0.1 g in 60 mL alcohol and dilute with water to 100 mL; pH range red 6.8—8.0 yellow-orange.

Nickel chloride, $NiCl_2 \cdot 6H_2O$—0.5*N*: 59 g per liter.

Nickel nitrate, $Ni(NO_3)_2 \cdot 6H_2O$—0.5*N*: 73 g per iter.

Nickel sulfate, $NiSO_4 \cdot 6H_2O$−0.5*N*: 66 g per liter.

Nitramine, picrylmethylnitramine, 2,4,6-trinitrophenylmethyl nitramine (indicator): dissolve 0.1 g in 60 mL alcohol and dilute with water to 100 mL; pH range colorless 10.8—13.0 red-brown; the solution should be kept in the dark as nitramine is unstable; on boiling with alkali it decomposes quickly. Fresh solutions should be prepared every few months.

Nitric acid, HNO_3—5*N*: 315 g per liter; sp. gr. 1.165.

Nitrohydrochloric acid: see *aqua regia.*

p-Nitrophenol (indicator): dissolve 0.2 g in 100 mL water; pH range colorless at about 5—7 yellow.

Nitroso-β-naphthol, $HOC_{10}H_6NO$—saturated solution: saturate 100 mL of 50% acetic acid with the solid.

Nylander's solution (detection of glucose): dissolve 40 g of rochelle salt and 20 g of bismuth subnitrate in 1000 mL of an 8% NaOH solution.

Obermayer's reagent (detection of indoxyl in urine): dissolve 4 g of $FeCl_3$ in a liter of concentrated HCl.

Orange III (indicator): see under *methyl orange.*

Oxalic acid, $H_2C_2O_4 \cdot 2H_2O$: dissolve in ten parts of water.

Pavy's solution (estimation of glucose): mix 120 mL of Fehling's solution and 300 mL of ammonium hydroxide (sp. gr. 0.88), and dilute to a liter with water.

Perchloric acid, $HClO_4$—60%: use as purchased.

Phenol solution: dissolve 20 g of phenol (carbolic acid) in a liter of water.

Phenol red, phenol-sulfonphthalein (indicator): dissolve 0.1 g in 14.20 mL 0.02*N* NaOH and dilute with water to 250 mL; or, 0.1 g in 100 mL 20% alcohol; pH range yellow 6.8—8.0 red.

Phenolphthalein (indicator): dissolve 1 g in 60 mL of alcohol and dilute with water to 100 mL; pH range colorless 8.2—10.0 red.

Phenol sulfonic acid (determination of nitrogen as nitrate; water analysis for nitrate): dissolve 25 g pure, white phenol in 150 mL of pure concentrated H_2SO_4, add 75 mL of fuming H_2SO_4 (15% SO_3), stir well and heat for two hours at 100°C.

TABLE 4.5   List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

Phosphoric acid, *ortho,* H_3PO_4—0.5N: 16 g per liter.

Poirrer blue C4B (indicator): dissolve 0.2 g in 100 mL water; pH range blue 11.0—13.0 red.

Potassium acid antimonate, KH_2SbO_4—0.1N: boil 23 g of the silt with 950 mL of water for 5 minutes, cool rapidly and add 35 mL of 6N KOH; allow to stand for one day, filter dilute filtrate to a liter.

Potassium arsenate, K_3AsO_4—0.5N ($K_3AsO_4/10$): 26 g per liter.

Potassium arsenite, $KAsO_2$−0.5N ($KAsO_2/6$): 24 g per liter.

Potassium bromate, $KBrO_3$—0.5N ($KBrO_3/12$): 14 g per liter.

Potassium bromide, KBr—0.5N: 60 g per liter.

Potassium carbonate, K_2CO_3—3N: 207 g per liter.

Potassium chloride, KCl—0.5N: 37 g per liter.

Potassium chromate, K_2CrO_4—0.5N: 49 g per liter.

Potassium cyanide, KCN—0.5N: 33 g per liter.

Potassium dichromate, $K_2Cr_2O_7$—0.5N ($K_2Cr_2O_7/8$): 38 g per liter.

Potassium ferricyanide, $K_3Fe(CN)_6$—0.5N: 55 g per liter.

Potassium ferrocyanide, $K_4Fe(CN)_6 \cdot 3H_2O$—0.5N: 53 g per liter.

Potassium hydroxide, KOH—5N: 312 g per liter.

Potassium iodate, KIO_3—0.5N ($KIO_3/12$): 18 g per liter.

Potassium iodide, KI—0.5N: 83 g per liter.

Potassium nitrate, KNO_3—0.5N: 50 g per liter.

Potassium nitrite, KNO_2—6N: 510 g per liter.

Potassium permanganate, $KMnO_4$—0.5N ($KMnO_4/10$): 16 g per liter.

Potassium pyrogallate (oxygen in gas analysis): weigh out 5 g of pyrogallol (pyrogallic acid), and pour upon it 100 mL of a KOH solution. If the gas contains less than 28% of oxygen, the KOH solution should be 500 g KOH in a liter of water; if there is more than 28% of oxygen in the gas, the KOH solution should be 120 g of KOH in 100 mL of water.

Potassium sulfate, K_2SO_4—0.5N: 44 g per liter.

Potassium thiocyanate, KCNS—0.5N: 49 g per liter.

Precipitating reagent (for group II, anions): dissolve 61 g of $BaCl_2 \cdot 2H_2O$ and 52 g of $CaCl_2 \cdot 6H_2O$ in water and dilute to one liter. If the solution becomes turbid, filter and use filtrate.

Quinaldine red (indicator): dissolve 0.1 g in 100 mL alcohol; pH range colorless 1.4—3.2 red.

Quinoline blue, cyanin (indicator): dissolve 1 g in 100 mL alcohol; pH range colorless 6.6—8.6 blue.

Rosolic acid, aurin, corallin, corallinphthalein, 4,4′-dihydroxy-fuchsone, 4,4′-dihydroxy-3-methyl-fuchsone (indicator): dissolve 0.5 g in 50 mL alcohol and dilute with water to 100 mL.

Salicyl yellow (indicator): see *alizarin yellow GG.*

Scheibler's reagent (precipitates alkaloids, albumoses and peptones): dissolve sodium tungstate in boiling water containing half its weight of phosphoric acid (sp. gr. 1.13); on evaporation of this solution, crystals of phosphotungstic acid are obtained. A 10% solution of phosphotungstic acid in water constitutes the reagent.

TABLE 4.5 List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

Schweitzer's reagent (dissolves cotton, linen, and silk, but not wool); add NH_4Cl and NaOH to a solution of copper sulfate. The blue precipitate is filtered off, washed, pressed, and dissolved in ammonia (sp. gr. 0.92).

Silver nitrate, $AgNO_3$—0.25N: 43 g per liter.

Silver sulfate, Ag_2SO_4—$N/13$ (saturated solution): stir mechanically 10 g of the salt in a liter of water for 3 hours; decant and use the clear liquid.

Soap solution (for hardness in water): (*a*) *Clark's or A.P.H.A. Stand. Methods*—prepare stock solution of 100 g of pure powdered castile soap in a liter of 80% ethyl alcohol; allow to stand over night and decant. Titrate against $CaCl_2$ solution (0.5 g $CaCO_3$ dissolved in a concentrated HCl, neutralized with NH_4OH to slight alkalinity using litmus as the indicator, make up to 500 mL; 1 mL of this solution is equivalent to 1 mg $CaCO_3$) and dilute with 80% alcohol until 1 mL of the resulting solution is equivalent to 1 mL of the standard $CaCl_2$ making due allowance for the lather factor (the lather factor is that amount of standard soap solution required to produce a permanent lather in a 50-mL portion of distilled water). One mL of this solution after subtracting the lather factor is equivalent to 1 mg of $CaCO_3$. (*b*) Boutron-Boudet—dissolve 100 g of pure castile soap in about 2500 mL of 56% ethyl alcohol and adjust so that 2.4 mL will give a permanent lather with 40 mL of a solution containing 0.59 g $Ba(NO_3)_2$ per liter of water; 2.4 mL of this solution is equivalent to 22 French degrees or 220 parts per million of hardness (as $CaCO_3$) on a 40-mL sample of water.

Sodium acetate, $NaC_2H_3O_2 \cdot 3H_2O$: dissolve 1 part of the salt in 10 parts of water.

Sodium acetate, acid: dissolve 100 g of sodium acetate and 30 mL of glacial acetic acid in water and dilute to one liter.

Sodium bismuthate (oxidation of manganese): heat 20 parts of NaOH nearly to redness in an iron or nickel crucible, and add slowly 10 parts of basic bismuth nitrate which has been previously dried. Add 2 parts of sodium peroxide, and pour the brownish-yellow fused mass on an iron plate to cool. When cold break up in a mortar, extract with water, and collect on an asbestos filter.

Sodium carbonate, Na_2CO_3—3N: 159 g per liter; one part Na_2CO_3, or 2.7 parts of the crystalline $Na_2CO_3 \cdot 10H_2O$ in 5 parts of water.

Sodium chloride, NaCl—0.5N: 29 g per liter.

Sodium chloroplatinite, Na_2PtCl_4: dissolve 1 part of the salt in 12 parts of water.

Sodium cobaltinitrite, $Na_3Co(NO_2)_6$—0.3N: dissolve 230 g of $NaNO_2$ in 500 mL of water, add 160 mL of 6N acetic acid and 35 g of $Co(NO_3)_2 \cdot 6H_2O$. Allow to stand one day, filter, and dilute the filtrate to a liter.

Sodium hydrogen phosphate, $Na_2HPO_4 \cdot 12H_2O$—0.5N: 60 g per liter.

Sodium hydroxide, NaOH—5N: 220 g per liter.

Sodium hydroxide, alcoholic: dissolve 20 g of NaOH in alcohol and dilute to one liter with alcohol.

Sodium hypobromite: dissolve 100 g of NaOH in 250 mL of water and add 25 mL of bromine.

Sodium nitrate, $NaNO_3$—0.5N: 43 g per liter.

Sodium nitroprusside (for sulfur detection): dissolve about one gram of sodium nitroprusside in 10 mL of water; as the solution deteriorates on standing, only freshly prepared solutions should be used. This compound is also called sodium nitroferricyanide and has the formula $Na_2Fe(NO)(CN)_5 \cdot 2H_2O$.

TABLE 4.5 List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

Sodium polysulfide, Na_2S_x: dissolve 480 g of $Na_2S \cdot 9H_2O$ in 500 mL of water, add 40 g of NaOH and 18 g of sulfur, stir mechanically and dilute to one liter with water.

Sodium sulfate, Na_2SO_4—0.5N: 35 g per liter.

Sodium sulfide, Na_2S; saturate NaOH solution with H_2S, then add as much NaOH as was used in the original solution.

Sodium sulfite, $Na_2SO_3 \cdot 7H_2O$—0.5N: 63 g per liter.

Sodium sulfite, acid (saturated): dissolve 600 g of $NaHSO_3$ in water and dilute to one liter; for the preparation of addition compounds with aldehydes and ketones:
prepare a saturated solution of sodium carbonate in water and saturate with sulfur dioxide.

Sodium tartrate, acid, $NaHC_4H_4O_6$: dissolve 1 part of the salt in 10 parts of water.

Sodium thiosulfate, $Na_2S_2O_3 \cdot 5H_2O$: one part of the salt in 40 parts of water.

Sonnenschein's reagent (alkaloid detection): a nitric acid solution of ammonium molybdate is treated with phosphoric acid. The precipitate so produced is washed and boiled with aqua regia until the ammonium salt is decomposed. The solution is evaporated to dryness and the residue is dissolved in 10% HNO_3.

Stannic chloride, $SnCl_4$—0.5N: 33 g per liter.

Stannous chloride, $SnCl_2 \cdot 2H_2O$—0.5N: 56 g per liter. The water should be acid with HCl and some metallic tin should be kept in the bottle.

Starch solution (iodine indicator): dissolve 5 g of soluble starch in cold water, pour the solution into 2 liters of water and boil for a few minutes. Keep in a glass-stoppered bottle.

Starch solution (other than soluble): make a thin paste of the starch with cold water, then stir in 200 times its weight of boiling water and boil for a few minutes. A few drops of chloroform added to the solution acts as a preservative.

Stoke's reagent: dissolve 30 g of ferrous sulfate and 20 g of tartaric acid in water and dilute to one liter. When required for use, add strong ammonia until the precipitate first formed is dissolved.

Strontium chloride, $SrCl_2 \cdot 6H_2O$—0.5N: 67 g per liter.

Strontium nitrate, $Sr(NO_3)_2$—0.5N: 53 g per liter.

Strontium sulfate, $SrSO_4$: prepare a saturated solution.

Sulfanilic acid (for detection of nitrites): dissolve 8 g of sulfanilic acid in one liter of acetic acid (sp. gr. 1.04).

Sulfuric acid, H_2SO_4—5N: 245 g per liter, sp. gr. 1.153.

Sulfurous acid, H_2SO_3: saturate water with sulfur dioxide.

Tartaric acid, $H_2C_4H_4O_6$: dissolve one part of the acid in 3 parts of water; for a saturated solution dissolve 750 g of tartaric acid in water and dilute to one liter.

Tannic acid: dissolve 1 g tannic acid in 1 mL alcohol and make up to 10 mL with water.

Tetrabromophenol blue, tetrabromophenol-tetrabromosulfonphthalein (indicator): dissolve 0.1 g in 5 mL 0.02N NaOH and dilute with water to 250 mL; pH range yellow 3.0—4.6 blue.

Thymol blue, thymol-sulfonphthalein (indicator): dissolve 0.1 g in 10.75 mL 0.02N NaOH and dilute with water to 250 mL; or, dissolve 0.1 g in 20 mL warm alcohol and dilute with water to 100 mL; pH range (acid) red 1.2—2.8 yellow, and (alkaline) yellow 8.0—9.6 blue.

Thymolphthalein (indicator): dissolve 0.1 g in 100 mL alcohol; pH range colorless 9.3—10.5 blue.

TABLE 4.5   List of Laboratory Solutions, Reagents, and Indicators (*Continued*)

Tincture of iodine (antiseptic): add 70 g of iodine and 50 g of KI to 50 mL of water; make up to one liter with alcohol.

o-**Tolidine solution** (for residual chlorine in water analysis): dissolve 1 g of pulverized *o*-tolidine, m.p. 129°C., in one liter of dilute hydrochloric acid (100 mL conc. HCl diluted to one liter).

Toluylene red (indicator): see *neutral red.*

Trichloroacetic acid: dissolve 100 g of the acid in water and dilute to one liter.

Trinitrobenzene, 1,3,5-trinitrobenzene (indicator): dissolve 0.1 g in 100 mL alcohol; pH range colorless 11.5—14.0 orange.

Trinitrobenzoic acid, 2,4,6-trinitrobenzoic acid (indicator): dissolve 0.1 g in 100 mL water; pH range colorless 12.0—13.4 orange-red.

Tropeolin D (indicator): see *methyl orange.*

Tropeolin O, sodium 2,4-dihydroxyazobenzene-4-sulfonate (indicator): dissolve 0.1 g in 100 mL water; pH range yellow 11.0—13.0 orange-brown.

Tropeolin OO, orange IV, sodium *p*-diphenylamino-azobenzene sulfonate, sodium 4′-anilino-azobenzene-4-sulfonate (indicator): dissolve 0.1 g in 100 mL water; pH range red 1.3—3.2 yellow.

Tropeolin OOO, sodium α-naphtholazobenzene sulfonate (indicator): dissolve 0.1 g in 100 mL water; pH range yellow 7.6—8.9 red.

Turmeric paper (gives a rose-brown coloration with boric acid); wash the ground root of turmeric with water and discard the washings. Digest with alcohol and filter, using the clear filtrate to impregnate white, unsized paper, which is then dried.

Uffelmann's reagent (gives a yellow coloration in the presence of lactic acid): add a ferric chloride solution to a 2% phenol solution until the solution becomes violet in color.

Wagner's solution (phosphate rock analysis): dissolve 25 g citric acid and 1 g salicylic acid in water, and make up to one liter. Twenty-five to 50 mL of this reagent prevents precipitation of iron and aluminum.

Wijs solution (for iodine number): dissolve 13 g resublimed iodine in one liter of glacial acetic acid (99.5%), and pass in washed and dried (over or through H_2SO_4) chlorine gas until the original thio titration of the solution is not quite doubled. There should be only a slight excess of iodine and no excess of chlorine. Preserve the solution in amber colored bottles sealed with paraffin. Do not use the solution after it has been prepared for more than 30 days.

Xylene cyanole-methyl orange indicator, Schoepfle modification (for partially color blind operators): dissolve 0.75 g xylene cyanole FF (Eastman No. T 1579) and 1.50 g methyl orange in 1 liter of water.

p-**Xylenol blue,** 1,4-dimethyl-5-hydroxybenzene-sulfonphthalein (indicator): dissolve 0.1 g in 250 mL alcohol; pH range (acid) red 1.2—2.8 yellow, and (alkaline) yellow 8.0—9.6 blue.

Zinc chloride, $ZnCl_2$—0.5*N*: 34 g per liter.

Zinc nitrate, $Zn(NO_3)_2 \cdot 6H_2O$—0.5*N*: 74 g per liter.

Zinc sulfate, $ZnSO_4 \cdot 7H_2O$—0.5*N*: 72 g per liter.

Source: From Dean.[2]

ipate in the reaction. The ratios obtained from the numerical coefficients in the chemical equation are the stoichiometric ratios that permit one to calculate the moles of one substance as related to the moles of another substance in the chemical equation. For example, the stoichiometric equation for the combustion of heptane

$$C_7H_{16} + 11O_2 \rightarrow 7CO_2 + 8H_2O$$
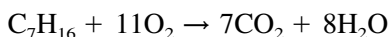
indicates that 1 mol (not kilogram) of heptane will react with 11 mol of oxygen to give 7 mol of carbon dioxide plus 8 mol of water. These may be pound-moles, kilogram-moles, gram-moles or any other type of mole, as shown in Fig. 4.2. One mol of CO_2 is formed from each $\frac{1}{7}$ mol of C_7H_{16}. Also, 1 mol of H_2O is formed with each $\frac{7}{8}$ mol of CO_2. Thus the equations tell us in terms of moles (not kilograms) the ratios among reactants and products.

Example 4.1. Determine the mass of oxygen that can be produced by the complete decomposition of 490 g of potassium chlorate.
Solution. The solution is based on the following information: There is 490 g $KClO_3$. Also,
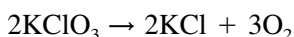
$$2KClO_3 \rightarrow 2KCl + 3O_2$$

$$\text{Mol wt of } KClO_3 = (39 + 35.5 + 48) = 122.5 \text{ g/g} \cdot \text{mol}$$

$$2KClO_3 = 2 \times 122.5 = 245$$

$$\text{Mol wt of } O_2 = (16)(2) = 32 \text{ g/g} \cdot \text{mol}$$

$$3O_2 = 3(32) = 96$$

Let X be the mass of O_2 liberated.
The proportion dictated by the law of combining weights is expressed by this scheme:

$$2KClO_3 \rightarrow 2KCl + 3O_2$$
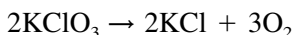
$$\frac{2(122.5) \text{ g}}{490 \text{ g}} = \frac{3(32) \text{ g}}{X \text{ g}}$$

Set a proportion

$$\frac{2(122.5)}{490} = \frac{3(32)}{X}$$

or

$$X = \frac{192 \text{ g } O_2}{490 \text{ g } KClO_3}$$

or expressed in another form: The chemical equation states that every 245 g of $KClO_3$ produces 96 g of O_2. How much O_2 can be produced with 490 g of $KClO_3$?

C_7H_{16}	+	$11O_2$	→	$7CO_2$	+	$8H_2O$

Qualitative description

Heptane	reacts with	oxygen	to give	carbon dioxide	and	water

Quantitative description

1 molecule of heptane	reacts with	11 molecules of oxygen	to give	7 molecules of carbon dioxide	and	8 molecules of water
6.023×10^{23} molecules of C_7H_{16}	+	$11(6.023 \times 10^{23})$ molecules of O_2	→	$7(6.023 \times 10^{23})$ molecules of CO_2	+	$8(6.023 \times 10^{23})$ molecules of H_2O
1 g · mol of C_7H_{16}	+	11 g · mol of O_2	→	7 g · mol of CO_2	+	8 g · mol of H_2O
1 lb · mol of C_7H_{16}	+	11 lb · mol of O_2	→	7 lb · mol of CO_2	+	8 lb · mol of H_2O
1 kg · mol of C_7H_{16}	+	11 kg · mol of O_2	→	7 kg · mol of CO_2	+	8 kg · mol of H_2O
1(100) g of C_7H_{16}	+	11(32) g of O_2	→	7(44) g of CO_2	+	8(18) g of H_2O
100 g		352 g		308 g		144 g

452 g	=	452 g
452 kg	=	452 kg
452 lb	=	452 lb

Note:

Component	Mol wt
C_7H_{16}	100
O_2	32
CO_2	44
H_2O	18

FIGURE 4.2 The chemical equation.

$$\frac{245 \text{ g KClO}_3}{490 \text{ g KClO}_3} = \frac{96 \text{ g O}_2}{X}$$

and

$$X = \frac{96 \text{ g O}_2 \times 490}{245}$$

$$= 192 \text{ g O}_2$$

CHEMICAL THERMODYNAMIC RELATIONS

Chemical Potential, $\hat{\mu}$. Chemical potential at constant temperature and pressure is Gibbs free energy per mole, expressed in J/mol. For a multicomponent system, $\hat{\mu}$ is the summation of potentials (partial molal free energies) of all the components present:

$$\hat{\mu} = \Sigma \frac{G_i}{n_i} = \Sigma \, \hat{g}_i \tag{4.1}$$

where G_i is Gibbs free energy of component i, and n_i is moles of i present.

Chemical Energy. Like other forms of energy, all products of intensive and extensive properties, chemical energy added to a system is $\hat{\mu}_i \cdot dn_i$ for one component, and $\Sigma\hat{\mu}_i \cdot dn_i$ for all components of the system.

Chemical Thermodynamic Equations. Basic differential equations of engineering thermodynamics are expanded by adding the chemical energy and other appropriate energy terms.

Relation of chemical potential to internal energy:

$$dU = T \, dS - P \, dV + \Sigma \, \hat{\mu}_i \cdot dn_i \tag{4.2}$$

Relation to enthalpy:

$$dH = T \, dS + V \, dP + \Sigma \, \hat{\mu}_i \cdot dn_i \tag{4.3}$$

Relation to Gibbs free energy:

$$dG = -S \, dT + V \, dP + \Sigma \, \hat{\mu}_i \cdot dn_i \tag{4.4}$$

Relation to Helmholtz function:

$$dF = -S \, dT - P \, dV + \Sigma \, \hat{\mu}_i \cdot dn_i \tag{4.5}$$

Activity, a. Activity is a thermodynamically effective concentration used in lieu of actual concentrations to compensate for deviations of gases from ideality, incomplete ionization of strong electrolytes in solution, and other discrepancies between calculated and experimental behavior.

Fugacity, f. Fugacity (escaping tendency) of a gas equals its activity. Fugacity is an effective partial pressure, expressed in atmospheres. For ideal gases fugacity exactly equals partial pressure. Since real gases in the standard state (in their usual phase at 1 atm pressure and 25°C) deviate slightly from ideality, activity of real gases is defined as the ratio of fugacity f to fugacity f^0, in a standard state where $f^0 = 1$ atm. For all practical purposes, f^0 is 1 atmosphere pressure. Nonidealities of real gases at other temperatures and pressures are handled by an activity coefficient, γ.

Activity Coefficient, γ. The activity coefficient is the ratio of fugacity to partial pressure, or the ratio of activity to molal concentration. For ideal gases: $\gamma = 1$.

Calculation of ΔG. At constant temperature and constant number of moles $dT = 0$ and $dn_i = 0$, Eq. (4.4) reduces to $dG = +V \, dP$. The calculation of ΔG for ideal gases, real gases, solids, and liquids at any pressure is dependent only on an expression for V as $f(P)$.

For ideal gases, integrating between limits gives:

$$\Delta G = RT \ln \frac{P_2}{P_1} \tag{4.6}$$

$$\Delta G = RT \ln \frac{f_2}{f_1} \tag{4.7}$$

For incompressible liquids and solids, V is constant, so

$$\Delta G = V(P_2 - P_1) \tag{4.8}$$

using appropriate units. The magnitude of this change for all condensed phases is very small.

THERMOCHEMISTRY*

1. Heat of Chemical Reaction. The heat of reaction or enthalpy of reaction $\Delta\hat{H}$, (T,P) is the difference $H_{\text{products}} - H_{\text{reactants}}$ for a reaction that takes place under the following circumstances:

1. Stoichiometric quantities of the reactants are fed, and the reaction proceeds to completion.
2. The reactants are fed at temperature T and pressure P, and the products emerge at the same temperature and pressure.

For example, for the reaction

$$CaC_2(s) + 2H_2O(l) \rightarrow Ca(OH)_2(s) + C_2H_2(g)$$
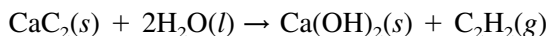
the heat of reaction at 25°C and a total pressure of 1 atm is

$$\Delta\hat{H}(25°C, 1 \text{ atm}) = -125.4 \text{ kJ/mol}$$

If the reaction is run under conditions such that the energy balance reduces to $Q = \Delta H$, then 125.4 kJ is emitted by the reaction system in the course of the reaction. (A negative Q implies flow of heat out of the system.)

If v_A is the stoichiometric coefficient of a reactant or reaction product A, and n_A moles of A are consumed or produced at $T = T_0$, the total enthalpy change is

$$\Delta H = \frac{\Delta\hat{H}_r(T_0)}{v_A} n_A \tag{4.9}$$

Several definitions and properties of heats of reaction are summarized below:

1. The standard heat of reaction $\Delta\hat{H}_r^0$ is the heat of reaction when both the reactions and products are at a specified temperature and pressure, usually 25°C and 1 atm.

---

2. If $\Delta\hat{H}_r(T)$ is negative, the reaction is said to be *exothermic* at temperature T; if $\Delta\hat{H}_r(T)$ is positive, the reaction is *endothermic* at T.

3. At low and moderate pressure, $\Delta\hat{H}_r$ is nearly independent of pressure.

4. The value of a heat of reaction depends on how the stoichiometric equation is written: for example, $\Delta\hat{H}_r^0$ for $A \rightarrow B$ is half of $\Delta\hat{H}_r^0$ for $2A \rightarrow 2B$. (This should seem reasonable taking into account the definition of $\Delta\hat{H}_r$.)

5. The value of a heat of reaction depends on the states of aggregation (gas, liquid, or solid) of the reactants and products.

Various types of heats of reaction are found in thermochemical systems. They include heats of formation, combustion, neutralization, solution, dilution, dissociation, and polymerization, among others. Since all are particular types of heats of reaction, they are all subject to similar treatments and thermodynamic requirements.

Hess's Law of Heat Summation. Hess showed in 1840 that the overall heat evolved or absorbed in a chemical reaction proceeding in several steps is equal to the algebraic sum of the enthalpies of the various stages. These chemical equations may be treated like ordinary algebraic equations.

2. Formation Reactions and Heat of Formation. A formation reaction of a compound is the reaction in which the compound is formed from its atomic constituents as they normally occur in nature (e.g., O_2 from O). The standard heat of such a reaction is the standard heat of formation, $\Delta\hat{H}_f^0$, of the compound.

Standard heats of formation for many compounds are listed in Tables 9.1 and 9.2 of Ref. 1. For example, $\Delta\hat{H}_f^0$ for crystalline ammonium nitrate is given in Table 9.1 of Ref. 2 as -365.14 kJ/mol, signifying

$$N_2(g) + 2H_2(g) + \tfrac{3}{2}O_2(g) \rightarrow NH_4NO_3(c): \Delta\hat{H}_r^0 = -365.14 \text{ kJ/mol}$$

Similarly, for liquid benzene $\Delta\hat{H}_f^0 = 48.66$ kJ/mol, or

$$6C(s) + 3H_2(g) \rightarrow C_6H_6(l): \Delta\hat{H}_r^0 = 48.66 \text{ kJ/mol}$$
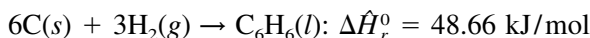
The standard heat of formation (pressure 1 atm and a temperature of 25°C) of an element is conveniently chosen to be zero.

It may be shown using Hess's law that if v_1 is the stoichiometric coefficient of the ith species participating in a reaction and $(\Delta\hat{H}_f^0)$ is the standard heat of formation of this species, the standard heat of the reaction is

$$\Delta\hat{H}_r^0 = \sum_{\text{products}} v_i(\Delta\hat{H}_f^0)_i - \sum_{\text{reactants}} v_i(\Delta\hat{H}_f^0)_i \tag{4.10}$$

3. Energy Balance on Reactive Processes. Several approaches to setting up energy balances on reactive systems are often adopted, which differ in the reference conditions used for enthalpy calculations. Suppose $\Delta\hat{H}_r^0$ for each reaction occurring in a system is known or calculable at a temperature T_0 (which is usually but not always 25°C). The following are two common choices of reference conditions, along with the procedures for calculating ΔH that correspond to each choice.

Choice 1. Reference conditions: the elements that constitute the reactants and products at 25°C and the nonreactive molecular species at any convenient temperature. A table of n_i's (n = number of moles) and $\hat{H}_i$'s for all stream components is

constructed, only now $\hat{H}_i$ for a reactant or product is the sum of the heat of formation of the species at 25°C and any sensible and latent heats required to bring the species from 25°C to its inlet or outlet state. The overall enthalpy change for the process is then

$$\Delta H = \sum_{\substack{\text{outlet} \\ \text{(products)}}} n_i \hat{H}_i - \sum_{\substack{\text{inlet} \\ \text{(reactants)}}} n_i \hat{H}_i \tag{4.11}$$

Example 4.2 *Energy balance on a methane oxidation reactor.* Methane is oxidized with air to produce formaldehyde in a continuous reactor. A competing reaction is the combustion of methane to form CO_2.

1. $CH_4(g) + O_2 \rightarrow HCHO(g) + H_2O(g)$
2. $CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O(g)$

A flow chart of the process for an assumed basis of 100 mol of methane fed to the reactor is shown in Fig. 4.3.

The pressure is low enough for ideal gas behavior to be assumed. If the methane enters the reactor at 25°C and the air enters at 100°C, how much heat must be withdrawn for the product stream to emerge at 150°C?

Solution. The problem is solved as follows:

$CH_4(150°C)$:

$$(\hat{H}_{CH_4})_{out} = (\Delta \hat{H}_f^0)_{CH_4} + \int_{25}^{150} (C_p)_{CH_4}\, dT$$

$$= -74.85 + 4.90 = -69.95 \text{ kJ/mol}$$

$HCHO(g, 150°C)$:

$$\hat{H} = (\Delta \hat{H}_f^0)_{HCHO(g)} + \int_{25}^{150} (C_p)_{HCHO}\, dT$$

$$= -115.90 + 4.75 = -111.15 \text{ kJ/mol}$$

In the last two integrals C_p is substituted from Eq. (4.17) (see below) (see Table 4.6 for values of coefficients of a, b, and c).
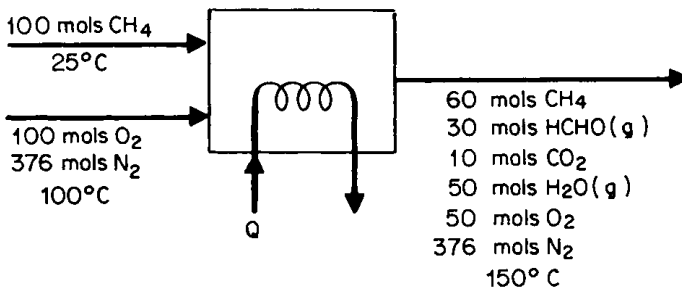

FIGURE 4.3 Flow chart of methane oxidation process.

TABLE 4.6   Molar Heat Capacities of Some Gases

$C_p(J/mol \cdot °C) = a + bT + cT^2 + dT^3$; T in °C

Compound	Formula	a	$b \cdot 10^2$	$c \cdot 10^5$	$d \cdot 10^9$
Oxygen	O_2	29.10	1.158	−0.6076	1.311
Hydrogen	H_2	28.84	0.00765	0.3288	−0.8698
Nitrogen	N_2	29.00	0.2199	0.5723	−2.871
Carbon monoxide	CO	28.95	0.4110	0.3548	−2.220
Carbon dioxide	CO_2	36.11	4.233	−2.887	7.464
Water	$H_2O(g)$	33.46	0.6880	0.7604	−3.593
Methane	CH_4	34.28	4.268	0.0	−8.864
Formaldehyde	CH_2O	34.31	5.469	0.366	−11.0

Note: Data shown in this table can also be found in Refs. 2 through 5.

$CO_2(150°C)$:

$$\hat{H} = (\Delta\hat{H}_f^0)_{CO_2(g)} + [125\overline{C}_p(150°C)]$$

Taking $\Delta\hat{H}_f^0$ from Table 9.2 of Ref. 1 and $\overline{C}_p$ from Table 4.7,

$$(\hat{H}_{CO_2})_{out} = -393.5 + 4.94 = -388.6 \text{ kJ/mol}$$

$H_2O(g, 150°C)$:

$$\hat{H} = (\Delta\hat{H}_f^0)_{H_2O(g)} + 125\overline{C}_p(150°C)$$

Taking $\Delta\hat{H}_f^0$ from Table 9.1 of Ref. 1 and $\overline{C}_p$ from Table 4.7,

TABLE 4.7   Mean Heat Capacities of Combustion Gases

$\overline{C}_p(J/mol \cdot °C)$; reference state: $P_{ref} = 1$ atm, $T_{ref} = 25°C$

$T(°C)$	Air	O_2	N_2	H_2	CO	CO_2	H_2O
0	28.94	29.24	29.03	28.84	29.00	36.63	33.55
25	29.05	29.39	29.06	28.84	29.06	37.15	33.63
100	29.21	29.80	29.16	28.86	29.23	38.63	33.92
200	29.45	30.32	29.32	28.90	29.47	40.45	34.34
300	29.71	30.80	29.52	28.95	29.72	42.10	34.80
400	29.97	31.24	29.74	29.03	29.99	43.59	35.29
500	30.25	31.65	29.98	29.12	30.27	44.93	35.81
600	30.53	32.02	30.24	29.23	30.56	46.14	36.36
700	30.81	32.39	30.51	29.35	30.85	47.23	36.92
800	31.10	32.71	30.79	29.48	31.14	48.20	37.49
900	31.38	33.02	31.07	29.63	31.42	49.07	38.08
1000	31.65	33.30	31.34	29.78	31.70	49.85	38.66
1100	31.92	33.55	31.62	29.94	31.97	50.54	39.24
1200	32.18	33.79	31.88	30.12	32.23	51.18	39.81
1300	32.42	34.02	32.13	30.29	32.47	51.75	40.37
1400	32.65	34.23	32.37	30.47	32.69	52.28	40.91
1500	32.85	34.42	32.58	30.66	32.89	52.77	41.42

Source: From Felder and Rousseau,[7] which in turn was taken from data presented in Himmelblau.[8]

$$\hat{H} = -241.83 + 4.27 = -237.56 \text{ kJ/mol}$$

Evaluate ΔH. From Eq. (4.11)

$$\Delta H = \sum_{\text{out}} n_i \hat{H}_i - \sum_{\text{in}} n_i \hat{H}_i = -15,300 \text{ kJ}$$

For the energy balance, neglecting kinetic energy changes

$$Q = \Delta H = -15,300 \text{ kJ}$$

Choice 2. Reference conditions: reactant and product species at T_0 in the state of aggregation for which $\Delta \hat{H}_r^0$ is known, and nonreactive species at any convenient temperature (such as the reactor inlet temperature, or the reference temperature of a mean heat capacity table).

If these reference conditions are chosen, the enthalpy change for the process may be determined by setting up and filling in a table of inlet and outlet stream component flow rates n_i and specific enthalpies $\hat{H}_i$, and calculating

$$\Delta H = \frac{n_{AR} \Delta \hat{H}_r^0}{v_A} + \sum_{\text{outlet}} n_i \hat{H}_i - \sum_{\text{inlet}} n_i \hat{H}_i \qquad (4.12)$$

where A = any reactant or product
   n_{AR} = moles of A produced or consumed in the process (not necessarily moles fed or moles present in the product)
   v_A = stoichiometric coefficient of A

Both n_{AR} and v_A are positive numbers.

If multiple reactions occur, a term of the form $n_A \Delta \hat{H}_r^0 / \omega_A$ must be included in Eq. (4.12) for each reaction. (Method 1 is generally easier to use in such a case.)

Example 4.3 *Simultaneous material and energy balances.* The ethanol dehydrogenation reaction is carried out with the feed entering at 300°C. The feed contains 90 mol % ethanol and the balance acetaldehyde. To keep the temperature from dropping too rapidly and hence quenching the reaction at a low conversion, heat is added to the reactor. Observe that when the heat addition rate is 5300 kJ per 100 mol of the feed gas, the outlet temperature is 265°C. Calculate the fractional conversion of ethanol achieved in the reactor using the heat capacity data given in Tables 4.6 and 4.7. A flow chart of the process is shown in Fig. 4.4.

Solution. Basis: 100 mol feed gas; also,

$$C_2H_5OH(g) \rightarrow CH_3CHO(g) + H_2(g): \Delta \hat{H}_r^0 = 68.95 \text{ kJ/mol}$$
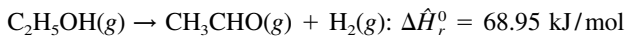
There are three unknowns, and no three material balance equations will enable one to
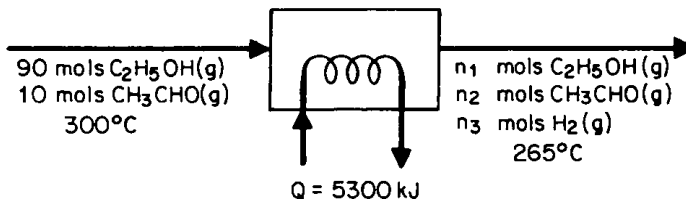

FIGURE 4.4 Flow chart of ethanol dehydrogenation process.

solve for n_i, n_2, and n_3. The problem must be solved by writing and simultaneously solving two material balance equations and an energy balance equation.

Balance on C:

$$(90)(2) + (10)(2) = 2n_1 + 2n_2$$

$$n_1 + n_2 = 100 \tag{1}$$

Balance on H:

$$(90)(6) + (10)(4) = 6n_1 + 4n_2 + 2n_3$$

$$6n_1 + 4n_2 + 2n_3 = 580 \tag{2}$$

Energy balance (at 25°C):

Substance	n_{in}	$\hat{H}_{in}$	n_{out}	$\hat{H}_{out}$	
$C_2H_5OH(g)$	90	30.3	n_1	26.4	n in mol
$CH_3CHO(g)$	10	22.0	n_2	19.2	$\hat{H}$ in kJ/mol
$H_2(g)$			n_3	7.0	

$\hat{H} = \overline{C}_p(T - 25)$: heat capacities from Table 4.7. The table* provides the following information:

$(C_2H_5OH)_{inlet}$:

$$\hat{H} = (0.110 \text{ kJ/mol} \cdot °C)(300°C - 25°C) = 30.3 \text{ kJ/mol}$$

$(CH_3CHO)_{inlet}$:

$$\hat{H} = (0.080)(275) = 22.0 \text{ kJ/mol}$$

$(C_2H_5OH)_{outlet}$:

$$\hat{H} = (0.110)(265 - 25) = 26.4 \text{ kJ/mol}$$

$(CH_3CHO)_{outlet}$:

$$\hat{H} = (0.080)(240) = 19.2 \text{ kJ/mol}$$

$(H_2)_{outlet}$:

$$\hat{H} = (0.029)(240) = 7.0 \text{ kJ/mol}$$

$Q = \Delta H$. Thus from Eq. (4.12):

---

*Heat capacities for C_2H_5OH and CH_3CHO are 0.110 kJ/kmol and 0.080 kJ/kmol respectively.

$$Q = \frac{(n_{H_2})_{out}\Delta\hat{H}_r^0}{v_{H_2}} + \sum_{out} n_i\hat{H}_i - \sum_{in} n_i\hat{H}_i$$

$$5300 = \frac{n_3(68.95)}{1} + 26.4n_1 + 19.2n_2 + 7.0n_3 - (90)(30.3) - (10)(22.0)$$

$$26.4n_1 + 19.2n_2 + 7.6n_3 - 8247 \qquad (3)$$

Solving Eqs. (1) through (3) simultaneously yields

$$n_1 = 7.5 \text{ mol } C_2H_5OH(g)$$

$$n_2 = 92.5 \text{ mol } CH_3CHO(g)$$

$$n_3 = 82.5 \text{ mol } H_2(g)$$

The fractional conversion of ethanol is

$$x = \frac{(n_{C_2H_5OH})_{in} - (n_{C_2H_5OH})_{out}}{(n_{C_2H_5OH})_{in}} = \frac{90 - 7.5}{90} = 0.917$$

Note on Heat Capacity. If the change in specific enthalpy of a substance that goes from T_1 to T_2 is $\hat{H}_2 - \hat{H}_1$, a mean heat capacity $\overline{C}_p$ may be defined as

$$\overline{C}_p = \frac{\hat{H}_2 - \hat{H}_1}{T_2 - T_1} \qquad (4.13)$$

Substituting for $\hat{H}_2 - \hat{H}_1$ from the relation

$$\Delta\hat{H} = \int_{T_1}^{T_2} C_p(T) \, dT$$

one gets

$$\overline{C}_p(T_1 \rightarrow T_2) = \frac{\int_{T_2}^{T_2} C_p(T) \, dT}{(T_2 - T_1)} \qquad (4.14)$$

Once $\overline{C}_p$ is known for the specified change in temperature (e.g., from a table of mean heat capacities), one may calculate the enthalpy change as

$$\Delta\hat{H} = \overline{C}_p\Delta T \qquad (4.15)$$

A simple multiplication thus replaces integration of $C_p \, dT$ in the calculation of $\Delta\hat{H}$. [Note, however, that if $\overline{C}_p$ is not known and the calculation of $\Delta\hat{H}$ is to be done only once, there is little point in first determining $\overline{C}_p$ from Eq. (4.13) since the numerator of this equation is the desired quantity.]

Table 4.7 gives mean heat capacities for the transitions of several common gases from a single base temperature T_{ref} to higher temperatures. These tables may be used to calculate the specific enthalpy of any of the tabulated species relative to the reference state as $\hat{H} = \overline{C}_p(T - T_{ref})$; the table may also be used to calculate $\Delta\hat{H}$ for a change between any two temperatures T_1 and T_2:

$$\Delta \hat{H}(T_1 \rightarrow T_2) = \hat{H}(T_2) - \hat{H}(T_1)$$

$$= (\overline{C}_p)_{T_2}(T_2 - T_{\text{ref}}) - (\overline{C}_p)_{T_1}(T_1 - T_{\text{ref}}) \qquad (4.16)$$

Introducing

$$C_p = a + bT + cT^2 + dT^3 \qquad (4.17)$$

in Eq. (4.14) allows one to calculate $\overline{C}_p$ when coefficients a, b, c, and d are given.

4. Basic Combustion Equations. All substances require the presence of oxygen for burning to take place. Any substance burning in air will combine with the oxygen. This process is called combustion and is an example of a chemical reaction between the burning substance and the oxygen in the air, the reaction producing heat. The chemical reaction is called **oxidation.**

The following are the basic equations normally used for combustion processes.

Carbon: $C + O_2 \rightarrow CO_2$; $2\,C + O_2 \rightarrow 2\,CO$
Hydrogen: $2\,H_2 + O_2 \rightarrow 2\,H_2O$
Sulfur: $S + O_2 \rightarrow SO_2$
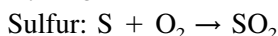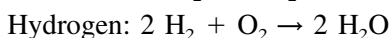
Typical hydrocarbon fuels:

$C_4H_8 + 6\,O_2 \rightarrow 4\,CO_2 + 4\,H_2O$
$C_2H_6O + 3\,O_2 \rightarrow 2\,CO_2 + 3\,H_2O$
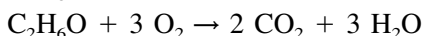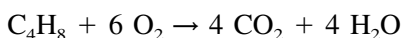
Carbon with air (assuming that air is composed of 79% nitrogen and 21% oxygen by volume):

$$C + O_2 + \frac{79}{21}\,N_2 \rightarrow CO_2 + \frac{79}{21}\,N_2 \text{ (by volume)}$$

$$12\,C + 32\,O_2 + \frac{28 \times 79}{21}\,N_2 \rightarrow 44\,CO_2 + \frac{28 \times 79}{21}\,N_2 \text{ (by mass)}$$
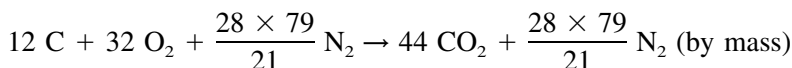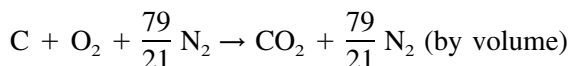
since the molecular weights of C, O_2, CO_2 and N_2 are 12, 32, 44 and 28.

Heating (Caloric) Value of Fuels. The calorific value of a fuel is the quantity of heat obtained per kilogram (solid or liquid) or per cubic meter (gas) when burnt with an excess of oxygen in a calorimeter (Table 4.7a).

If H_2O is present in the products as a liquid then the 'higher calorific value' (HV) is obtained. If the H_2O is present as a vapor then the 'lower calorific value' (LV) is obtained. LV $=$ HV $-$ 207.4% H_2 (by mass).

CHEMICAL EQUILIBRIUM

1. Equilibrium Constant. The equilibrium constant for the reaction

$$aA + bB = cC + dD$$
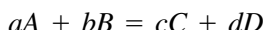
where the reaction is in solution is

TABLE 4.7a Caloric Value of Fuels

	Higher heating value	Lower heating value
Solid (kJ/kg; 15°C)		
Anthracite	34 600	34 000
Bituminous coal	33 500	32 450
Coke	30 750	30 500
Lignite	21 650	20 400
Liquid (kJ/kg; 15°C)		
Petrol (gasoline)	47 000	43 900
Benzole (crude benzene)	42 000	40 200
Kerosene (paraffin)	46 250	43 250
Diesel	46 000	43 250
Light fuel oil	44 800	42 100
Heavy fuel oil	44 000	41 300
Gas (MJ/m³, 15°C, 1 bar)		
Coal gas	20.00	18.00
Producer gas	6.04	6.00
Natural gas	36.20	32.60
Blast-furnace gas	3.41	3.38
Carbon monoxide	11.80	11.80
Hydrogen	11.85	10.00

Source: Ref. 6

$$K_c = \frac{[C]^c[D]^d}{[A]^a[B]^b} \tag{4.18}$$

([ ] refers to molarity). If reaction is in the gas phase, the equilibrium constant is

$$K_p = \frac{p_C^c \cdot p_D^d}{p_A^a \cdot p_B^b} \tag{4.19}$$

where p = partial pressure.

Note that equilibrium constants are dimensional; their units are useful in reconstructing the equilibrium constant expression and the reaction equation on which it is based.

2. Gibbs Free Energy Change and Reactivity. For the reaction

$$aA + bB \rightleftharpoons cC + dD$$
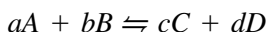
standard free energy is

$$(\Delta G)^0 = -RT \ln \frac{(a_C^c)(a_D^d)}{(a_A)^a(a_B)^b} = -RT \ln K_a \tag{4.20}$$

where a = activity, ( ) refers to molality, and the superscript ^{0} designates standard state at 25°C (298 K) and unit activity or 1-atm fugacity.

- If $(\Delta G)^0$ is negative, the reaction can occur as written (a forward driving force).
- If $(\Delta G)^0$ is zero, no driving force exists (the system is at equilibrium).
- If $(\Delta G)^0$ is positive, a reverse reaction can occur (a reverse driving force).

Values of $(\Delta G)^0$ are tabulated in many handbooks under the "standard free energy change" for the reaction (see Refs. 2 and 3).

PHASE EQUILIBRIA*

See Chap. 7 for information on temperature-pressure phase relations, equation of state, gas mixture, and Gibb's phase rule.

Ideal Solutions.   The activity of each constituent of ideal liquid solutions is equal to its mole fraction under all conditions of temperature, pressure, and concentration. The total volume of the solution exactly equals the sum of the volumes of its components. The enthalpy when the components are mixed is zero. The total vapor pressure is the sum of the contribution of the individual components following Raoult's law: the vapor pressure contribution of each individual component is the product of its mole fraction and the vapor pressure of the pure component. This also applies to the vapor pressure of solutions containing nonvolatile components. The freezing point of the solvent in ideal solutions occurs at the temperature where the vapor pressure of the solution equals the vapor pressure of the solid solvent.

Real Solutions.   Actual liquid solutions are seldom ideal, often showing deviations from the conditions of ideality described above. Most significant are positive or negative deviations in the direct summation of vapor pressure component contributions; these affect distillation behavior in the separation of components. Deviations from ideality increase with solute concentration; i.e., dilute solutions behave reasonably ideally.

Henry's Law.   At a constant temperature, the concentration of a gas dissolved in a liquid is directly proportional to the partial pressure of the gas above the liquid.

Raoult's Law.   This states that

$$p_a = x_a P_a \qquad\qquad (4.21)$$

where p_a = partial pressure of component A in vapor
$\quad x_a$ = mole fraction of A in liquid solution
$\quad P_a$ = vapor pressure of pure liquid A

Binary Solution Vapor-Liquid Equilibria. In the vapor phase:

$$P_{\text{total}} = p_i + p_j \tag{4.22}$$

$$n_{\text{total}} = n_i + n_j$$

$$y_i = \frac{p_i}{P_{\text{total}}}$$

$$y_i + y_i = 1$$

where p_i is the partial pressure, y_i is the mole fraction, and n_i is the number of moles of i.

In the liquid phase:

$$x_i = \frac{n_i}{n_{\text{total}}} \tag{4.23}$$

$$x_i + x_j = 1$$

$$n_{\text{total}} = n_i + n_j$$

where x_i is the mole fraction of i.

Ideal Solutions. Each component of an ideal solution obeys Raoult's law [Eq. (4.21)] relating concentrations in vapor and liquid phases.

Real Solutions. Numerical distillation calculations often use a vapor-liquid equilibrium ratio k for each component:

$$k_i = \frac{y_i}{x_i} = \frac{f \text{ of pure } i \text{ at its vapor pressure at } T \text{ of the system}}{f \text{ of pure } i \text{ at } T, P \text{ of the system}} \tag{4.24}$$

The volatility ratio between components of binary solutions is:

$$\alpha = \frac{k_i}{k_j} = \frac{y_i x_j}{x_i y_i} \tag{4.25}$$

where α is the volatility ratio. For ideal binary solutions, where α is constant, manipulation of Eqs. (4.24) and (4.25) relates mole fraction in the vapor phase y_i to volatility ratio α and mole fraction in the liquid phase x_i:

$$y_i = \frac{\alpha x_i}{1 + x_i(\alpha - 1)} \tag{4.26}$$

Binary Solution Vapor Pressure Composition Diagrams

Ideal Solutions. These follow Raoult's law [Eq. (4.21)]. As shown in Fig. 4.5*a*, the vapor pressure of each component is linear and proportional to the mole fraction, and the vapor pressure of the mixture is the simple sum of the component vapor pressures. The diagrams shown represent one fixed temperature.
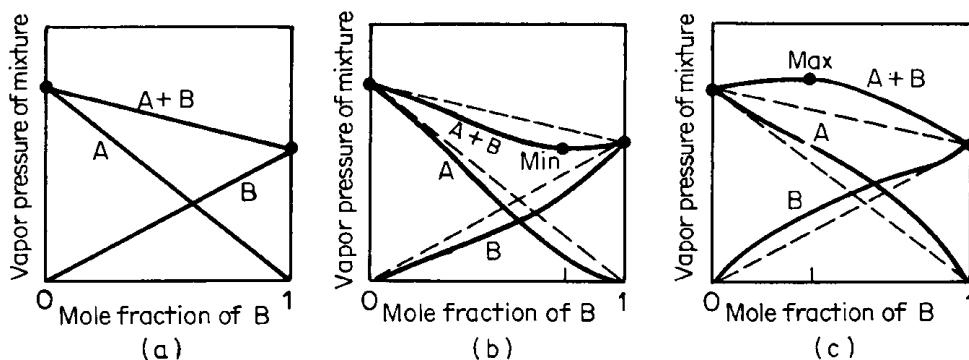
FIGURE 4.5   Binary solution vapor pressure composition diagrams.

Real Solutions.   These show significant deviations from linearity of individual vapor pressures with the mole fraction. At low mole fractions in the liquid phase, Henry's law is followed, while at high mole fractions, Raoult's law tends to be followed. These deviations from linearity are shown in Fig. 4.5*b* and *c*.

Boiling Point Composition Diagrams.   Figure 4.6 shows three types of boiling point composition diagrams. These are drawn for a single pressure. Figure 4.6*a* is the usual case that exists for ideal solutions; it corresponds with the vapor pressure diagram shown in Fig. 4.5*a*. At any given temperature, vapor composition *y* is in equilibrium with liquid composition *x*. This ideal type of boiling point diagram exists for many combinations of chemically similar materials.

Figure 4.6*b* and *c* correspond to the vapor pressure diagrams shown in Fig. 4.5*b* and *c*.

Azeotropes (constant boiling mixtures) exist at 1 and 2. Their composition can be altered by changing pressure. These arise from nonideality of the solutions. For more data on azeotropes, see Refs. 2, 9, and 10.
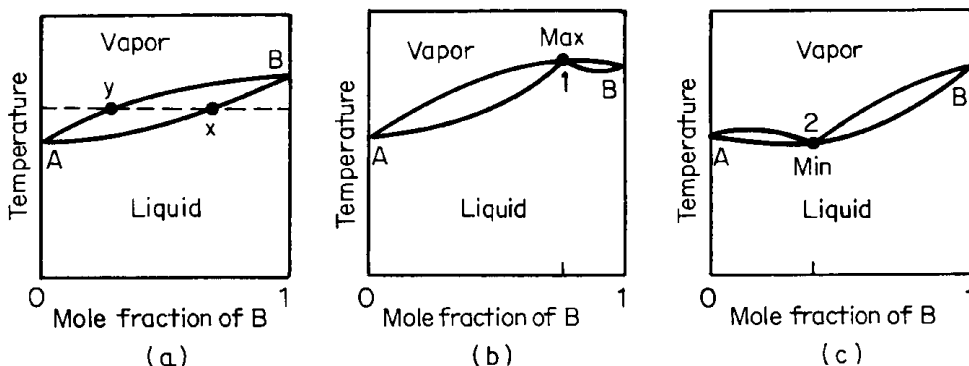
FIGURE 4.6   Binary solution boiling point composition diagrams.

CHEMICAL REACTION RATES*

1. Kinetics

Rate of Reaction.   The rate of reaction of any component A based on unit volume of fluid is

$$r_A = \frac{1}{V}\frac{dN_A}{dt} \qquad (4.27)$$

and where density remains unchanged

$$r_A = \frac{dC_A}{dt} \qquad (4.28)$$

Frequently

$$r_A = \begin{pmatrix} \text{temperature-} \\ \text{dependent term} \end{pmatrix}\begin{pmatrix} \text{concentration-} \\ \text{dependent term} \end{pmatrix} = kf(C_A, C_B, \ldots) \qquad (4.29)$$

N_A = number of moles of component A

Order, Molecularity, Elementary Reactions.   Where the rate can be expressed as

$$-r_A = kC_A^a C_B^b \cdots \qquad (4.30)$$

the reaction is ath order with respect to A and nth order overall where $n = a + b + \cdots$.

   Note. $a, b, \ldots$ are empirically observed and are not necessarily equal to the stoichiometric coefficients. In the special case where $a, b, \ldots$ are the stoichiometric coefficients, the reaction is elementary: unimolecular ($n = 1$), bimolecular ($n = 2$), trimolecular ($n = 3$).

Rate Constant k and Temperature Dependency of a Reaction

$$k = (\text{conc})^{1-n}\,(\text{time})^{-1}$$

From Arrhenius's law the variation with temperature is

$$k = k_0 e^{-E/RT} \qquad \text{or} \qquad \ln\frac{k_2}{k_1} = \frac{E}{R}\left[\frac{1}{T_1} - \frac{1}{T_2}\right] \qquad (4.31)$$

where E is the activation energy of the reaction.

---

2. Homogeneous, Constant Fluid Density, Bath Kinetics

Irreversible First-Order Reaction. For the reaction $A \rightarrow$ products, with rate

$$-\frac{dC_A}{dt} = kC_A \quad \text{or} \quad \frac{dX_A}{dt} = k(1 - X_A) \tag{4.32}$$

the integrated form is

$$-\ln \left(\frac{C_A}{C_{A_0}} \right) = -\ln (1 - X_A) = kt \tag{4.33}$$

Irreversible Second-Order Reaction. For the reaction $A + B \rightarrow$ products with rate

$$-\frac{dC_A}{dt} = kC_A C_B \tag{4.34}$$

when $M = C_{B_0}/C_{A_0} \neq 1$, the integrated form is

$$\ln \left(\frac{C_B C_{A_0}}{C_{B_0} C_A} \right) = \ln \left(\frac{M - X_A}{M(1 - X_A)} \right) = (C_{B_0} - C_{A_0})kt$$

When $C_{A_0} = C_{B_0}$, the integrated form is

$$\frac{1}{C_A} - \frac{1}{C_{A_0}} = \frac{1}{C_{A_0}} \left(\frac{X_A}{1 - X_A} \right) = kt \tag{4.35}$$

Irreversible nth-Order Reaction. For the reaction with rate

$$-r_A = -\frac{dC_A}{dt} = kC_A^n \tag{4.36}$$

the integrated form for $n \neq 1$ is

$$C_A^{1-n} - C_{A_0}^{1-n} = (n - 1)kt \tag{4.37}$$

Reversible First-Order Reaction. For the reaction

$$A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} R \qquad K = \frac{k_1}{k_2} \tag{4.38}$$

with rate

$$-\frac{dC_A}{dt} = \frac{dC_R}{dt} = k_1 C_A - k_2 C_R$$

the integrated form is

$$-\ln \left(\frac{X_{Ae} - X_A}{X_{Ae}} \right) = -\ln \left(\frac{C_A - C_{Ae}}{C_{A_0} - C_{Ae}} \right) = (k_1 + k_2)t \tag{4.39}$$

Integration of Rate in General. For the reaction with rate

$$-r_A = -\frac{dC_A}{dt} = k \cdot f(C_A, C_B, \ldots) \tag{4.40}$$

the integration is

$$t = C_{A_0} \int_0^{X_A} \frac{dX_A}{(-r_A)} = \int_{C_{A_0}}^{C_A} \frac{dC_A}{kf(C_A, C_B, \ldots)} \tag{4.41}$$

which is solved analytically or graphically.

3. Bath Reaction with Changing Fluid Density. Where density change is proportional to the fractional conversion of any reactant *A* (isothermal system),

$$\frac{C_A}{C_{A_0}} = \frac{1 - X_A}{1 + \epsilon_A X_A} \tag{4.42}$$

where

$$\epsilon_A = \frac{(\text{volume where } X_A = 1) - V_{x_A = 0}}{V_{x_A = 0}} \tag{4.43}$$

The rate for any reactant *A* is then

$$-r_A = -\frac{1}{V}\frac{dN_A}{dt} = \frac{C_{A_0}}{(1 + \epsilon_A X_A)}\frac{dX_A}{dt} = kf(C_A, C_B, \ldots) \tag{4.44}$$

Integrating in the general case:

$$t = C_{A_0} \int_0^{X_A} \frac{dX_A}{(1 + \epsilon_A X_A)(-r_A)} \tag{4.45}$$

ELECTROCHEMISTRY*

1. General. The chemical energy of redox reactions can be equated at constant *T*, *P* with Gibbs free energy using Eq. (4.4).

Since *dT* and *dP* are both zero at constant *T*, *P* then $dG = \Sigma\hat{\mu}_i \, dn_i$. Gibbs free energy can then be manifested as electric energy by an electron flow through an external conductor provided there is electrical isolation of the oxidation reaction (electron source) and reduction reaction (electron sink). The external conductor transfers electrons from source to sink, and a liquid junction permits internal migration of ions for preservation of charge neutrality of the system as a whole.

Chemical energy at constant T, P can be considered convertible to electric and /or thermal energy, for it can be shown by conservation of energy that $\Sigma \hat{\mu}_i \, dn_i = -E \, dq - T \, dS$. In the absence of $T \, dS$ thermal energy:

$$dG = -E \, dq \qquad (4.46)$$

where E is volts, dq is coulombs/gram-mole, and dG is energy in joules/grammole.

2. Redox Cells

Definitions of Anode and Cathode

Anode: Oxidation occurs at this electrode; it is the site where the reducing agent is oxidized; it is the source of electrons to the external circuit whether in a battery or an electrolytic cell.

Cathode: Reduction occurs at this electrode; its characteristics are the reverse of those of the anode.

Batteries. Batteries are redox cells that are physically arranged for external flow and internal ion mobility. When different electrolytes are used, gel structures or a porous membrane prevent mixing. The electrolyte(s) must permit ionization of the reactant species and be conductive. Mobility of ions in solution provides the mechanism for maintaining a charge balance and electrical neutrality within the electrolyte. By convention, batteries are labeled negative at the anode of the spontaneous discharge reaction.

Battery reactions on discharge are listed below. Only a few are reversible.

1. *Lead storage battery* (H_2SO_4 *electrolyte*) (*reversible*): Oxidation reaction at lead plates, labeled electrically negative:

$$Pb + HSO_4^- \rightarrow PbSO_4 + 2e^- + H^+ \qquad E_0 = +0.36 \text{ V}$$
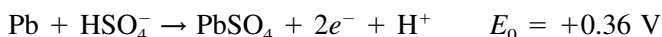
Reduction reaction at lead dioxide plates, labeled electrically positive:

$$PbO_2 + 3H^+ + HSO_4^- + 2e^- \rightarrow PbSO_4 + 2H_2O \qquad E^0 = +1.68V$$
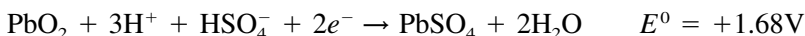
2. *Mercury cell* (*KOH electrolyte saturated with ZnO*): Oxidation (negative terminal):

$$Zn + 4OH^- \rightarrow ZnO_2^{2-} + 2H_2O + 2e^- \qquad E_B^0 = +1.22 \text{ V}$$

Reduction (positive terminal):

$$HgO + H_2O + 2e^- \rightarrow Hg + 2OH^- \qquad E_B^0 = +0.10 \text{ V}$$
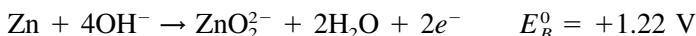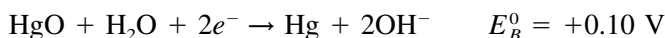
3. *Zinc-silver peroxide cell* (*KOH electrolyte saturated with ZnO*): Oxidation (negative terminal):

$$Zn + 4OH^- \rightarrow ZnO_2^{2-} + 2H_2O + 2e^- \qquad E_B^0 = + 1.22 \text{ V}$$

Reduction (positive terminal):

$$Ag_2O + H_2O + 3e^- \rightarrow 2Ag + 2OH^- \qquad E_B^0 = +0.34 \text{ V}$$
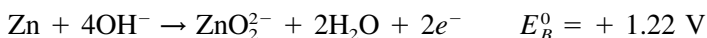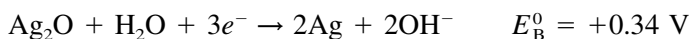
4. *LeClanche cell* (*flashlight battery, NH_4Cl electrolyte*): Oxidation (negative terminal):

$$Zn \rightarrow Zn^{2+} + 2e^- \qquad E^0 = +0.76 \text{ V}$$

Reduction (positive terminal):

$$2MnO_2 + 2NH_4^+ + 2e^- \rightarrow Mn_2O_3 + H_2O$$

$$+ 2NH_3 \text{ (for complexing } Zn^{2+}) \qquad E^0 = +0.74 \text{ V}$$
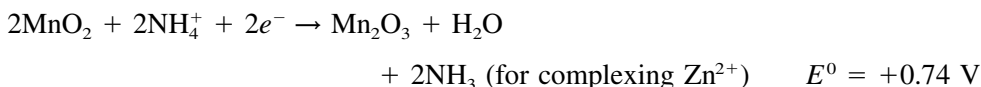
5. *Alkaline-zinc-manganese dioxide "alkaline flashlight battery"* (*KOH electrolyte*): Oxidation (negative terminal):

$$Zn + 4OH^- \rightarrow ZnO_2^{2-} + 2H_2O + 2e^- \qquad E_B^0 = +1.22 \text{ V}$$

Reduction (positive terminal):

$$MnO_2 + 2H_2O + e^- \rightarrow Mn(OH)_3 + OH^- \qquad E_B^0 = +0.35 \text{ V}$$
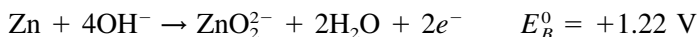
6. *Nickel-cadmium storage battery* (*KOH electrolyte*) (*reversible*): Oxidation (negative terminal):

$$Cd + 2OH^- \rightarrow Cd(OH)_2 + 2e^- \qquad E_B^0 = + 0.81 \text{ V}$$

Reduction (positive terminal):

$$NiO_4 + 2H_2O + 2e^- \rightarrow Ni(OH)_2 \qquad E_B^0 = + 0.49 \text{ V}$$
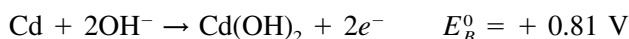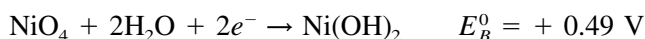
In all these equations

$$E^0 = \text{standard potential (acid solution)}$$

$$E_B^0 = \text{standard potential (basic solution)}$$

3. Summary: Chemical Effects of Electricity.   Therefore, a material must contain charged particles to be able to conduct electric current. In solids, electrons carry the current. Copper, lead, aluminum, iron and carbon are some examples of solid conductors. In liquids and gases the current is carried by the part of a molecule that has acquired an electric charge called ions. These can possess a positive or negative charge, and examples include hydrogen ion H^+, copper ion Cu^{++} and hydroxyl ion OH^-. Distilled water contains no ions and is a poor conductor of electricity, whereas salt water contains ions and is a fairly good conductor of electricity.

Electrolysis.   Electrolysis is the decomposition or a liquid compound by the passage of electric current through it. Practical applications of electrolysis include the electroplating of metals, the refining of copper and the extraction of aluminium from its ore. An electrolyte is a compound that will undergo electrolysis. Examples include salt water, copper sulphate and sulphuric acid. The electrodes are the two conductors carrying current to the electrolyte. The positive-connected electrode is called the anode and the negative-connected electrode the cathode. When two cop-

per wires connected to a battery are placed in a beaker containing a salt-water solution, current will flow through the solution. Air bubbles appear around the wires as the water is changed into hydrogen and oxygen by electrolysis.

Electroplating.   Electroplating uses the principle of electrolysis to apply a thin coat of a metal to another metal. Some practical applications include the tin-plating of steel, silver-plating of nickel alloys and chromium plating of steel. If two copper electrodes connected to a battery are placed in a beaker containing copper sulphate as the electrolyte it is found that the cathode (i.e., the electrode connected to the negative terminal of the battery) gains copper whilst the anode loses copper.

Corrosion.   Corrosion is the gradual destruction of a metal in a damp atmosphere by means of simple cell action. In addition to the presence of moisture and air required for rusting, an electrolyte, an anode and a cathode are required for corrosion. Thus, if metals widely spaced in the electrochemical series are used in contact with each other in the presence of an electrolyte, corrosion will occur. For example, if a brass valve is fitted to a heating system made of steel, corrosion will occur. The effects of corrosion include the weakening of structures, the reduction of the life of components and materials, the wastage of materials and the expense of replacement.

Rusting of iron (and iron-based materials) is due to the formation on its surface of hydrated oxide of iron produced by a chemical reaction. Rusting of iron always requires the presence of oxygen and water.

Corrosion Prevention.   Corrosion may be prevented by considering the following points.

Material Selection.   Metals and alloys that resist corrosion in a particular environment can be used. Proximity of metals with large potential difference, e.g. a copper pipe on a steel tank, should be avoided. Galvanic protection can be used, e.g. by use of a 'sacrificial anode' of zinc close to buried steel pipe or a ship's hull.

Appropriate Design.   Crevices that hold water, e.g. bad joints and incomplete welds, should be avoided as should high tensile stresses in material subject to stress corrosion. Locked-in internal stress due to forming should be avoided.

Modified Environment.   Metals can be enclosured against a corrosive atmosphere, water, etc. Drying agents, e.g. silica gel, and corrosion inhibitors, e.g. in central-heating radiators, can be used.

Protective Coating.   Metals can be coated to make them impervious to the atmosphere, water, etc. by use of a coating of grease, plasticizer, bitumen, resins, polymers, rubber latex, corrosion-resistant paints or metal coating.

Stress Corrosion Cracking.   Under tensile stress and in a corrosive environment some metals develop surface cracks called 'stress corrosion cracking,' which is time dependent and may take months to develop. It is avoided by minimizing stress and/or improving the environment.

Galvanic Corrosion.   For a pair of metals, that highest up the 'galvanic table' is the 'negative electrode' or 'cathode'; that lower down is the 'positive electrode' or 'anode.' The anode loses metal, i.e. corrodes, whilst the cathode remains unchanged. The greater the potential, the greater the rate of corrosion. Hydrogen is assumed to have zero potential (Table 4.7*b*).

TABLE 4.7b Galvanic Table for Metals (Relative to Hydrogen)

Metal		Potential difference (ΔV)	
Gold	NOBLE	+1.70	CATHODIC
Platinum		+0.86	
Silver		+0.80	
Copper		+0.34	
Hydrogen		0	
Lead		−0.13	
Tin		−0.14	
Nickel		−0.25	
Iron		−0.44	
Chromium		−0.74	
Zinc		−0.76	
Aluminium		−1.66	
Magnesium		−2.34	
	BASE		ANODIC

Source: Ref. 6.

ORGANIC CHEMISTRY

Reactions. Organic reactions relate to changes in the bonding to carbon atoms; their course is greatly influenced by reaction conditions that are experimentally established on the basis of reaction mechanism studies. Complications arise because several different reactions may occur simultaneously, accompanied by cyclization, molecular rearrangements, and oxidation-reduction reactions. All occur in the direction of minimum energy and increased stability. There is an encyclopedic literature on the subject, replete with tens of thousands of reactions, several hundred of which are generally useful in synthesis and have been given names in honor of early investigators in the field.

General Classes of Compounds. These include

1. The straight and branched chain types of compounds (see Table 4.8)
2. Cyclic compounds (see Table 4.9)

Note. For conciseness the following symbols are used in diagrammatic representations of compounds:

R = H atom or saturated hydrocarbon group
R' = hydrocarbon group only
X = halogen
n = an integer

TABLE 4.8 Straight and Branched Chain Types of Compounds

Type of name	General formula
1. Alkane or paraffin (also saturated hydrocarbons)	$\begin{array}{c} R \quad R \\ \diagdown \diagup \\ C \\ \diagup \diagdown \\ R \quad R \end{array}$
2. Alkene or olefin (unsaturated hydrocarbons)	$\begin{array}{c} R \quad\quad R \\ \diagdown \quad\quad \diagup \\ C=C \\ \diagup \quad\quad \diagdown \\ R \quad\quad R \end{array}$
3. Alkyne	$R-C\equiv C-R$
4. Alcohol	$\begin{array}{c} R \quad OH \\ \diagdown \diagup \\ C \\ \diagup \diagdown \\ R \quad R \end{array}$
5. Ether	$R-O-R'$
6. Aldehyde	$\begin{array}{c} H \\ \mid \\ R-C=O \end{array}$
7. Ketone	$\begin{array}{c} O \\ \parallel \\ R'-C-R' \end{array}$
8. Carboxylic acid	$\begin{array}{c} O \\ \parallel \\ R-C-OH \end{array}$
9. Grignard reagent	$\begin{array}{c} R \\ \diagdown \\ R-C-Mg-X \\ \diagup \\ R \end{array}$
10. Acyl halide	$\begin{array}{c} O \\ \parallel \\ R'-C-X \end{array}$
11. Anhydride	$\begin{array}{c} O \quad\quad O \\ \parallel \quad\quad \parallel \\ R-C-O-C-R \end{array}$
12. Ester	$\begin{array}{c} O \quad\quad R \\ \parallel \quad\quad \mid \\ R-C-O-C-R \\ \quad\quad\quad\quad \mid \\ \quad\quad\quad\quad R \end{array}$
13. Amide	$\begin{array}{c} O \\ \parallel \\ R-C-NH_2 \end{array}$

TABLE 4.8 Straight and Branched Chain Types of Compounds (*Continued*)

Type of name	General formula
14. Amine (base)	$\begin{array}{c} R \quad\quad R \\ \diagdown\!\!\diagup \\ C \\ \diagup\;\;\diagdown \\ R \quad\quad NH_2 \end{array}$
15. Nitrile	$\begin{array}{c} R \quad\quad R \\ \diagdown\!\!\diagup \\ C \\ \diagup\;\;\diagdown \\ R \quad\quad C\!\equiv\!N \end{array}$

TABLE 4.9 Cyclic Compounds

Type of name	General formula
1. Cycloparaffin (Naphthene)	
2. Cycloalkene	
3. Aromatic	
4. Naphthalenic	

NUCLEAR REACTIONS

The conventional notation for a nuclear reaction is

$$(_zB^A)_1 + (_zB^A)_2 \rightarrow (_zB^A)_3 + (_zB^A)_4 + Q$$
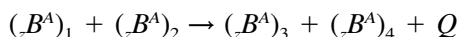
in which z is the number of protons, A is the mass number, B is the chemical symbol for the atom, electron, or nucleon, and Q is the energy released.

A specific example is the fission reaction:

$$_{92}U^{235} + _0n^1 \rightarrow _{92}U^{236} \rightarrow _{38}Sr^{94} + _{54}Xe^{140} + 2_0n^1 + Q$$
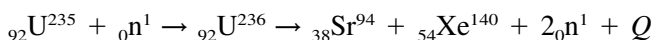
The strontium and xenon products are highly radioactive and decay further into other products. The final result is a spectrum of products.

The conservation equations are

$$\Sigma z_i = 0 \quad \text{and} \quad \Sigma A_i = 0$$

$$\text{Initial mass} - \text{final mass} = Q$$

BIOCHEMISTRY

The elements that make up the major organic constituents of the body are C, H, O, N, P, and S. These organic constituents are of two classes: very large molecules, which may be classed as elements of storage, structure, or function, and small molecules which are (1) the constituents of the macromolecules, (2) derivatives of these constituents with special functions (e.g., coenzymes, vitamins), and (3) sources of energy, through fermentation or oxidation processes.

The organic elements are also of two classes: (1) monovalent and divalent ions, which are major constituents of the body fluids: Na^+, K^+, Cl^-, HCO_3^-, Ca^{2+}, Mg^{2+}, HPO_4^{2-}, $H_2PO_4^-$; and (2) trace elements: Fe, Cu, Mn, Zn, I, Co, Mo, Si, F.

The most abundant single compound in the body is water, which makes up 65 to 70 percent of the lean body mass, or 55 percent of the whole body weight.

Cell. The unit of organization and of metabolic operations in the body is the cell. Every cell is bounded by a plasma membrane, which actively guards the internal environment of the cell by excluding certain materials from, allowing some to pass into, and expending energy to facilitate the transport of others through the cell. The fabric of the membrane is a bimolecular layer of lipid, mainly phosphoglyceride molecules presenting their polar ends toward the surface of the membrane and their hydrocarbon tails toward the interior and one another.

Proteins. The principal nitrogen-containing compounds of plant and animal tissues are the proteins. Their essential nature is illustrated in their role as biological catalysts, or enzymes, but they act also as constituents of membranes; as prominent components of connective tissues, bone matrix, cartilage, hair, horn, claws, nails, as parts of hormones, antigens, and antibodies; as osmotic regulators; and as both specific and general carriers, not only of the respiratory gases but also of a host of simpler organic compounds and many inorganic ions. Their elementary composition is simple: C = 50 to 55 percent; H = 6.0 to 7.3 percent; O = 19 to 24 percent; N

= 13 to 19 percent; S = 0 to 4 percent. Many other elements are associated with proteins in more or less firm combination, but only iodine is found as a constituent of an α-amino acid in a derived protein, thyroglobulin.

Proteins are polymers of α-amino acids. Their sizes vary enormously, from tiny proteins such as insulin (mol wt approximately 6000) to very large aggregates with molecular weights of many millions. The structure common to them all is the peptide bond, formed by the loss of water between the carboxyl group of one α-amino acid and the α-amino group of another. This yields a repeating backbone structure as follows:

$$
\ldots N{-}\underset{R}{\overset{H}{\underset{|}{\overset{|}{C}}}}{-}\overset{H}{\underset{}{\overset{|}{C}}}{-}\overset{O}{\underset{}{\overset{\|}{C}}}{-}\overset{H}{\underset{}{\overset{|}{N}}}{-}\underset{R}{\overset{H}{\underset{|}{\overset{|}{C}}}}{-}\overset{O}{\underset{}{\overset{\|}{C}}}\ldots
$$
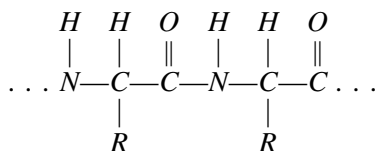
Compounds formed in this way are called *peptides*. Proteins, containing many such bonds, are called *polypeptides*.

In addition to the proteins, the other class of nitrogenous high-molecular-weight constituents of cells is the *nucleic acids*.

Enzymes. All of the chemical reactions in the body take place at a constant, relatively low temperature, and most of them would be infinitely slow unless they were catalyzed. A major function of proteins is to serve as the biological catalysts synthesized in living cells known as enzymes. The characteristics of a differentiated cell are determined not only by its structural proteins but also by the nature, variety, and amounts of its constituent enzymes and by their ordering into enzyme systems. Over 1500 enzymes are known, and many of these have been isolated in crystalline form and studied in detail. All enzymes have proven to be proteins.

Energy Exchange and Production. Living organisms are enclaves of order open to an environment continually tending to disorder. Organisms are maintained by the constant expenditure of energy derived from foodstuffs by processes that are, in the long run, oxidative. All of the life processes of growth, reproduction, repair, maintenance, and chemical, electrical, and mechanical work are supported by chemical reactions which, taking the entire system that can be described as animal + input + output, proceed with a net loss of free energy, that is, of the component of the total energy that is available for useful work. Thus, although an organism may maintain a constancy of composition, internal order, and total internal energy for a large part of its lifetime, it does so at the continuous expense of its environment.

The quantitative aspects of the energy exchanges of the whole organism and of the energy of individual chemical reactions are important to an understanding of bioenergetics in human beings.

NOMENCLATURE

Symbol = Definition, SI unit (U.S. Customary unit)

C_A = concentration of A, mol A/volume

C_{A_0} = initial concentration of A, mol A/volume

C_p = specific heat, J/mol · °C (Btu/mol · °F)

e^- = electron (sign $^-$ indicates negative charge)

F = Helmholtz function, J/mol (Btu/mol)

f = fugacity, N/m² (lb$_f$/ft²)

G = Gibbs free energy, J/mol (Btu/mol)

H = enthalpy, J/mol (Btu/mol)

H = enthalpy (in ''Thermochemistry'' section), J (Btu)

ΔH = enthalpy difference, J (Btu)

$\hat{H}$ = enthalpy (in ''Thermochemistry'' section), J/mol (Btu/mol)

N_A = number of moles (in ''Thermochemistry'' section)

n = number of moles

p_a = partial pressure, N/m² (lb$_f$/ft²)

P = pressure, N/m² (lb$_f$/ft²)

Q = heat, J (Btu)

R = gas constant, J/mol · K (Btu/mol · °R)

S = entropy, J/mol · K (Btu/mol · °F)

T = temperature, K (°R)

t = time, s

V = volume, m³ (ft³)

U = internal energy, J/mol (Btu/mol)

X_A = fraction of reactant A converted

X_a = mole fraction

ΔV = potential difference, V

Greek

$\hat{\mu}$ = chemical potential, J/mol (Btu/mol)

γ = activity coefficient

Subscripts

a = component A

A = component A

f = reaction

0 = standard condition

0 = initial conditions (1 atm, 25°C)

s = solid

g = gas

l = liquid

Others

[_] = concentration (C_A = [A])

=average

Note.   Other symbols used in this chapter are defined in the text.

*REFERENCES**

1. K. Gieck, *Engineering Formulas,* 5th ed., McGraw-Hill, New York, 1986.

2. J. A. Dean, *Lange's Handbook of Chemistry,* 13th ed., McGraw-Hill, New York, 1985.

3. R. H. Perry and D. W. Green, *Peny's Chemical Engineers' Handbook,* 6th ed., McGraw-HiIl, New York, 1984.

4. N. Radosevic, *Priručnik za hemičare i tehnologe* (*Handbook for Chemists*), Tehnička knjiga, Belgrade, 1968.

5. R. H. Perry, *Engineering Manual,* 3d ed., McGraw-HiIl, New York, 1976.

6. E. N. Ganić, *Tehnička hemija.* (*Chemistry for Engineers—Class Notes*), Dept. of Mechanical Engineering, University of Sarajevo, 1984–2002.

7. R. M. Felder and R. W. Rousseau, *Elementary Principles of Chemical Processes,* Wiley, New York, 1978.

8. D. M. Himmelblau, *Basic Principles and Calculations in Chemical Engineering,* 3d ed., Prentice-Hall, Englewood Cliffs, N.J., 1974.

9. E. U. Condon and H. Odishaw, *Handbook of Physics,* 2d ed., McGraw-Hill, New York, 1967.

10. S. P. Parker (ed.), *McGraw-Hill Encyclopedia of Physics,* McGraw-Hill, New York, 1983.

11. O. Levenshpiel, *Chemical Reaction Engineering,* 2d ed., Wiley, New York, 1969.

12. J. F. Mulligan, *Practical Physics,* McGraw-Hill, New York, 1980.

13. J. M. Smith and H. C. Van Ness, *lntroduction to Chemical Engineering Thermodynamics,* 2d ed., McGraw-Hill, New York, 1959.

14. J. Carvill, *Mechanical Engineer's Data Handbook,* Butterworth-Heinemann, 2001.

15. J. Bird, *Newnes Engineering Science Pocket Book,* 3d ed., Newnes, 2001.

16. E. A. Reeves, *Newnes Electrical Pocket Book,* 22d ed., Newnes, 2001.

17. R. A. Corbitt, *Standard Handbook of Environmental Engineering*, 2d ed., McGraw-Hill, 1999.

18. C. C. Lee and S. D. Lin, *Handbook of Environmental Engineering Calculations,* McGraw-Hill, 2000.

---

*Those references listed above but not cited in the text were used for comparison between different data sources, clarification, clarity of presentation, and, most important, reader's convenience when further interest in the subject exists.

CHAPTER 5
MECHANICS OF RIGID BODIES

1. Definitions

Statics. That branch of mechanics which deals with the equilibrium of forces on bodies at rest (or moving at a uniform velocity in a straight line).

Force. The action of one body on another. It is a vector quantity defined by its magnitude, direction, and point of application (*P*) (Fig. 5.1). The forces by which the individual particles of a body act on each other are *internal.* All other forces are *external.* The forces exerted by the body on its supports are called *reactions.* They are equal in magnitude and opposite in direction to the forces, called *supporting forces*, with which the supports act on the body.

Gravitation. The force of the earth's attraction directed toward the earth's center.

Moment. The moment of a force **F** about a point *O* (Fig. 5.7) is

$$\mathbf{M} = \mathbf{F} \cdot l \tag{5.1}$$

where l = perpendicular distance (called level arm l) from point *O* to the line of action of force **F**. This moment has the dimension of force × length and is considered positive when tending to produce counterclockwise rotation around point *O.*

Free-Body Diagram. A structure or a part of a structure which is considered to be separated from everything else and mode "free." When a free body is diagrammed, vector arrows are placed representing the forces which the other parts exert upon the free body. The forces which the free body exerts upon the other parts are not represented. See Fig. 5.2. (The roller at *B* in Fig. 5.2*a* is replaced in Fig. 5.2*b* by a vertical force F_B, since the roller offers constraint only in the vertical direction. Similarly. the pin at *A* is replaced by horizontal and vertical reactions F_{XA} and F_{YA}, since the pin offers two degrees of constraint. The drawing shown in Fig. 5.2*b,* in which the beam is shown completely isolated from its supports and where all forces are shown by vectors, is the actual *free-body diagram.*)

Basic Laws

Newton's Laws

1. If a body is at rest, it will remain at rest, or if in motion, it will move uniformly in a straight line, until acted on by some force.
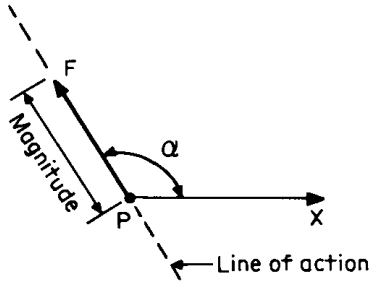
FIGURE 5.1
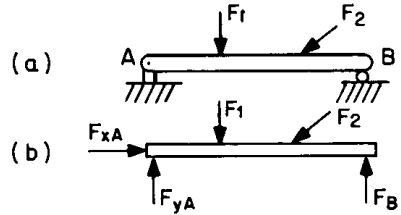
FIGURE 5.2

2. If a body is acted on by several forces, it will obey each as though the others did not exist, and this will happen whether the body is at rest or in motion. Change of the motion of a body is proportional both to the force exerted against it and to the time during which the force is exerted, and is in the same direction as the force.

3. If a force acts to change the state of a body with respect to rest or motion, the body will offer a resistance equal to and directly opposed to the force. Or, to every action there is opposed an equal and opposite reaction.

Linear Momentum. The product of mass and the linear velocity of a particle; it is a vector. The moment of the linear momentum vector about a fixed axis is the *angular momentum* of the particle about that fixed axis. For a rigid body rotating about a fixed axis, angular momentum is defined as the product of the moment of inertia and angular velocity, each measured about the fixed axis.

The relation between mass (m), acceleration (a), and force ($\mathbf{F}$) is contained in Newton's second law of motion:

$$\mathbf{F} = m \cdot \mathbf{a} \tag{5.2}$$

The direction of $\mathbf{F}$ is the same as the direction of $\mathbf{a}$. An alternative form of Newton's second law states that the resultant force is equal to the time rate of change of momentum:

$$\mathbf{F} = \frac{d(m \cdot \mathbf{v})}{dt} \tag{5.3}$$

The linear momentum of a system of bodies is unchanged if there is no resultant external force on the system. The angular momentum of a system of bodies about a fixed axis is unchanged if there is no resultant external moment about this axis.

2. Composition of Forces

Classification of System of Forces. The classification of a system of forces is made according to the arrangement of the forces' action lines. If the action lines lie in the same plane, the system is *coplanar;* otherwise it is *noncoplanar.* If they pass through the same point, the system is *concurrent;* otherwise it is *nonconcurrent.* If two or more forces have the same action line, they are *collinear.* A system

of two equal forces, parallel, opposite in sense, and having different action lines is a *couple.*

Graphical Composition of Forces. A force may be represented by a straight line in a determined position, and its magnitude by the length of the straight line. The direction in which it acts may be indicated by an arrow.

A parallelogram of two forces intersecting each other (Fig. 5.3*a*) leads directly to the graphical composition by means of a triangle of forces. In Fig. 5.3*b*, F_R is called the *closing side,* and represents the resultant of the forces of F_1 and F_2 in magnitude and direction. Its position is given by the point of application *O*. By means of the repeated use of the triangle of forces and by omitting the closing sides of the individual triangles, the magnitude and direction of the resultant F_R of any number of forces in the same plane and intersecting at a single point can be found (see Fig. 5.4). If the forces are in equilibrium, F_R must be equal zero, i.e., the force polygon must close.

If in a closed polygon one of the forces is reversed in direction, this force becomes the resultant of all the others.

If the forces do not all lie in the same plane, the diagram becomes a polygon in space. The resultant F_R of this system may be obtained by adding the forces in space. The resultant is the vector which closes the space polygon. The space polygon may be projected onto three coordinate planes, giving three related plane polygons. Any two of these projections will involve all the static equilibrium conditions and will be sufficient for a full description of the force system (see Fig. 5.5).

Mathematical Composition of Forces. The resultant F_R of two forces F_1 and F_2 applied at the same point is defined analytically as follows (Fig. 5.3*a*):

$$F_R = (F_1^2 + F_2^2 + 2F_1F_2 \cos \alpha)^{1/2} \tag{5.4}$$

where

(a)　　　　　　　　(b)
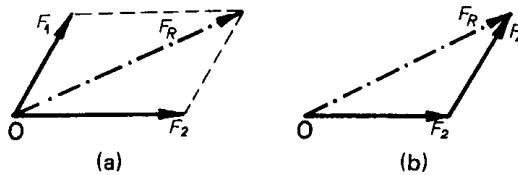
FIGURE 5.3 (*a*) Diagram of forces; (*b*) force polygon. (*From Giecki.*[1])

(a)　　　　　　　　(b)
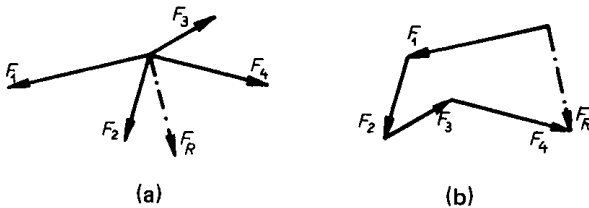
FIGURE 5.4 (*a*) Diagram of forces; (*b*) force polygon. (*From Gieck.*[1])
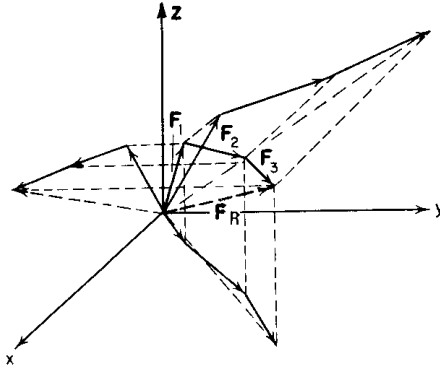
FIGURE 5.5 (*From Avallone and Baumeister.*[2])

$$\tan \alpha = \frac{F_2 \sin \alpha}{F_1 + F_2 \cos \alpha} \tag{5.5}$$

Force **F** may be resolved into two component forces (Fig. 5.6) as:

$$F = (F_x^2 + F_y^2)^{1/2} \tag{5.6}$$

where
$$F_x = F \cdot \cos \alpha$$

$$F_y = F \cdot \sin \alpha$$

$$\tan \alpha = \frac{F_y}{F_x}$$

Moment $\mathbf{M}_0$ of a force **F** about a point O (Fig. 5.7) is

$$M_0 = F \cdot l_0 = F_y \cdot \overline{x} - F_x \cdot \overline{y} \tag{5.7}$$

The resultant force F_R of any given forces is (Fig. 5.8):

$$F_R = (F_{Rx}^2 + F_{Ry}^2)^{1/2} \tag{5.8}$$

$$F_{Rx} = \Sigma F_x \qquad F_{Ry} = \Sigma F_y$$

$$\tan \alpha_R = \frac{F_{Ry}}{F_{Rx}}$$

$$\Sigma M_0 = F_R \cdot l_R \tag{5.9}$$

where l_R is distance of F_R from reference point.
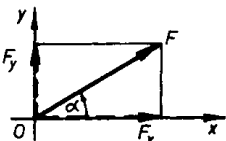
FIGURE 5.6 (*From Geick.*[1])   **FIGURE 5.7** (*From Gieck.*[1])   **FIGURE 5.8** (*From Gieck.*[1])
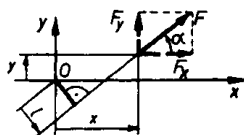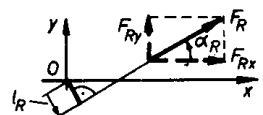
3. Conditions of Equilibrium.*   A body is in equilibrium when both the resultant force and the sum of the moments of all external forces about any random point are equal to zero.

Concurrent Forces in a Plane.   If a system of concurrent forces in a plane is in equilibrium, the sum of the components of the forces along any axis in the plane of the forces is equal to zero. Each force may be resolved into two rectangular components along the X and Y axes. Let F_x and F_y represent the X and Y components of force **F,** and let ΣF_x and ΣF_y represent the summation of all of the X and Y components, respectively. Then $\Sigma F_x = 0$ and $\Sigma F_y = 0$ are the two independent equations of equilibrium by means of which two unknown quantities may be determined.

EXAMPLE 5.1   In Fig. 5.9, $\overline{BC}$ is a boom which is hinged at C and supported by the tie bar $\overline{AB}$. The stresses in the boom $\overline{BC}$ and in the tie bar $\overline{AB}$ due to the load of $F_N = 1000$ lb$_f$ (4448 N) are required. Point B is the free body. Let the X axis be horizontal and the Y axis vertical:

$$\Sigma F_x = 0.966F_{AB} - 0.866F_{BC} = 0$$

$$\Sigma F_y = 0.500F_{BC} - 0.2588F_{AB} - F_N = 0$$

where $\alpha_1 = 15°$, $\alpha_2 = 30°$, and $l = 40$ ft (12.2 m); cos 15° = 0.966, and cos 30° = 0.866, sin 15° = 0.2588, and sin 30° = 0.500.

Solution.   Solving for these two equations gives

$$F_{AB} = 3350 \text{ lb}_f \text{ (14,900 N)} - \text{tension}$$

$$F_{BC} = 3730 \text{ lb}_f \text{ (16,591 N)} - \text{compression}$$

Example 5.1 could also have been solved by the method of moments. Since the free body is in equilibrium, the summation of the moments with respect to any axis is equal to zero. The equations $\Sigma M_A = 0$ and $\Sigma M_C = 0$ give the two unknown forces:
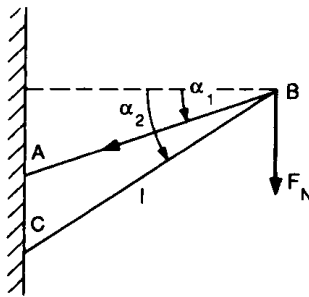
FIGURE 5.9

---

*Examples 5.1 through 5.5 in this Section taken partly from *General Engineering Handbook,* 2d ed., by C. E. O'Rourke. Copyright © 1940. Used by permission of McGraw-Hill, Inc. All rights reserved.

$$\Sigma M_A = F_N \times 34.64 - F_{BC} \times 9.3 = 0$$

$$\Sigma M_C = F_N \times 34.64 - F_{AB} \times 10.33 = 0$$

Thus

$$F_{BC} = 3730 \text{ lb}_f \text{ (16,591 N)} \qquad F_{AB} = 3350 \text{ lb}_f \text{ (14,900 N)}$$

Parallel Forces in a Plane. If a system of parallel forces in a plane is in equilibrium, the equation $\Sigma F = 0$ gives only one equation. The equation $\Sigma M = 0$ gives another independent equation, however, so two unknown quantities can be determined. The moment equation should be written first, with a point on the line of action of one of the unknown forces as the center of moments. The resulting equation has only one unknown quantity which is thus determined. The equation $\Sigma F = 0$ completes the solution.

EXAMPLE 5.2 Figure 5.10 represents a beam 32 ft (9.754 in) long with three vertical loads and two unknown vertical reactions, F_{R_1} and F_{R_2}. If

$$F_1 = 2000 \text{ lb}_f \text{ (8896 N)}$$

$$F_2 = 5000 \text{ lb}_f \text{ (22,240 N)}$$

$$F_3 = 6000 \text{ lb}_f \text{ (26,688 N)}$$

$$l_1 = 16 \text{ ft (4.87 m)}$$

$$l_2 = 12 \text{ ft (3.65 m)}$$

$$l_3 = 4 \text{ ft (1.22 m)}$$

$$l_4 = 7 \text{ ft (2.13 m)}$$

$$l_5 = 25 \text{ ft (7.62 m)}$$

then equation $\Sigma M_{R_2} = 0$ gives the following:
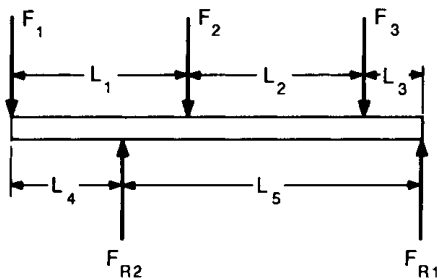
FIGURE 5.10

$$25F_{R_1} - 2000 \times 32 - 5000 \times 16 - 6000 \times 4 = 0$$

or

$$F_{R_1} = 6720 \text{ lb}_f \text{ (29,890 N)}$$

Equation $\Sigma F_y = 0$ gives the following:

$$F_{R_2} + 6720 - 13,000 = 0 \quad \text{or} \quad F_{R_2} = 6280 \text{ lb}_f \text{ (27,933 N)}$$

This result may be checked by using the equation $\Sigma M_{R_1} = 0$:

$$25F_{R_2} + 2000 \times 7 - 5000 \times 9 - 6000 \times 21 = 0$$

or

$$F_{R_2} = 6280 \text{ lb}_f \text{ (27,933 N)}$$

Any Force System in a Plane.   If a coplanar system of forces in equilibrium is nonconcurrent and nonparallel, three independent equations of equilibrium may be written, and therefore three unknown quantities may be determined. The three equations may consist of one moment equation and two force equations, of two moment equations and one force equation, or of three moment equations. Obtaining the solution is simplified if each equation is written in such a way as to contain only one unknown quantity.

EXAMPLE 5.3   In Fig. 5.11, the boom $\overline{CR}$ is hinged at C and supported by the tie bar $\overline{AB}$. The boom is the free body, and the three uknown quantities consist of the stress in AB and the amount and direction of reaction C. If reaction C is replaced by its vertical and horizontal components, the unknown quantities are three forces, whose directions are all known.
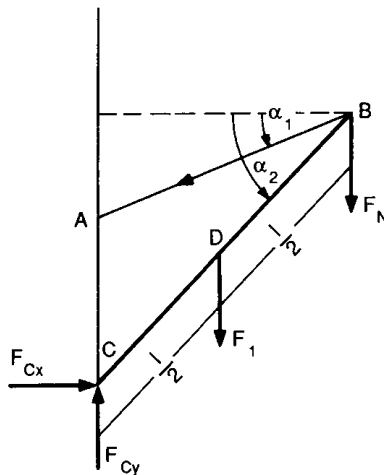
FIGURE 5.11

The equation $\Sigma M_C = 0$ gives the following [note that $l = 40$ ft (12.2 m), $F_1 = 2000$ lb$_f$ (8896 N), $F_N = 3000$ lb$_f$ (13,344 N), and $\alpha_1 = 15°$, $\alpha_2 = 30°$):

$$F_{AB} \times 40 \times 0.2588 - F_1 \times 20 \times 0.866$$
$$- F_N \times 40 \times 0.866 = 0$$

or $F_{AB} = 13,400$ lb$_f$ (59,603 N). The equation $\Sigma F_x = 0$ gives $F_{Cx} - 13,400 \times 0.966 = 0$, or $F_{Cx} = 12,900$ lb$_f$ (57,379 N).

The equation $\Sigma F_y = 0$ gives $F_{Cy} = 5000 - 13,400 \times 0.2588 = 0$, or $F_{Cy} = 8460$ lb$_f$ (37,630 N). Reaction $F_C = (F_{Cx}^2 + F_{Cy}^2)^{1/2} = 15,400$ lb$_f$ (68,499 N).

Values of F_{Cx} and F_{Cy} could have been obtained by writing the equations $\Sigma M_A = 0$ and $\Sigma M_B = 0$.

Concurrent Forces in Space. If a system of concurrent forces in space is in equilibrium, three independent equations may be written, and therefore three unknown quantities may be determined. The three equations may consist of three moment equations, of two moment equations and one force equation, of one moment equation and two force equations, or of three force equations. The solution is simplified if equations can be written containing only one unknown quantity.

EXAMPLE 5.4 If the weights of the members of the shear-legs crane (Fig. 5.12) are neglected, the force system is concurrent. The stresses in members AB, BC, and BE by the 1000-lb$_f$ (4448-N) load are required. Point B is the free body.

Given

$$l_1 = 40 \text{ ft (12.2 m)}$$

$$l_2 = 20 \text{ ft (6.1 m)}$$

$$l_3 = 6 \text{ ft (1.83 m)}$$

$$l_4 = 8 \text{ ft (2.44 m)}$$

$$\text{Length } \overline{BD} = (20^2 - 6^2)^{1/2} = 19.08 \text{ ft (5.81 m)}$$

$$\overline{BF} = (19.08^2 - 8^2)^{1/2} = 17.32 \text{ ft (5.28 m)}$$

$$\overline{FA} = (40^2 - 17.32^2)^{1/2} = 36.06 \text{ ft (10.99 m)}$$

$$\overline{DA} = 28.06 \text{ ft (8.55 m)}$$

and $$F_N = 1000 \text{ lb}_f \text{ (4448 N)}$$

Solution. With line CE as the axis, the moment equation is

$$F_{AB} \times 28.06 \times \frac{17.32}{40} - F_N \times 8 = 0$$

or $F_{AB} = 660$ lb$_f$ (2935 N), $F_{BC} = F_{BE}$ by symmetry.
The equation $\Sigma F_x = 0$ gives:
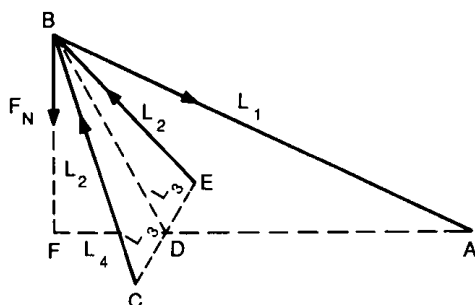
FIGURE 5.12
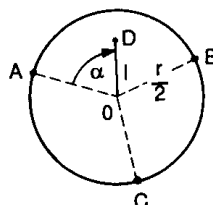
FIGURE 5.13

$$2F_{BC} \times 0.4 - 660 \times \frac{36.06}{40} = 0 \quad \text{or} \quad F_{BC} = 745 \text{ lb}_f \text{ (3313 N)}$$

Instead of the equation $\Sigma F_x = 0$, either equation $\Sigma F_y = 0$ or a moment equation could have been used. An axis through E parallel to AD would have been a satisfactory axis of moments.

Parallel Forces in Space. If a system of parallel forces in space is in equilibrium, three independent equations may be written; hence three unknown quantities may be determined. These three equations may consist of one force equation and two moment equations, or of three moment equations. The moment equation with respect to an axis intersecting two of the unknown forces will give the value of the third unknown force.

EXAMPLE 5.5 The table of diameter r shown in Fig. 5.13 is supported by legs A, B, and C, 120° apart. A load F_N is applied at point D. Find reactions due to the F_N load if:

$$r = 5 \text{ ft (1.52 m)}$$

$$l = 8 \text{ in (0.2032 m)}$$

$$\alpha = 45°$$

$$F_N = 200 \text{ lb}_f \text{ (889.6 N)}$$

Solution. Use BC as the axis of moments:

$$3.75F_A - 200 \times 1.721 = 0 \quad \text{or} \quad F_A = 91.8 \text{ lb}_f \text{ (408.3 N)}$$

Next, use an axis through C parallel to OA as the axis of moments. Then

$$4.33F_B + 91.8 \times 2.165 - 200 \times 2.636 = 0 \quad \text{or} \quad F_B = 75.9 \text{ lb}_f \text{ (337.6 N)}$$

$$F_y = 91.8 + 75.9 + F_C - 200 = 0 \quad \text{or} \quad F_C = 32.3 \text{ lb}_f \text{ (145.6 N)}$$

Any Force System in Space. If a force system in equilibrium consists of forces which are noncoplanar, nonconcurrent, and nonparallel, six independent equations of equilibrium may be written, three force equations, and three moment equations:

$$\Sigma F_x = 0 \qquad \Sigma F_y = 0 \qquad \Sigma F_z = 0 \tag{5.10}$$

$$\Sigma M_x = 0 \qquad \Sigma M_y = 0 \qquad \Sigma M_z = 0$$

This makes it possible to solve for six unknown quantities in such a system.

It is also convenient to apply the algebraical condition ΣM about every axis $= 0$ instead of using so many resolution equations. For graphical presentation, the projection of the system on any plane is in equilibrium, and algebraical and graphical (force polygon) conditions can be used to solve such projected systems.

Note. Combining the solutions discussed in the above five subsections, one can solve other more elaborate problems.

Flexible Cord: Load Uniformly Distributed Horizontally. Flexible cords, ropes, cables, or chains which are suspended horizontally between two points and are loaded uniformly horizontally take the shape of a parabola. If the cord carries equally concentrated loads evenly spaced horizontally, the smooth curve through the points of application of the loads will take the shape of a parabola, and the cord itself will not deviate very much from the curve. The suspension bridge is a good example of this type of loading, the chief weight carried being that of the roadway.

If F_N is the weight carried by a cord, l the span between supports, F the tension at the lowest point, and s the length of the cord, the length s is given by the rapidly converging series

$$s = l + \frac{F_N^2 \cdot l}{24F^2} - \frac{F_N^4 \cdot l}{640F^4} + \cdots \tag{5.11}$$

In terms of the span l and the sag d,

$$s = l + \frac{8d^2}{3l} - \frac{32d^4}{5l^3} + \cdots \tag{5.12}$$

When considering the right half of a loaded cord as the free body, the three forces must act through a common point.
The equation $\Sigma F_x = 0$ gives:

$$F_R \cdot \cos \alpha = F$$

The equation $\Sigma F_y = 0$ gives:

$$F_R \cdot \sin \alpha = \frac{F_N}{2}$$

If these two equations are squared and added, the resulting equation is

$$F_R^2 = F^2 + \left(\frac{F_N}{2}\right)^2$$

The equation $\Sigma M_c = 0$ gives:

$$F \cdot d - \frac{F_N \cdot l}{8} = 0$$

(a)
Stable
equilibrium

(b)
Unstable
equilibrium

(c)
Neutral
equilibrium
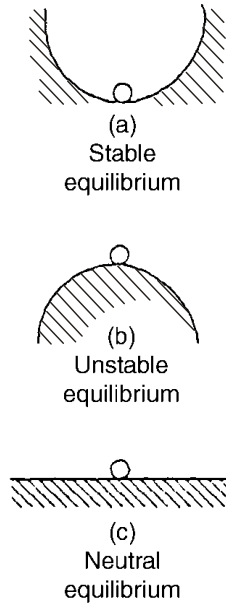
FIGURE 5.14

4. Center of Gravity

Definitions. The attraction of the earth for each particle of a body constitutes a system of forces with fixed application points. No matter how a body is turned, the resultant pull of the earth upon it always acts through a certain fixed point, called its *center of gravity.* The centroid of a geometric solid coincides with the center of gravity of a homogeneous body which would occupy the same volume. The centroid of a surface is the limiting position of the center of gravity of a homogeneous thin plate, one face of which coincides with the surface as the thickness of the plate approaches zero. The centroid of a line is the limiting position of the center of gravity of a homogeneous thin rod, the axis of which coincides with the line as the thickness of the rod approaches zero.

Planes of Symmetry and Axes of Symmetry. If a solid, surface, or line has a plane of symmetry, this plane contains the centroid. The line of intersection of two planes of symmetry contains the centroid. The point of intersection of three planes of symmetry locates the centroid.

Centroids by Integration. For a line of length *l*,

$$x_c = \frac{1}{l} \int x \, dl \qquad (5.13)$$

For a surface of area *A*,

$$x_c = \frac{1}{A} \int x \, dA \qquad (5.14)$$

For a solid of volume *V*,

$$x_c = \frac{1}{V} \int x \, dV \tag{5.15}$$

Similar expressions hold true for y_c and z_c. Centroids of some simple figures are given in Tables 5.1 and 5.2. Note that

- If a line, surface, or solid is composed of several simple parts whose centroids are known, the moment of the entire line, surface, or solid is equal to the sum of the moments of its several simple parts.
- If a mass is made up of several component parts of different material, and if the centers of gravity of these component masses are known, the moment equation for the component weights will locate the center of gravity of the entire mass.

*Note**: Based on previous sections it is clear that an object is in equilibrium when the forces acting on the object are such that there is no tendency for the object to move. The state of equilibrium of an object can be divided into three groups.

- If an object is in stable equilibrium and it is slightly disturbed by pushing or pulling (i.e. a disturbing force is applied), the center of gravity is raised, and when the disturbing force is removed, the object returns to its original position. Thus a ball-bearing in a hemispherical cup is in stable equilibrium, as shown in Fig. 5.14a.
- An object is in unstable equilibrium if, when a disturbing force is applied, the center of gravity is lowered and the object moves away from its original position. Thus, a ball-bearing balanced on top of a hemispherical cup is in unstable equilibrium, as shown in Fig. 5.14b.
- When an object in neutral equilibrium has a disturbing force applied, the center of gravity remains at the same height and the object does not move when the disturbing force is removed. Thus, a ball-bearing on a flat horizontal surface is in neutral equilibrium, as shown in Fig. 5.14c.

5. Moment of Inertia*

Moment of Inertia and Radius of Gyration of an Area.   The moment of inertia of an area with respect to any axis is the sum of the products of the differential areas and the squares of their distances from the given axis. Moment of inertia is denoted by I, usually with a subscript to indicate the axis. Thus

$$I_x = \int y^2 \, dA \quad \text{and} \quad I_y = \int x^2 \, dA \tag{5.16}$$

are the rectangular moments of inertia with respect to the X and Y axes in the plane of the area. Also,

---

TABLE 5.1 Properites of Plane Sections
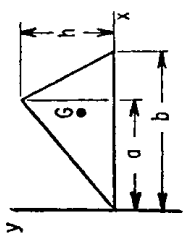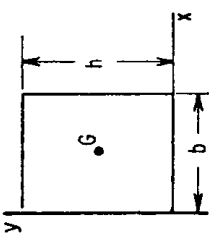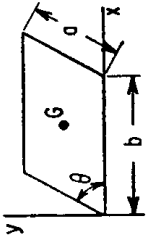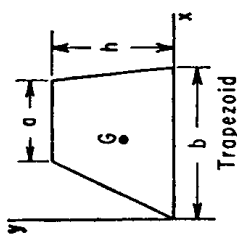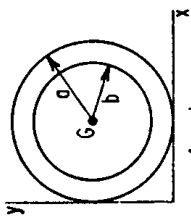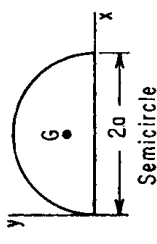
Figure	Area and centroid	Moment of inertia	r^2
Triangle	$A = \frac{1}{2}bh$ $x_c = \frac{1}{3}(a + b)$ $y_c = \frac{1}{3}h$	$I_{x_c} = \dfrac{bh^3}{36}$ $I_{y_c} = \dfrac{bh}{36}(b^2 - ab + a^2)$ $I_x = \dfrac{bh^3}{12}$ $I_y = \dfrac{bh}{12}(b^2 + ab + a^2)$	$r_{x_c}^2 = \frac{1}{18}h^2$ $r_{y_c}^2 = \frac{1}{18}(b^2 - ab + a^2)$ $r_x^2 = \frac{1}{6}h^2$ $r_y^2 = \frac{1}{6}(b^2 + ab + a^2)$
Rectangle	$A = bh$ $x_c = \frac{1}{2}b$ $y_c = \frac{1}{2}h$	$I_{x_c} = \dfrac{bh^3}{12}$ $I_{y_c} = \dfrac{b^3h}{12}$ $I_x = \dfrac{bh^3}{3}$ $I_y = \dfrac{b^3h}{3}$ $I_p = \dfrac{bh}{12}(b^2 + h^2)$	$r_{x_c}^2 = \frac{1}{12}h^2$ $r_{y_c}^2 = \frac{1}{12}b^2$ $r_x^2 = \frac{1}{3}h^2$ $r_y^2 = \frac{1}{3}b^2$ $r_p^2 = \frac{1}{12}(b^2 + h^2)$
Parallelogram	$A = ab\sin\theta$ $x_c = \frac{1}{2}(b + a\cos\theta)$ $y_c = \frac{1}{2}(a\sin\theta)$	$I_{x_c} = \dfrac{a^3b}{12}\sin^3\theta$ $I_{y_c} = \dfrac{ab}{12}\sin\theta(b^2 + a^2\cos^2\theta)$ $I_x = \dfrac{a^3b}{3}\sin^3\theta$ $I_y = \dfrac{ab}{3}\sin\theta\,(b + a\cos\theta)^2$ $\quad - \dfrac{a^2b^2}{6}\sin\theta\cos\theta$	$r_{x_c}^2 = \frac{1}{12}(a\sin\theta)^2$ $r_{y_c}^2 = \frac{1}{12}(b^2 + a^2\cos^2\theta)$ $r_x^2 = \frac{1}{3}(a\sin\theta)^2$ $r_y^2 = \frac{1}{3}(b + a\cos\theta)^2$ $\quad - \frac{1}{6}(ab\cos\theta)$

TABLE 5.1 Properites of Plane Sections (*Continued*)

Figure	Area and centroid	Moment of inertia	r^2

Trapezoid

$$A = \tfrac{1}{2}h\,(a + b)$$

$$y_c = \tfrac{1}{3}h\,\frac{2a + b}{a + b}$$

$$I_{x_c} = \frac{h^3(a^2 + 4ab + b^2)}{36(a + b)}$$

$$I_x = \frac{h^3(3a + b)}{12}$$

$$r_{x_c}^2 = \frac{h^2(a^2 + 4ab + b^2)}{18(a + b)^2}$$

$$r_x^2 = \frac{h^2(3a + b)}{6(a + b)}$$

Annulus

$$A = \pi(a^2 - b^2)$$

$$x_c = a$$

$$y_c = a$$

$$I_{x_c} = I_{y_c} = \frac{\pi}{4}(a^4 - b^4)$$

$$I_x = I_y = \frac{5}{4}\pi a^4 - \pi a^2 b^2 - \frac{\pi}{4}b^4$$

$$I_p = \frac{\pi}{2}(a^4 - b^4)$$

$$r_{x_c}^2 = r_{y_c}^2 = \tfrac{1}{4}(a^2 + b^2)$$

$$r_x^2 = r_y^2 = \tfrac{1}{4}(5a^2 + b^2)$$

$$r_P^2 = \tfrac{1}{2}(a^2 + b^2)$$

Semicircle

$$A = \tfrac{1}{2}\pi a^2$$

$$x_c = a$$

$$y_c = \frac{4a}{3\pi}$$

$$I_{x_c} = \frac{a^4(9\pi^2 - 64)}{72\pi}$$

$$I_{y_c} = \tfrac{1}{8}\pi a^4$$

$$I_x = \tfrac{1}{8}\pi a^4$$

$$I_y = \tfrac{5}{8}\pi a^4$$

$$r_{x_c}^2 = \frac{a^2(9\pi^2 - 64)}{36\pi^2}$$

$$r_{y_c}^2 = \tfrac{1}{4}a^2$$

$$r_x^2 = \tfrac{1}{4}a^2$$

$$r_y^2 = \tfrac{5}{4}a^2$$
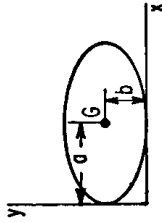
5.14

Circular sector

$$A = a^2\theta$$

$$x_c = \frac{2a}{3}\frac{\sin\theta}{\theta}$$

$$y_c = 0$$

$$I = \tfrac{1}{4}a^4(\theta - \sin\theta\cos\theta)$$

$$I_y = \tfrac{1}{4}a^4(\theta + \sin\theta\cos\theta)$$

$$r_x^2 = \tfrac{1}{4}a^2\,\frac{\theta - \sin\theta\cos\theta}{\theta}$$

$$r_y^2 = \tfrac{1}{4}a^2\,\frac{\theta + \sin\theta\cos\theta}{\theta}$$

Circular segment

$$A = a^2(\theta - \tfrac{1}{2}\sin 2\theta)$$

$$x_c = \frac{2a}{3}\frac{\sin^3\theta}{\theta - \sin\theta\cos\theta}$$

$$y_c = 0$$

$$I_x = \frac{Aa^2}{4}\left[1 - \frac{2\sin^3\theta\cos\theta}{3(\theta - \sin\theta\cos\theta)}\right]$$

$$I_y = \frac{Aa^2}{4}\left(1 + \frac{2\sin^3\theta\cos\theta}{\theta - \sin\theta\cos\theta}\right)$$

$$r_x^2 = \frac{a^2}{4}\left[1 - \frac{2\sin^3\theta\cos\theta}{3(\theta - \sin\theta\cos\theta)}\right]$$

$$r_y^2 = \frac{a^2}{4}\left(1 + \frac{2\sin^3\theta\cos\theta}{\theta - \sin\theta\cos\theta}\right)$$

Ellipse

$$A = \pi ab$$

$$x_c = a$$

$$y_c = b$$

$$I_{xc} = \frac{\pi}{4}ab^3$$

$$I_{yc} = \frac{\pi}{4}a^3b$$

$$I_x = \tfrac{5}{4}\pi ab^3$$

$$I_y = \tfrac{5}{4}\pi a^3b$$

$$I_P = \frac{\pi ab}{4}(a^2 + b^2)$$

$$r_{xc}^2 = \tfrac{1}{4}b^2$$
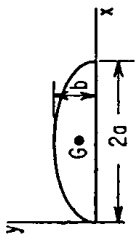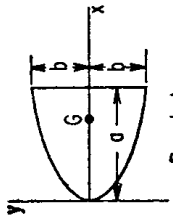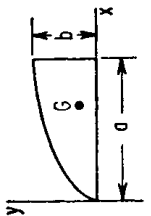
$$r_{yc}^2 = \tfrac{1}{4}a^2$$

$$r_x^2 = \tfrac{5}{4}b^2$$

$$r_y^2 = \tfrac{5}{4}a^2$$

$$r_P^2 = \tfrac{1}{4}(a^2 + b^2)$$

5.15

TABLE 5.1 Properites of Plane Sections (*Continued*)

Figure	Area and centroid	Moment of inertia	r^2
Semiellipse	$A = \tfrac{1}{2}\pi ab$ $x_c = a$ $y_c = \dfrac{4b}{3\pi}$	$I_{x_c} = \dfrac{ab^3}{72\pi}(9\pi^2 - 64)$ $I_{y_c} = \dfrac{\pi}{8}a^3 b$ $I_x = \dfrac{\pi}{8}ab^3$ $I_y = \tfrac{5}{8}\pi a^3 b$	$r_{x_c}^2 = \dfrac{b^2}{36\pi^2}(9\pi^2 - 64)$ $r_{y_c}^2 = \tfrac{1}{4}a^2$ $r_x^2 = \tfrac{1}{4}b^2$ $r_y^2 = \tfrac{5}{4}a^2$
Parabola	$A = \tfrac{4}{3}ab$ $x_c = \tfrac{3}{5}a$ $y_c = 0$	$I_{x_c} = I_x = \tfrac{4}{15}ab^3$ $I_{y_c} = \tfrac{16}{175}a^3 b$ $I_y = \tfrac{4}{7}a^3 b$	$r_{x_c}^2 = r_x^2 = \tfrac{1}{5}b^2$ $r_{y_c}^2 = \tfrac{12}{175}a^2$ $r_y^2 = \tfrac{3}{7}a^2$
Semiparabola	$A = \tfrac{2}{3}ab$ $x_c = \tfrac{3}{5}a$ $y_c = \tfrac{3}{8}b$	$I_x = \tfrac{2}{15}ab^3$ $I_y = \tfrac{2}{7}a^3 b$	$r_x^2 = \dfrac{1}{5}b^2$ $r_y^2 = \tfrac{3}{7}a^2$

Note: A = area
x_c, y_c = coordinates of centroid of section in XY coordinate system
r_x, r_y = radius of gyration of the section with respect to the centroidal parallel to the X and Y axes
r_{x_c}, r_{y_c} = radius of gyration of the section with respect to the centroidal axes parallel to the X and Y axes
r_p = radius of gyration of the section about the polar axis passing through the centroid
I_x, I_y = moment of inertia with respect to the X and Y axes shown
I_p = polar moment of inertia about an axis passing through the centroid
I_{x_c}, I_{y_c} = moment of inertia about an axis through the centroid parallel to the X and Y axes
G = marks the centroid
Source: From Rothbart.[4]

5.16

TABLE 5.2 Properties of Homogeneous Bodies
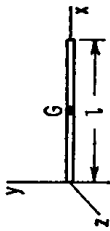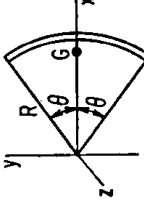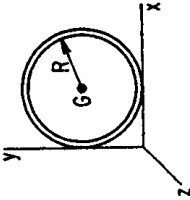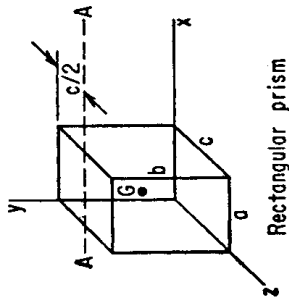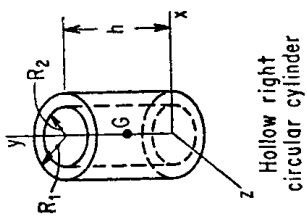
Body	Mass and centroid	Moment of inertia	r^2
 Thin rod	$m = \rho l$ $x_c = \frac{1}{2}l$ $y_c = 0$ $x_c = 0$	$I_x = I_{x_c} = 0$ $I_{y_c} = I_{z_c} = \dfrac{m}{12}\,l^2$ $I_y = I_z = \dfrac{m}{3}\,l^2$	$r_x^2 = r_{x_c}^2 = 0$ $r_{y_c}^2 = r_y^2 = \tfrac{1}{12}\,L^2$ $r_y^2 = r_z^2 = \tfrac{1}{3}l^2$
 Thin circular rod	$m = 2\rho R\theta$ $x_c = \dfrac{R\sin\theta}{\theta}$ $y_c = 0$ $z_c = 0$	$I_x = I_{x_c}$ $= \dfrac{mR^2(\theta - \sin\theta\cos\theta)}{2\theta}$ $I_y = \dfrac{mR^2(\theta + \sin\theta\cos\theta)}{2\theta}$ $I_z = mR^2$	$r_x^2 = r_{x_c}^2 = \dfrac{R^2(\theta - \sin\theta\cos\theta)}{2\theta}$ $r_y^2 = \dfrac{R^2(\theta + \sin\theta\cos\theta)}{2\theta}$ $r_z^2 = R^2$
 Thin hoop	$m = 2\pi\rho R$ $x_c = R$ $y_c = R$ $z_c = 0$	$I_{x_c} = I_{y_c} = \dfrac{m}{2}\,R^2$ $I_{z_c} = mR^2$ $I_x = I_y = \tfrac{3}{2}mR^2$ $I_z = 3mR^2$	$r_{x_c}^2 = r_{y_c}^2 = \tfrac{1}{2}R^2$ $r_{z_c}^2 = R^2$ $r_x^2 = r_y^2 = \tfrac{3}{2}R^2$ $r_z = 3R^2$

TABLE 5.2 Properties of Homogeneous Bodies (*Continued*)

Body	Mass and centroid	Moment of inertia	r^2
 Rectangular prism	$m = \rho abc$ $x_c = \frac{1}{2}a$ $y_c = \frac{1}{2}b$ $z_c = \frac{1}{2}c$	$I_{x_c} = \frac{1}{12}m(b^2 + c^2)$ $I_x = \frac{1}{3}m(b^2 + c^2)$ $I_{AA} = \frac{1}{12}m(4b^2 + c^2)$	$r^2_{x_c} = \frac{1}{12}(b^2 + c^2)$ $r^2_x = \frac{1}{3}(b^2 + c^2)$ $r^2_{AA} = \frac{1}{12}(4b^2 + c^2)$
 Hollow right circular cylinder	$m = \pi \rho h(R_1^2 - R_2^2)$ $x_c = 0$ $y_c = \frac{1}{2}h$ $z_c = 0$	$I_{x_c} = I_{z_c} = \frac{1}{12}m(3R_1^2 + 3R_2^2 + h^2)$ $I_{y_c} = I_y = \frac{1}{2}m(R_1^2 + R_2^2)$ $I_x = I_z = \frac{1}{12}m(3R_1^2 + 3R_2^2 + 4h^2)$	$r^2_{x_c} = r^2_{z_c} = \frac{1}{12}(3R_1^2 + 3R_2^2 + h^2)$ $r^2_{y_c} = r^2_y = \frac{1}{2}(R_1^2 + R_2^2)$ $r^2_x = r^2_z = \frac{1}{12}(3R_1^2 + 3R_2^2 + 4h^2)$
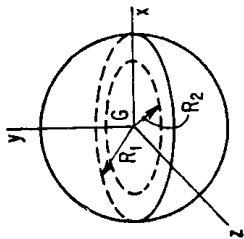
Hollow sphere
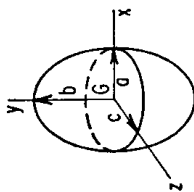
$$m = \frac{4}{3}\pi\rho(R_1^3 - R_2^3)$$
$$x_c = 0$$
$$y_c = 0$$
$$z_c = 0$$

$$I_x = I_y = I_z = \frac{2}{5}m\,\frac{R_1^5 - R_2^5}{R_1^3 - R_2^3}$$

$$r_x^2 = R_y^2 = r_z^2 = \frac{2}{5}\frac{R_1^5 - R_2^5}{R_1^3 - R_2^3}$$

Ellipsoid

$$m = \frac{4}{3}\pi\rho abc$$
$$x_c = 0$$
$$y_c = 0$$
$$z_c = 0$$

$$E_x = \frac{1}{5}m(b^2 + c^2)$$
$$I_y = \frac{1}{5}m(a^2 + c^2)$$
$$I_z = \frac{1}{5}m(a^2 + b^2)$$

$$r_x^2 = \frac{1}{5}(b^2 + c^2)$$
$$r_y^2 = \frac{1}{5}(a^2 + c^2)$$
$$r_z^2 = \frac{1}{5}(a^2 + b^2)$$

Note: ρ = mass density
m = mass
x_c, y_c, z_c = coordinates of centroid in *XYZ* coordinate system
r_{xc}, r_{yc}, r_{zc} = radius of gyration of the body with respect to the centroidal axes parallel to the *X, Y,* and *Z* axes shown
I_x, I_y, I_z = moment of inertia with respect to the *X, Y,* and *Z* axes shown
I_{xc}, I_{yc}, I_{zc} = moment of inertia about an axis through the centroid parallel to the *X, Y,* and *Z* axes shown
I_{AA} = moment of inertia with respect to special axes shown
G = marks the centroid
Source: From Rothbart.[4]

5.19

$$I_p = \int \rho^2 \, dA \tag{5.17}$$

is the polar moment of inertia with respect to the Z axis normal to the area.

In terms of the total area, $I = A\mathbf{r}^2$, $\mathbf{r}^2$ being the mean value of x^2 or y^2. Since $I = A\mathbf{r}^2$,

$$\mathbf{r} = \sqrt{\frac{I}{A}} \tag{5.18}$$

The quantity $\mathbf{r}$ is called the *radius of gyration.*

If X is any axis in the plane of an area A, X_0 the parallel centroidal axis, and d the distance between the axes, the relationship between the two moments of inertia is given by the transfer formula

$$I_x = I_{x_0} + Ad^2 \tag{5.19}$$

If Z is any polar axis of an area A, Z_0 the centroidal axis, and d the distance between the two axes,

$$I_p = I_{po} + Ad^2 \tag{5.20}$$

If X and Y are two rectangular axes in the plane of the area and Z is the polar axis through their point of intersection,

$$I_p = I_x + I_y \tag{5.21}$$

If an area is composed of several simple parts, the moment of inertia of each part with respect to any given axis may be computed separately and added to obtain the moment of inertia of the entire area.

Moment of Inertia of Mass. In problems involving the rotation of a mass about an axis, it is necessary to compute the quantity $\int \rho^2 \, dm$. In this expression, dm is the mass of any differential particle and ρ is its distance from the axis of rotation. The quantity $\int \rho^2 \, dm$ is called the *moment of inertia of the mass* and is represented by I, usually with a subscript to indicate the axis. The radius of gyration $\mathbf{r}$ is the distance from the axis at which all of the mass could be concentrated and has the same moment of inertia; or, in other words, r^2 is the mean value of the variable ρ^2. Thus

$$I = \int \rho^2 \, dm = mr^2 \tag{5.22}$$

The relation between the moment of inertia with respect to the centroidal axis O and that with respect to any parallel axis C is given by the transfer formula

$$I_C = I_0 + md^2 \tag{5.23}$$

d being the perpendicular distance between the axes.

Values of the moment of inertia for different plane sections and homogeneous bodies are given in Tables 5.1 and 5.2.

FRICTION

1. Definitions of Sliding and Static Friction.   Friction is the resistance that is encountered when two solid surfaces slide or tend to slide over each other.

The direction of the friction force is always opposite to the direction of motion. Thus, the friction force opposes motion, as shown in Fig. 5.15a.

In stationary systems, friction manifests itself as a force equal and opposite to the shear force applied to the interface. Thus, as shown in Fig. 5.15b, if a force F_0 is applied, a friction force F will be generated, equal and opposite to F_0, so that the surfaces remain at rest. F can take any magnitude up to a limiting value F', and can therefore prevent sliding whenever F_0 is less than F'. If F_0 exceeds F', slipping of the body occurs.

The friction force is proportional to the normal force F_N (note that $F_N = -W$, where W is equal to the weight of the body), and the coefficient of proportionality is defined as the *static* friction coefficient f_0. This is expressed by the equation

$$f_0 = \frac{F'}{F_N} \tag{5.24}$$

The angle of static friction α (the largest angle relative to the horizontal at which a surface may be tilted, so that the body placed on the surface does not slide down) is defined by

$$\tan \alpha = \frac{F'}{F_N} = f_0 \tag{5.25}$$

If one surface slides over the other, a frictional force resulting from the motion
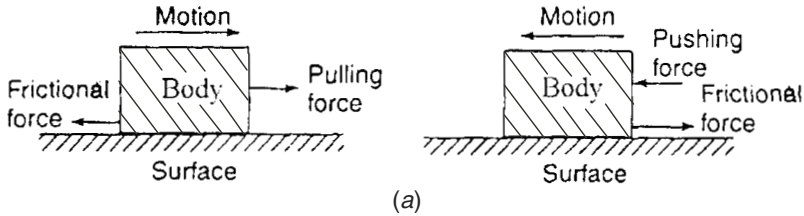
FIGURE 5.15a
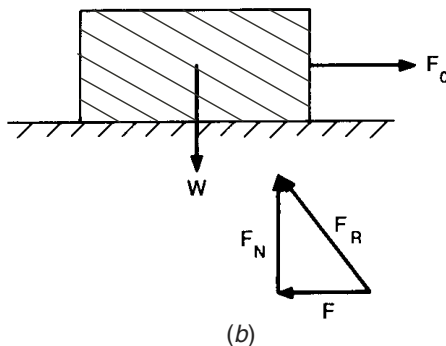
FIGURE 5.15b

must be overcome. This force is usually smaller, by about 20 percent, than F' and is expressed as

$$F = f \cdot F_N \tag{5.26}$$

where f is the *sliding* friction coefficient or kinetic friction. Angle of kinetic friction is defined as

$$\tan \alpha = \frac{F}{F_N} \tag{5.27}$$

In the range of practical velocities of sliding, the coefficients of sliding friction are smaller than the coefficients of static friction. With small velocities of sliding and very clean surfaces, the two coefficients do not differ appreciably.

When the surfaces are free from contaminating fluids, or films, the resistance is called *dry friction*. When the surfaces are separated from each other by a very thin film of lubricant, the friction is that of *boundary* (or *greasy*) *lubrication*.

Values of sliding and static coefficients are given in Table 5.3. Additional data for sliding coefficients are given in Table 5.4.

2. Rolling Friction. If a sphere touches a plane or a cylindrical surface, the contact is only a point, while if a cylinder touches a plane or another cylindrical surface, the contact is a line. Assuming that all materials are elastic, the sphere or cylinder under pressure and its supporting surface are distorted so that there is an area of contact. As the sphere or cylinder is rolled along under pressure, it is assumed that the material of the supporting surface is rolled ahead as a miniature wave and the center of contact is a small distance a in front of the normal radius, as shown in Fig. 5.16. With the center of contact as the center of moments,

$$F \frac{r}{2} = Wa \tag{5.28}$$

where F is the horizontal force and W is the normal force (in this case the weight of the body). With hard rollers and hard bearings, the value of a is very small and the resistance to rolling is very much smaller than the resistance to sliding. Values of a as recommended in the literature are given in Table 5.5.

3. Friction of Flexible Belts and Bands. See Fig. 5.17. Let T_1 be the tension in the taut side of a belt over a pulley, T_2 the tension in the slack side, f the coefficient of friction, and α the angle of contact in radians. Then the expression giving the limiting relation between T_1 and T_2 when slipping impends is

$$T_1 = T_2 e^{\alpha f} \tag{5.29}$$

The quantity e is the base of the natural system of logarithms and its value is 2.71828. The same relation holds true for a flexible band brake and for a rope over a spar or around a snubbing post. When slipping occurs, the quantity f is the kinetic coefficient of friction.

Note that $T_1 - T_2$ is equal to the circumferential force transfered by friction. Values of T_1/T_2 for various values of f and α are shown in Table 5.6.

Note. Values of friction for different machine elements are given in Ref. 2.

TABLE 5.3 Coefficients of Static and Sliding Friction

(*Reference letters indicate the lubricant used; see footnote*)

Materials	Static		Sliding	
	Dry	Greasy	Dry	Greasy
Hard steel on hard steel	0.78	0.11 (*a*)	0.42	0.029 (*h*)
	....	0.23 (*b*)	....	0.081 (*c*)
	....	0.15 (*c*)	....	0.080 (*i*)
	....	0.11 (*d*)	....	0.058 (*j*)
	....	0.0075 (*p*)	....	0.084 (*d*)
	....	0.0052 (*h*)	....	0.105 (*k*)
	....	....	....	0.096 (*l*)
	....	....	....	0.108 (*m*)
	....	....	....	0.12 (*a*)
Mild steel on mild steel	0.74	....	0.57	0.09 (*a*)
	....	....	....	0.19 (*u*)
Hard steel on graphite	0.21	0.09 (*a*)		
Hard steel on babbitt (ASTM No. 1) ...	0.70	0.23 (*b*)	0.33	0.16 (*b*)
	....	0.15 (*c*)	....	0.06 (*c*)
	....	0.08 (*d*)	....	0.11 (*d*)
	....	0.085 (*e*)		
Hard steel on babbitt (ASTM No. 8) ...	0.42	0.17 (*b*)	0.35	0.14 (*b*)
	....	0.11 (*c*)	....	0.065 (*c*)
	....	0.09 (*d*)	....	0.07 (*d*)
	....	0.08 (*e*)	....	0.08 (*h*)
Hard steel on babbitt (ASTM No. 10) ..	....	0.25 (*b*)	....	0.13 (*b*)
	....	0.12 (*c*)	....	0.06 (*c*)
	....	0.10 (*d*)	....	0.055 (*d*)
	....	0.11 (*e*)		
Mild steel on cadmium silver	....	....	....	0.097 (*f*)
Mild steel on phosphor bronze	....	....	0.34	0.173 (*f*)
Mild steel on copper lead	....	....	....	0.145 (*f*)
Mild steel on cast iron	....	0.183 (*c*)	0.23	0.133 (*f*)
Mild steel on lead	0.95	0.5 (*f*)	0.95	0.3 (*f*)
Nickel on mild steel	....	....	0.64	0.178 (*x*)
Aluminum on mild steel	0.61	....	0.47	
Magnesium on mild steel	....	....	0.42	
Magnesium on magnesium	0.6	0.08 (*y*)		
Teflon on Teflon	0.04	....	....	0.04 (*f*)
Teflon on steel	0.04	....	....	0.04 (*f*)
Tungsten carbide on tungsten carbide ..	0.2	0.12 (*a*)		
Tungsten carbide on steel	0.5	0.08 (*a*)		
Tungsten carbide on copper	0.35			
Tungsten carbide on iron	0.8			
Bonded carbide on copper	0.35			
Bonded carbide on iron	0.8			
Cadmium on mild steel	....	....	0.46	
Copper on mild steel	0.53	....	0.36	0.18 (*a*)
Nickel on nickel	1.10	....	0.53	0.12 (*w*)
Brass on mild steel	0.51	....	0.44	
Brass on cast iron	....	....	0.30	
Zinc on cast iron	0.85	....	0.21	
Magnesium on cast iron	....	....	0.25	

TABLE 5.3 Coefficients of Static and Sliding Friction (*Continued*)

Materials	Static		Sliding	
	Dry	Greasy	Dry	Greasy
Copper on cast iron	1.05	....	0.29	
Tin on cast iron	....	....	0.32	
Lead on cast iron	....	....	0.43	
Aluminum on aluminum	1.05	....	1.4	
Glass on glass	0.94	0.01 (*p*)	0.40	0.09 (*a*)
	....	0.005 (*q*)	....	0.116 (*v*)
Carbon on glass	....	....	0.18	
Garnet on mild steel	....	....	0.39	
Glass on nickel	0.78	....	0.56	
Copper on glass	0.68	....	0.53	
Cast iron on cast iron	1.10	....	0.15	0.070 (*d*)
	....	....	....	0.064 (*n*)
Bronze on cast iron	....	....	0.22	0.077 (*n*)
Oak on oak (parallel to grain)	0.62	....	0.48	0.164 (*r*)
	....	....	....	0.067 (*s*)
Oak on oak (prependicular)	0.54	....	0.32	0.072 (*s*)
Leather on oak (parallel)	0.61	....	0.52	
Cast iron on oak	....	....	0.49	0.075 (*n*)
Leather on cast iron	....	....	0.56	0.36 (*t*)
	....	....	....	0.13 (*n*)
Laminated plastic on steel	....	....	0.35	0.05 (*t*)
Fluted rubber bearing on steel	....	....	....	0.05 (*t*)

(*a*) Oleic acid; (*b*) Atlantic spindle oil (light mineral); (*c*) castor oil; (*d*) lard oil; (*e*) Atlantic spindle oil plus 2 percent oleic acid; (*f*) medium mineral oil; (*g*) medium mineral oil plus ½ percent oleic acid; (*h*) stearic acid; (*i*) grease (zinc oxide base); (*j*) graphite; (*k*) turbine oil plus 1 percent graphite; (*l*) turbine oil plus 1 percent stearic acid; (*m*) turbine oil (medium mineral); (*n*) olive oil; (*p*) palmitic acid; (*q*) ricinoleic acid; (*r*) dry soap; (*s*) lard; (*t*) water; (*u*) rape oil; (*v*) 3-in-1 oil; (*w*) octyl alcohol; (*x*) triolein; (*y*) 1 percent lauric acid in paraffin oil.

Source: From Avallone and Baumeister.[2]

TABLE 5.4 Values of *f* for Mild Steel on Medium Steel—Effect of Sliding Velocity

in/s	0.0001	0.001	0.01	0.1	1	10	100
cm/s	0.000254	0.00254	0.0254	0.254	2.54	25.4	254
f	0.53	0.48	0.39	0.31	0.23	0.19	0.18

Source: Avallone and Baumeister.[2]

KINEMATICS

1. Definitions

Kinematics.   The study of the motion of bodies without reference to the forces causing that motion or the mass of the bodies.

Motion.   Motion of a particle with respect to other particles or objects is its state of continual changing of position with respect to them.
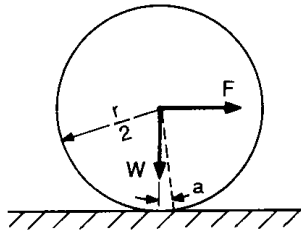
FIGURE 5.16

TABLE 5.5 Rolling Resistance

Materials	a, inches (cm)
Hardwood on hardwood	0.02 (0.05)
Iron on iron (steel on steel)	0.002 (0.005)
Hard polish steel on hard polish steel	0.0002 to 0.0004 (0.0005 to 0.001)
Pneumatic tires on good road of asphalt	0.02 to 0.022 (0.0508 to 0.0558)
Pneumatic tires on heavy mud of asphalt	0.04 to 0.06 (0.1016 to 0.1524)
Iron or steel wheels on wood track	0.06 to 0.10 (0.1524 to 0.254)

Source: Updated from O'Rourke.[3]
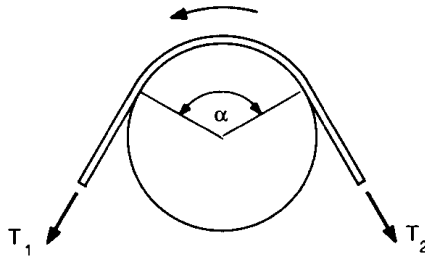
FIGURE 5.17

Rectilinear Motion.   Motion along a straight path.

Curvilinear Motion.   Motion along a curved path which may be either planar or skewed.

Displacement.   Displacement of a particle is its change of position and is a vector quantity. If point *A* is the position of a particle at a time *t,* and *B* its position at a later time t_2, its displacement in the time interval $t_2 - t_1 = \Delta t$ is the vector **AB**, no matter whether the path is straight or curved.

Velocity.   Velocity of a particle is its time rate of displacement and is a vector quantity.

Speed.   The magnitude of velocity without reference to direction.

Acceleration.   Acceleration of a particle is its time rate of change of velocity and is a vector quantity.

Relative Motion.   The displacement, velocity, and acceleration of a body with respect to a fixed point on the earth are called its *absolute* displacement, velocity, and acceleration, respectively. The displacement, velocity, and acceleration of one

TABLE 5.6 Values of T_1/T_2

$\dfrac{a^\circ}{360^\circ}$	f								
	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
0.1	1.06	1.1	1.13	1.17	1.21	1.25	1.29	1.33	1.37
0.2	1.13	1.21	1.29	1.37	1.46	1.55	1.65	1.76	1.87
0.3	1.21	1.32	1.45	1.60	1.76	1.93	2.13	2.34	2.57
0.4	1.29	1.46	1.65	1.87	2.12	2.41	2.73	3.10	3.51
0.425	1.31	1.49	1.70	1.95	2.23	2.55	2.91	3.33	3.80
0.45	1.33	1.53	1.76	2.03	2.34	2.69	3.10	3.57	4.11
0.475	1.35	1.56	1.82	2.11	2.45	2.84	3.30	3.83	4.45
0.5	1.37	1.60	1.87	2.19	2.57	3.00	3.51	4.11	4.81
0.525	1.39	1.64	1.93	2.28	2.69	3.17	3.74	4.41	5.20
0.55	1.41	1.68	2.00	2.37	2.82	3.35	3.98	4.74	5.63
0.6	1.46	1.76	2.13	2.57	3.10	3.74	4.52	5.45	6.59
0.7	1.55	1.93	2.41	3.00	3.74	4.66	5.81	7.24	9.02
0.8	1.65	2.13	2.73	3.51	4.52	5.81	7.47	9.60	12.35
0.9	1.76	2.34	3.10	4.11	5.45	7.24	9.60	12.74	16.90
1.0	1.87	2.57	3.51	4.81	6.59	9.02	12.35	16.90	23.14
1.5	2.57	4.11	6.59	10.55	16.90	27.08	43.38	69.49	111.32
2.0	3.51	6.59	12.35	23.14	43.38	81.31	152.40	285.68	535.49
2.5	4.81	10.55	23.14	50.75	111.32	244.15	535.49	1,174.5	2,575.9
3.0	6.59	16.90	43.38	111.32	285.68	733.14	1,881.5	4,828.5	12,391
3.5	9.02	27.08	81.31	244.15	733.14	2,199.90	6,610.7	19,851	59,608
4.0	12.35	43.38	152.40	535.49	1,881.5	6,610.7	23,227	81,610	286,744

Source: From Avallone and Baumeister.[2]

body with respect to another which may also have displacement, velocity, and acceleration with respect to the earth are called its *relative* displacement, velocity, and acceleration, respectively.

2. Rectilinear Motion

Velocity

$$\text{Instantaneous velocity} = v = \frac{ds}{dt} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} \tag{5.30}$$

$$\text{Average velocity} = \bar{v} = \frac{\Delta s}{\Delta t} \tag{5.31}$$

where s = distance measured along the path of a particle.

Acceleration

$$\text{Instantaneous acceleration} = a = \frac{dv}{dt} = \frac{d^2s}{dt^2} = \lim_{\Delta t \to 0} \frac{\Delta v}{\Delta t} \tag{5.32}$$

$$\text{Average acceleration} = \bar{a} = \frac{\Delta v}{\Delta t} \tag{5.33}$$

Relations among a, v, s, t

$$v = \frac{ds}{dt} \tag{5.34}$$

$$a = \frac{dv}{dt} = \frac{d^2s}{dt^2} \tag{5.35}$$

$$\frac{a}{v} = \frac{dv}{ds} \tag{5.36}$$

$$s_2 - s_1 = \int_{t_1}^{t_2} v \, dt \tag{5.37}$$

$$v_2 - v_1 = \int_{t_1}^{t_2} a \, dt \tag{5.38}$$

$$t_2 - t_1 = \int_{s_1}^{s_2} \frac{ds}{v} = \int_{v_1}^{v_2} \frac{dv}{a} \tag{5.39}$$

$$v_2^2 - v_1^2 = 2 \int_{s_1}^{s_2} a \, ds \tag{5.40}$$

where s_1 = distance from origin of time t_1
s_2 = distance at a later time t_2
v_1 = velocity of particle at time t_1
v_2 = velocity at later time t_2

Motion Graphs. These include

- A *distance-time curve,* which offers a convenient means for the study of motion of a point. The slope of the curve at any point will represent the velocity at that time.
- A *velocity-time curve,* which offers a convenient means for the study of acceleration. The slope of the curve at any point will represent the acceleration at that time.
- An *acceleration-time curve,* which may be constructed by plotting accelerations as ordinates, and times as abscissas. The area under this curve between any two ordinates will represent the total increase in velocity during the time interval.

Examples of Rectilinear Motion

Uniform Acceleration (a = constant). Let v_0 be initial velocity and s_0 initial distance. Then from Eqs. (5. 37) through (5.40):

$$v = at + v_0 \tag{5.41}$$

$$s = \tfrac{1}{2}at^2 + v_0 t \tag{5.42}$$

$$v^2 = 2a(s - s_0) + v_0^2 \tag{5.43}$$

Free Fall. If a body falls from rest in a vacuum, $v_0 = 0$, $s_0 = 0$, and $a = g$ (= acceleration due to gravity). Then:

$$v = gt = (2gs)^{1/2} \tag{5.44}$$

$$s = \tfrac{1}{2}gt^2 \tag{5.45}$$

If a body is projected upward at an initial velocity v_0, then

$$a = -g$$

$$v = -gt + v_0 = (-2gs + v_0^2)^{1/2} \tag{5.46}$$

$$s = -\tfrac{1}{2}gt^2 + v_0 t \tag{5.47}$$

H (total ascent to highest position) $= \dfrac{v_0^2}{2g}$ and t_H (time required) $= \dfrac{v_0}{g}$

Crank and Connecting-Rod Mechanism. The problem is to find expressions for the velocity and acceleration of any point in the crosshead *A,* as shown in Fig. 5.18. Such a point describes rectilinearmotion. Let $\lambda = (r/l) = \frac{1}{4} \cdots \frac{1}{6}$, $n =$ revolutions-per-second assumed constant, $\omega =$ radians of angle described by crank per second, and $s =$ distance of *A* from its extreme position *O.* Then

$$s = r(1 - \cos \varphi) + \frac{\lambda}{2} r \sin^2 \varphi \tag{5.48}$$

$$v = \omega r \sin \varphi (1 + \lambda \cos \varphi) \tag{5.49}$$

$$a = \omega^2 r (\cos \varphi + \lambda \cos 2\varphi) \tag{5.50}$$

$$\varphi = \omega t = 2\pi n t$$

Simple Harmonic Motion. This type of motion has wide application in physics and engineering (for example, a body supported by a spring performs a linear harmonic oscillation). For this kind of motion, quantities *s, v,* and *a* are defined as:

$$s = A \sin (\omega t + \varphi_0) \tag{5.51}$$

$$v = \frac{ds}{dt} = A\omega \cos (\omega t + \varphi_0) \tag{5.52}$$

$$a = \frac{dv}{dt} = -A\omega^2 \sin (\omega t + \varphi_0) \tag{5.53}$$
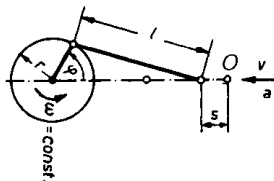
FIGURE 5.18 (*From Gieck.*[1])

where A = amplitude (maximum displacement)
   $\varphi = (\omega t + \varphi_0)$ = angular position at time t
   φ_0 = angular position at $t = 0$
   s = displacement
   ω = angular frequency, radians per unit time

Sliding Motion on an Inclined Plane. See Fig. 5.19. Here the quantities are defined as:

$$a = g \, (\sin \alpha - f \cos \alpha) \tag{5.54}$$

$$v = (2as)^{1/2} \tag{5.55}$$

$$s = \frac{v^2}{2a} \tag{5.56}$$

where f = coefficient of sliding friction.

Rolling Motion on an Inclined Plane. See Fig. 5.20. Here the quantities are defined as:

$$a = gr^2 \, \frac{\sin \alpha - (r_0/r) \cos \alpha}{r^2 + K^2} \tag{5.57}$$

v = value defined by Eq. (5.55)

s = value defined by Eq. (5.56)

r_0 = values given in Table 5.5, i.e., use $r_0 = a$

$$K^2 = \tfrac{2}{5}r^2 \qquad = \frac{r^2}{2} \qquad = r^2$$

for ball, solid cylinder, and pipe with low wall thickness, respectively.

3. Curvilinear motion

Velocity and Acceleration. The magnitude of velocity at any instant is

$$v = \frac{ds}{dt} \tag{5.58}$$

where s is measured along the curved path of a particle (Fig. 5.21a). The linear direction of the velocity is the tangent to the path of the point. In Fig. 5.21a, let ABC be the path of a moving point and $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ velocity vectors at A, B, and C.
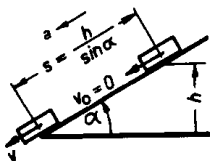
FIGURE 5.19 (*From Gieck.*[1])
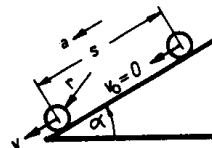
FIGURE 5.20 (*From Gieck.*[1])
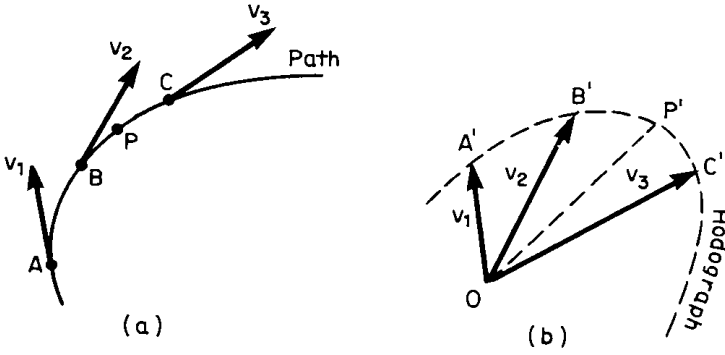
FIGURE 5.21

If O is taken as a pole (Fig. 5.21b) and vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ representing the velocities at corresponding points are drawn, the curve connecting the terminal points of these vectors is known as the *hodograph* of the motion.

The acceleration is given in terms of its tangential and normal components a_t and a_n, respectively. The tangential component a_t may be positive, zero, or negative, assuming the direction of the velocity is positive. Its amount is

$$a_t = \frac{dv}{dt} \tag{5.59}$$

the *normal* component is always directed toward the center of curvature along the radius ρ at the point. Its amount is

$$a_n = \frac{v^2}{\rho} \tag{5.60}$$

where v is the velocity at the point. If the radius of curvature ρ increases and becomes infinity, the path of the point becomes a straight line. If the radius of curvature ρ ceases to vary and becomes the constant radius r, the path of the point becomes a circle and the motion is rotation. The normal acceleration is

$$a_n = \frac{v^2}{r} \tag{5.61}$$

Resultant acceleration is

$$a = (a_t^2 + a_n^2)^{1/2} \tag{5.62}$$

Components of Velocity and Acceleration in Different Coordinate Systems. See Figs. 3.1 to 3.3 in Chap. 3.

Rectangular, $P(x,y,z)$

$$v = (v_x^2 + v_y^2 + v_z^2)^{1/2} \tag{5.63}$$

$$a = (a_x^2 + a_y^2 + a_z^2)^{1/2} \tag{5.64}$$

$$v_x = \dot{x} \qquad v_y = \dot{y} \qquad v_z = \dot{z} \tag{5.65}$$

$$a_x = \ddot{x} \qquad a_y = \ddot{y} \qquad a_z = \ddot{z} \tag{5.66}$$

Cylindrical, P(r, θ, z)

$$v = (v_r^2 + v_\theta^2 + v_z^2)^{1/2} \tag{5.67}$$

$$a = (a_r^2 + a_\theta^2 + a_z^2)^{1/2} \tag{5.68}$$

$$v_r = \dot{r} \qquad v_\theta = r\dot{\theta} \qquad v_z = \dot{z} \tag{5.69}$$

$$a_r = \ddot{r} - r\dot{\theta}^2 \qquad a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta} \qquad a_z = \ddot{z} \tag{5.70}$$

Spherical, P(r, θ, φ)

$$v = (v_r^2 + v_\theta^2 + v_\phi^2)^{1/2} \tag{5.71}$$

$$a = (a_r^2 + a_\theta^2 + a_\phi^2)^{1/2} \tag{5.72}$$

$$v_r = \dot{r} \qquad v_\phi = r\dot{\phi} \qquad v_\theta = r\dot{\theta} \sin \phi \tag{5.73}$$

$$a_r = \ddot{r} - r\dot{\phi}^2 - r\dot{\theta}^2 \sin^2 \phi \tag{5.74}$$

$$a_\phi = 2\dot{r}\dot{\phi} + r\ddot{\phi} - r\dot{\theta}^2 \sin \phi \cos \phi \tag{5.75}$$

$$a_0 = 2\dot{r}\dot{\theta} \sin \phi + r\ddot{\theta} \sin \phi + 2r\dot{\theta}\dot{\phi} \cos \phi \tag{5.76}$$

Note. The above relations show that velocities and acceleration (like forces) may be composed or resolved according to the parallelogram and parallelepipedon laws discussed above in "Graphical Composition of Forces."

Motion of a Projectile (Example). Assuming that the resistance of air is neglected, the vertical component of the motion of a projectile is the same as that of a falling body, and the horizontal component of the motion is that of a body with constant velocity. Let v_0 be the initial velocity, at an angle α with the horizontal (Fig. 5.22). In the vertical direction,

$$a_y = -g$$

$$v_y = v_0 \sin \alpha - g \cdot t$$

$$y = v_0 t \sin \alpha - \frac{gt^2}{2} \tag{5.77}$$

In the horizontal direction,

FIGURE 5.22

$$a_x = 0$$
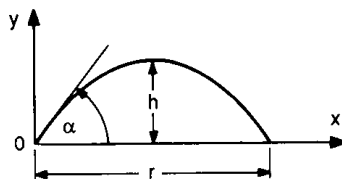
$$v_x = v_0 \cos \alpha$$

$$x = v_0 t \cos \alpha \tag{5.78}$$

Eliminating t from Eqs. (5.77) and (5.78),

$$y = x \tan \alpha - \frac{gx^2}{2v_0^2 \cos^2 \alpha} \tag{5.79}$$

Also,

$$h = \frac{(\sin^2 \alpha)v_0^2}{2g}$$

$$r = \frac{(\sin 2\alpha)v_0^2}{g}$$

$$T = \frac{2v_0 \sin \alpha}{g}$$

where $\alpha = \alpha_1 = 45° =$ value of α for maximum r
     $h =$ greatest height attained
     $T =$ time of flight

   Note.   Motion graphs can be constructed for curvilinear motion of a particle, as they can for rectilinear motion.

4. Rotation

Angular Displacement θ.   If a body rotates about a fixed axis, each particle of the body describes a circle (Fig. 5.23). The angle described by any radius (or by any straight line in a plane normal to the axis) is called the *angular displacement.*
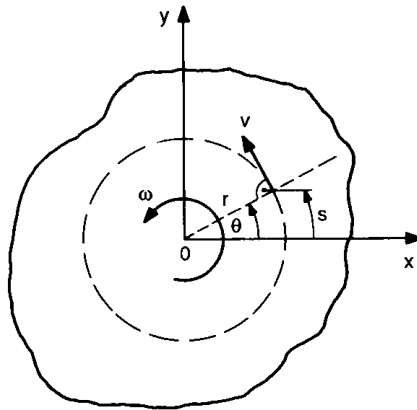
FIGURE 5.23

The unit of angular displacement is the radian, the angle at the center subtended by an arc s equal in length to the radius r.

If θ is the angle of displacement, the arc $s = r\theta$. One circumference or $360° = 2\pi$ rad, or 1 rad $= 180°/\pi$.

Angular Velocity. The angular velocity ω is the time rate of angular displacement:

$$\omega = \frac{d\theta}{dt} = \dot{\theta} \tag{5.80}$$

As $s = r\theta$, by differentiation with respect to t the following relation is obtained:

$$v = r\omega \tag{5.81}$$

Angular Acceleration. The angular acceleration α is the time rate of change of angular velocity:

$$\alpha = \frac{d\omega}{dt} = \dot{\omega} = \ddot{\theta} \tag{5.82}$$

As $v = r\omega$, by differentiation with respect to t the following relation is obtained:

$$a_t = r \cdot \alpha \tag{5.83}$$

As $a_n = v^2/r$ [Eq. (5.60)] and $v = r\omega$, then

$$a_n = r\omega^2 \tag{5.84}$$

If $\alpha = 0$, ω and a_n are constant. The angle moved through in time t is

$$\theta = \omega t \tag{5.85}$$

and the period of one complete rotation is

$$T = \frac{2\pi}{\omega} \tag{5.86}$$

The frequency f is the reciprocal of the period T:

$$f = \frac{1}{T}$$

In general, as with rectilinear motion, the following apply:

$$\frac{\alpha}{\omega} = \frac{d\omega}{d\theta} \tag{5.87}$$

$$\theta_2 - \theta_1 = \int_{t_1}^{t_2} \omega \, dt \tag{5.88}$$

$$\omega_2 - \omega_1 = \int_{t_1}^{t_2} \alpha \, dt \tag{5.89}$$

$$t_2 - t_1 = \int_{\theta_1}^{\theta_2} \frac{d\theta}{\omega} = \int_{\omega_1}^{\omega_2} \frac{d\omega}{\alpha} \tag{5.90}$$

$$\omega_2^2 - \omega_1^2 = 2 \int_{\theta_1}^{\theta_2} \alpha \, d\theta \tag{5.91}$$

If $\alpha = $ const., the following relations are obtained ($\omega_0 = \omega_1$, $\omega_2 = \omega$, $t_1 = 0$, $t_2 = t$):

$$\omega = \omega_0 + \alpha t \tag{5.92}$$

$$\theta = \omega_0 t + \frac{\alpha t^2}{2} \tag{5.93}$$

$$\omega^2 = \omega_0^2 + 2\alpha\theta \tag{5.94}$$

Note also that for rotation of a body about a fixed axis through the center of gravity, if a constant unbalanced moment M is applied,

$$M = I\alpha \tag{5.95}$$

where $I = $ moment of inertia (mass).

DYNAMICS

1. Definitions

Dynamics. That branch of mechanics which deals with the effects of unbalanced external forces in modifying the motion of bodies. Dynamics also deals with the terms "work," "energy," and "power."

Force (Gravitational). The force F is the product of mass m and acceleration a:

$$F = ma \tag{5.96}$$

The gravitational force W is the force acting on a mass m due to the earth's acceleration g:

$$W = mg \tag{5.97}$$

From Eqs. (5.96) and (5.97)

$$F = \frac{W}{g} a \qquad (5.98)$$

Being a gravitational force, the weight W is measured by means of a spring balance.

Conservation of Mass. The mass of a body remains unchanged by ordinary physical changes to which it may be subjected.

Conservation of Energy. The principle of the conservation of energy requires that the total mechanical energy of a system remain unchanged if it is subjected only to forces which depend on position or configuration.

Conservation of Momentum. The linear momentum of a system of bodies remains unchanged if there is no resultant external force on the system. The angular momentum of a system of bodies about a fixed axis is unchanged if there is no resultant external momentum about this axis.

Mutual Attraction (*Gravitation*). Two particles attract each other with a force F proportional to their masses m_1 and m_2 and inversely proportional to the square of the distance r between them, or

$$F = \frac{K m_1 m_2}{r^2} \qquad (5.99)$$

where K ($= 6.673 \times 10^{-11}$ m^3/kg · s^2 or 3.44×10^{-8} ft^3/lb · s^2) is the gravitational constant.

2. Particle Dynamics—Selected Problems*. In many engineering problems involving the motion of a particle or the motion of a mass center, the magnitude and direction of the force are specified, and the solution of the problem is given by the integration of the component differential equations

$$F_x = m\ddot{x} \qquad F_y = m\ddot{y} \qquad F_z = m\ddot{z} \qquad \text{etc.}$$

The differential equations and the solutions of some commonly discussed problems in particle dynamics are given below.

1. Body of mass m under a constant gravitational acceleration g:

$$m\ddot{y} = mg \qquad m\ddot{x} = 0$$

$$\dot{y} = gt + \dot{y}_0 \qquad \dot{x} = \dot{x}_0$$

$$y = \frac{1}{2gt^2} + \dot{y}_0 t + y_0 \qquad x = \dot{x}_0 t + x_0$$

$$y - y_0 = \tfrac{1}{2}g \left(\frac{x - x_0}{\dot{x}_0} \right)^2 + \dot{y}_0 \frac{x - x_0}{\dot{x}_0}$$

The details of the motion depend on the initial values ($t = 0$) of the displacements and velocities. For the particular case $y_0 = \dot{y}_0 = x_0 = \dot{x}_0 = 0$:

$$\dot{y} = gt \qquad y = \frac{1}{2gt^2} \qquad \dot{y} = \sqrt{2gy}$$

2. Body falling under constant gravitational force and a drag force proportional to the velocity ($F_d = -kv$):

$$m\ddot{y} = mg - ky \qquad \frac{\dot{y}k}{mg} = 1 - \left(1 - \frac{\dot{y}_0 k}{mg}\right)e^{-kt/m}$$

$$\frac{yk^2}{m^2g} = \frac{y_0 k^2}{m^2 g} + \frac{kt}{m} - \left(1 - \frac{\dot{y}_0 k}{mg}\right)(1 - e^{-kt/m}) \qquad \dot{y}_{max} = \frac{mg}{k} = \text{terminal velocity}$$

For a sphere falling in a fluid medium, Stokes' law states that $F_d = -3\pi d\mu v$, where μ = dynamic viscosity, d = diameter, and $\rho v d/\mu < 1.0$, ρ being the density of the medium.

3. Body falling with drag force proportional to velocity squared ($F_d = -kv^2$), and $y_0 = \dot{y} = 0$:

$$m\ddot{y} = mg - k\dot{y}^2 \qquad \dot{y}\left(\frac{k}{mg}\right)^{1/2} = \tanh\left(\frac{kgt^2}{m}\right)^{1/2} \qquad \frac{yk}{m} = \ln\cosh\left(\frac{kgt^2}{m}\right)^{1/2}$$

4. Projectile retarded by resisting force $F_d = -a - bv^2$, and $x_0 = \dot{y} = 0$ (Poncelet's penetration problem):

$$m\ddot{x} = -a - b\dot{x}^2 \qquad m\ddot{y} = 0$$

$$\tan^{-1}\dot{x}\left(\frac{b}{a}\right)^{1/2} = (ab)^{1/2}\frac{t}{m} + \tan^{-1}\dot{x}_0\left(\frac{b}{a}\right)^{1/2}$$

$$\frac{bx}{m} = \tfrac{1}{2}\ln\frac{(b/a)\dot{x}_0^2 + 1}{(b/a)\dot{x}^2 + 1}$$

5. Projectile with drag force $F_d = -kv^2$, and $x_0 = y_0 = 0$; approximate solution for flat trajectory:

$$m\ddot{x} = -kv^2\cos\phi = -k\dot{x}^2\left[1 + \left(\frac{\dot{y}}{\dot{x}}\right)^2\right]^{1/2} \simeq -k\dot{x}^2$$

$$m\ddot{y} = -mg - kv^2\sin\phi = -mg - k\dot{x}\dot{y}\left[1 + \left(\frac{\dot{y}}{\dot{x}}\right)^2\right]^{1/2} \simeq -mg - k\dot{x}\dot{y}$$

$$\frac{\dot{x}}{\dot{x}_0} = \left(1 + \frac{k\dot{x}_0}{m}t\right)^{-1} \qquad \frac{2k\dot{x}_0\dot{y}}{mg} = \left(1 + \frac{2k\dot{x}_0 y_0}{mg}\right)\left(1 + \frac{k\dot{x}_0 t}{m}\right)^{-1} - \left(1 + \frac{k\dot{x}_0 t}{m}\right)$$

6. Projectile with drag force $F_d = -kv$, and $x_0 = y_0 = 0$:

$$m\ddot{x} = -k\dot{x} \qquad m\ddot{y} = -k\dot{y} - mg$$

$$\frac{xk}{\dot{x}_0 m} = 1 - e^{-kt/m} \qquad \frac{yk^2}{gm^2} = \left(1 + \frac{\dot{y}_0 k}{gm}\right)(1 - e^{-kt/m}) - \frac{k}{m}t$$

7. Mass m performing forced vibrations under action of force F:

$$m\ddot{x} + c\dot{x} + kx = F(t)$$

$$x = \frac{T}{2\pi m} \int_0^t F(r)e^{-c(t-r)/2m} \sin \frac{2\pi}{T}(t-r)\,dr + \frac{\dot{x}_0 T}{2\pi} \sin \frac{2\pi}{T} t + x_0 \cos \frac{2\pi}{T} t$$

$$T = \frac{2\pi}{[k/m - (c/2m)^2]^{1/2}}$$

8. Planetary motion of a particle of mass m about a fixed point with inverse-square attraction. For plane motion (polar coordinates):

$$m(2\dot{r}\dot{\theta} + r\ddot{\theta}) = 0 \qquad m(\ddot{r} - r\dot{\theta}^2) = -\frac{k}{r^2}$$

$$mr^2\dot{\theta} = h = \text{constant angular momentum}$$

$$\frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - \frac{k}{r} = E = \text{constant energy}$$

$$r = \frac{(h/m)^2}{E(1 + e \cos \theta)} \text{ (orbit)} \qquad e^2 = 1 + \frac{2Eh^2}{mk^2}$$

The orbital equation is that of a conic having a focus at the origin. If $e < 1$, the orbit is an ellipse; if $e = 1$, it is a parabola; and if $e > 1$, it is a hyperbola; e depends on the magnitudes of the initial E and h but not on the direction of initial velocity.

3. Centrifugal and Centipetal Forces. Let a particle of mass m move in a circle of radius r about a fixed axis (Fig. 5.24). The resultant of all forces acting on the particle has a normal component

$$F_c = mr\omega^2 = \frac{mv^2}{r} \tag{5.100}$$

and a tangential component $= mr\alpha$. If ω is constant and $\alpha = 0$, the resultant force acting on the particle to make it rotate in its circular path is $mr\omega^2$ toward the axis, and is called *centripetal* force. *Centrifugal* force for the particle is equal and op-
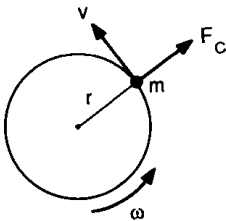
FIGURE 5.24
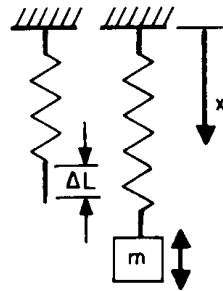
FIGURE 5.25

posite to centripetal force, and is exerted by the particle upon its neighboring particles, or upon the axis of rotation.

4. Harmonic Oscillation

Mechanical Oscillation. For the spring-mass system (Fig. 5.25), motion is described by the following equation:

$$m\ddot{x} = -kx \tag{5.101}$$

and the solution for period of oscilation and frequency are

$$T = 2\pi\sqrt{\frac{m}{k}} \tag{5.102}$$

$$f = \frac{1}{T} = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \tag{5.103}$$

where k (spring stiffness) $= G/\Delta l = mg/\Delta l$

Also, the angular frequency (angular velocity) $\omega = 2\pi f$. Values of k [Eq. (5.103)] for various types of springs are given in Table 5.7.

Pendulum. Basic relations are given for simple, conical, and compound pendulums only.

  Simple Pendulum. See Fig. 5.26. The basic equation for this is

$$T = 2\pi\sqrt{\frac{l}{g}} \tag{5.104}$$

  Conical Pendulum. See Fig. 5.27. The basic equations are

$$T = 2\pi\sqrt{\frac{h}{g}}$$

$$= 2\pi\sqrt{\frac{l\cos\beta}{g}} \tag{5.105}$$

$$\tan\beta = \frac{r}{h} = \frac{r\omega^2}{g} \tag{5.106}$$

  Compound Pendulum. See Fig. 5.28. The basic relations are
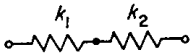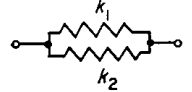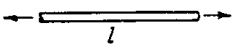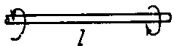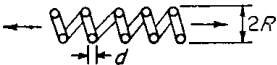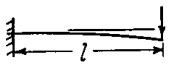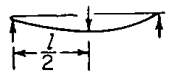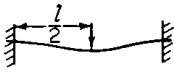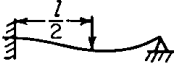
$$T = 2\pi\sqrt{\frac{I_0}{mg\bar{l}}} \tag{5.107}$$

$$I_0 = I_G + m\bar{l}^2 \tag{5.108}$$

$$I_G \approx mg\bar{l}\left(\frac{T^2}{4\pi^2} - \frac{\bar{l}}{g}\right)$$

Point G in Fig. 5.28 is the center of gravity of mass m.

TABLE 5.7 Stiffness of Various Types of Springs

	$k = \dfrac{1}{1/k_1 + 1/k_2}$
	$k = k_1 + k_2$
	$k = \dfrac{EI}{l}$
	$k = \dfrac{EA}{l}$
	$k = \dfrac{GJ}{l}$
	$k = \dfrac{Gd^4}{64nR^3}$
	$k = \dfrac{3EI}{l^3}$
	$k = \dfrac{48EI}{l^3}$
	$k = \dfrac{192EI}{l^3}$
	$k = \dfrac{768EI}{7l^3}$
	$k = \dfrac{3EIl}{a^2b^2}$

Note:
I = moment of inertia of cross-sectional area
l = total length
A = cross-sectional area
J = torsion constant of cross section ($= \frac{1}{2}\pi r^4$ for circular cross section)
n = number of turns
E = modulus of elasticity
G = shear modulus $\simeq 0.385E$
Source: From Flügge.[5]

5. Work

Work of a Force. The work of a variable force in moving a body through the distance $\Delta s = s_2 - s_1$ is given as

$$W = \int_{s_1}^{s_2} F \cos \alpha \, ds \tag{5.109}$$

where F is the variable force, ds is the elementary length of the path, and α is angle between the force and the element ds.
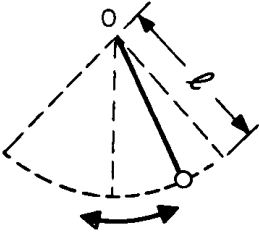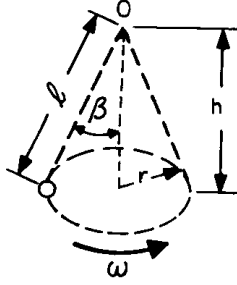
FIGURE 5.26   **FIGURE 5.27**   **FIGURE 5.28**
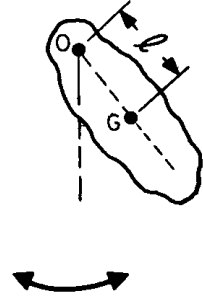
Work of Gravity. The work of gravity on a body in any motion is equal to the product of the weight and the change in height of the mass center.

Work of a Torque. The work of a torque (twisting moment) T on a rotating body for an angular displacement $d\theta$ (in radians) is

$$W = \int_{\theta_1}^{\theta_2} T \, d\theta \tag{5.110}$$

If T is constant, $W = T(\theta_2 - \theta_1)$.

Mechanical Efficiency. Mechanical efficiency η of a machine is the ratio of useful output to total input of work:

$$\eta = \frac{W_u}{W_a} \tag{5.111}$$

and $W_a = W_u + W_f$ where W_a = work applied to the machine, W_u = useful work performed, and W_f = work required to overcome friction or other types of resistance.

6. Energy. Energy (potential and kinetic) of a body is the amount of work it can do by virtue of its position or its motion against forces applied to it. The unit of energy is same as the unit of work.

Potential Energy. The potential energy of a body is that possessed by virtue of its configuration. For example, a body of weight W, located at a height above the earth's surface such that its mass center can descend h, has a potential energy

$$E_p = Wh \tag{5.112}$$

Kinetic Energy of Translation. If a body has a motion of translation, each particle of the body is moving with the same velocity; so at each instant of velocity v, the kinetic energy of the body is

$$E_k = \tfrac{1}{2}mv^2 \tag{5.113}$$

Kinetic Energy of Rotation. In a motion of rotation, each particle of the body is moving with a different velocity, but if ρ represents the radius of any particle and ω the angular velocity of the body, the kinetic energy of the body is

$$E_k = \tfrac{1}{2} \int \rho^2 \omega^2 \; dm = \tfrac{1}{2} I \omega^2 \tag{5.114}$$

where I = moment of inertia (mass) about the axis of rotation.

Kinetic Energy of Translation and Rotation. A body which has a motion of comined translation and rotation has both a kinetic energy of translation and a kinetic energy of rotation. If v is the velocity of the center of gravity at any instant and ω is the angular velocity, the kinetic energy of the body is

$$E_k = \tfrac{1}{2} m v^2 + I \omega^2 \tag{5.115}$$

7. Power. Power is the rate at which work is being done. The unit of power is one unit of work performed in one unit of time. If a force F is acting upon a body and moving it in the direction of the force with a velocity v, the power is

$$P = Fv \tag{5.116}$$

The power of a torque at any instant is

$$P = T\omega \tag{5.117}$$

where ω is instantaneous angular velocity of the body.

8. Impulse and Impact

Linear Impulse and Momentum. The *impulse* of a force is the product of the force and the time during which it acts. The total impulse during time t is

$$\int_0^t F \; dt$$

If the force is constant, this becomes $F \cdot t$. Impulse is a vector quantity, its direction being the same as the direction of the force F.

  Momentum is the product of a mass m and velocity v:

$$\text{momentum} = mv$$

Momentum is also a vector quantity, its direction being the same as the direction of the velocity v.

  An alternative statement of Newton's second law of motion is that the resultant of an unbalanced force system must be equal to the time rate of change of linear momentum:

$$\Sigma F = \frac{d(mv)}{dt} \tag{5.118}$$

Also,

$$\int F \, dt = m(v_1 - v_2) = \text{the change of the momentum of the body}$$

Direct Central Impact. If two inelastic bodies collide in a direct central impact, they have a common velocity v after their impact. Since the impulses are equal and opposite, there is no change in momentum, or

$$m_1 v_1 + m_2 v_2 = (m_1 + m_2)v$$

If the two bodies are elastic, they separate, but their relative velocities after impact are less than before impact. The ratio of the relative velocity of each after impact to that before is e, the coefficient of restitution. There is no change in linear momentum, so

$$m_1 v_1 + m_2 v_2 = m_1 v_1' + m_2 v_2' \tag{5.119}$$

Also,

$$e = \frac{(v_2' - v_1')}{(v_1 - v_2)} \tag{5.120}$$

Angular Impulse and Angular Momentum. The moment of an impulse is the product of the impulse and the distance from the force to the center of moments. It is also called *angular impulse.* Angular impulse and linear impulse cannot be added.

The moment of a momentum is the product of the momentum and the distance from the center of mass to the center of moments. It is also called *angular momentum.* Its unit is $I\omega$.

In any motion of rotation, the initial angular momentum plus the positive angular impulse minus the negative angular impulse is equal to the final angular momentum. In any mutual action between two rotating bodies or two parts of the same rotating body, the angular impulses are equal in amount and opposite in direction; hence for the entire system the change in angular momentum must be zero.

NOMENCLATURE

Symbol = Definition, SI units (U.S. Customary unit)

A = area, m² (ft²)

a = acceleration, m/s² (ft/s²)

E_k = kinetic energy, J (lb$_f$ · ft)

E_p = potential energy, J (lb$_f$ · ft)

F = force, N (lb$_f$)

f = sliding friction coefficient

f_0 = static friction coefficient

f = frequency of oscillation, s^{-1} or 1/s

g = acceleration of gravity, m/s² (ft/s²)

I = moment of inertia of an area, m⁴ (ft⁴)
I = moment of inertia of mass, kg · m² (lb · ft²)
l = distance, m (ft)
m = mass, kg (lb)
M = momentum, N · m (lb$_f$ · ft)
P = power, J/s(lb$_f$ · ft/s)
r = radius (or diameter), m (ft)
r = radius of gyration, m (ft)
s = distance (displacement), m (ft)
T = time period of oscillation, s
T = torque (twisting moment), N · m (lb$_f$ · ft)
t = time, s
v = velocity, m/s (ft/s)
V = volume, m³ (ft³)
W = work, N · m (lb$_f$ · ft)
W = weight of a body, N (lb$_f$)
x = rectangular coordinate, m (ft)
y = rectangular coordinate, m (ft)
z = rectangular coordinate, m (ft)

Subscripts

0 = initial condition
t = tangential
n = normal
x = x direction
y = y direction
z = z direction
N = normal direction
R = reaction

Greek

α = angular acceleration, s⁻², rad/s² (in "Rotation" subsection)
α = angle, rad, deg
ρ = mass density, kg/m³ (lb/ft³)
ρ = distance, m (ft) (in "Moment of Inertia" subsection)
μ = dynamic viscosity, N · s/m² [lb/(ft · s)]
ω = angular velocity, s⁻¹, rad/s

Superscripts

$'$ = ft

$''$ = inch

$^-$ = time average ($\bar{x}$ is time average of x)

Mathematical Operation Symbols

d/dt = derivative with respect to t, s^{-1}

$\dot{}$ = first derivative ($\dot{x} = dx/dt$)

$\ddot{}$ = second derivative ($\ddot{x} = d^2x/dt^2$)

Note. Other symbols are defined in the text. Boldfaced symbols in the text denote vector quantities.

REFERENCES*

1. K. Gieck, *Engineering Formulas,* 5th ed., McGraw-Hill, New York, 1986.

2. E. A. Avalione and T. Baumeister III (eds.), *Mark's Standard Handbook for Mechanical Engineers,* 9th ed., McGraw-Hill, New York, 1987.

3. C. E. O'Rourke, *General Engineering Handbook,* 2d ed., McGraw-Hill, New York, 1940.

4. H. A. Rothbart, *Mechanical Design and Systems Handbook,* 2d ed., McGraw-Hill, New York, 1985.

5. W. Flügge, *Handbook of Engineering Mechanics,* McGraw-Hill, New York, 1962.

6. F. P. Beer and E. R. Johnson, *Vector Mechanics for Engineers,* McGraw-Hill, New York, 1964.

7. N. H. Cook, ''Mechanics of Solids,'' Course 2.01, MIT class notes, 1973–1976.

8. S. H. Crandall, ''Dynamics,'' Course 2.032, MIT class notes, 1973–1976.

9. S. H. Crandall, et al., *Dynamics of Mechanical and Electromechanical Systems,* McGraw-Hill, New York, 1968.

10. J. P. Den Hartog, *Mechanics,* Dover, New York, 1961.

11. O. W. Eshbach and M. Souders, *Handbook of Engineering Fundamentals,* 3d ed., John Wiley & Sons, New York, 1975.

12. C. W. Ham, E. J. Crane, and W. L. Rogers, *Mechanics of Machinery,* 4th ed., McGraw-Hill, New York, 1958.

13. J. L. Synge and B. A. Griffith, *Principles of Mechanics,* 3d ed., McGraw-Hill, New York, 1959.

14. S. Timoshenko and D. H. Young, *Engineering Mechanics,* 4th ed., McGraw-Hill, New York, 1956.

*Those references listed above but not cited in the text were used for comparison between different data sources, clarification, clarity of presentation, and, most important, reader's convenience when further interest in the subject exists.

15. S. Timoshenko and D. H. Young, *Theory of Structures,* McGraw-Hill, New York, 1945.

16. D. N. Wormley, ''Dynamic Systems,'' Course 2.023, MIT class notes, 1973–1976.

17. J. Bird, *Newnes Engineering Science Pocket Book,* 3d ed., Newnes, 2001.

18. F. S. Merritt, M. K. Lofin and J. T. Ricketts, *Standard Handbook for Civil Engineers,* 4th ed., McGraw-Hill, 1996.

CHAPTER 6
MECHANICS OF DEFORMABLE BODIES

STATIC STRESSES*

1. Compression and Tension. If a bar with a cross-sectional area of A is acted upon by two equal and oppositely directed axial forces F, F, the load per unit area, called the *unit stress S,* is given by the expression

$$S = \frac{F}{A} \tag{6.1}$$

If the forces are acting toward each other, the stress is *compression* and is denoted by S_c. If the forces are acting away from each other, the stress is *tension* and is denoted by S_t.

2. Shear. If two equal and oppositely directed forces F, F, are applied normal to the axis of a bar and in different planes of action, the part of the bar between the planes of action of the forces is subjected to a shearing action in which any cross section tends to slide over the one next to it. Unit *shearing stress* is denoted by S_s. The shearing stress is not uniform across the cross section, but the average unit stress is given by the expression $S_s = F/A$. (See Fig. 6.1.)

3. Modulus of Elasticity and Shearing Modulus of Elasticity. Elasticity is the ability of a material to return to its original dimension after the removal of stresses. Nearly all of the materials used in engineering work are elastic, and within certain limits obey fairly well *Hooke's law* of proportionality of stress to deformation or strain. If l is the length of a bar and e its total change in length, the *unit deformation or strain* is

$$\delta = \frac{e}{l} \tag{6.2}$$

The ratio of the unit stress S to the corresponding unit strain δ is called the *modulus of elasticity* and is denoted by E:

---

*This section is taken in part from Refs. [1, 11, 13].
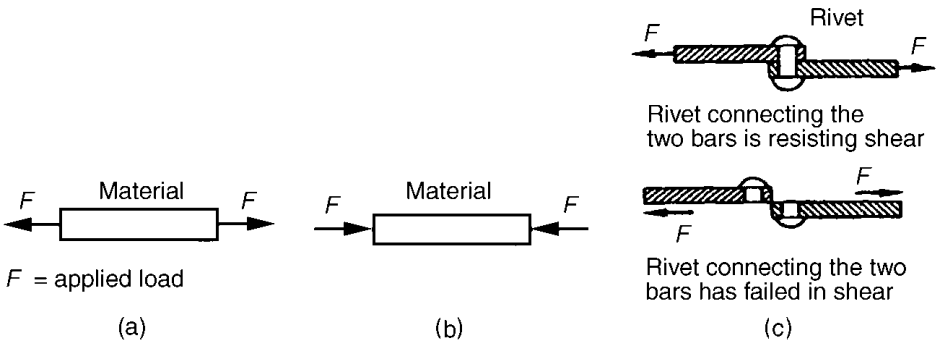
FIGURE 6.1   (*a*) Tension; (*b*) compression; (*c*) shear.

$$E = \frac{S}{\delta} \tag{6.3}$$

For nearly all materials, the values of E in tension and compression are practically the same. The modulus of elasticity in tension is also known as *Young's modulus.*

In a similar way, the ratio of the unit shearing stress S_s to the corresponding unit shearing strain δ_s is called the *shearing modulus of elasticity E_s:*

$$E_s = \frac{S_s}{\delta_s} \tag{6.4}$$

For homogeneous materials, E_s is about two-fifths of the value of E. Symbol G is often used instead of E_s and is often referred to as the *modulus of rigidity.*

4. Elastic Limit, Yield Point, and Ultimate Strength.   The maximum unit stress for which Hooke's law is valid is called the *proportional elastic limit.* The elastic limit is also defined as the maximum stress to which the body may be subjected without permanent deformation or set. As the axial dimension of a bar is changed by stress, any lateral dimension is changed oppositely. The ratio of the unit lateral strain to the unit axial strain is called *Poisson's ratio μ:*

$$\mu = \frac{\text{lateral strain}}{\text{longitudinal strain}} \tag{6.5}$$

If the load on a bar of ductile material is increased above the elastic limit, it soon reaches a value at which the deformation continues to increase with little or no increase of the load. The unit stress at which this occurs is called the *yield point.* Materials which are not ductile have no yield point.

The greatest load a bar will hold is called the *maximum load,* and the corresponding unit stress is the *ultimate strength.* Ductile materials tested in tension form a neck at about the time the ultimate strength is reached, and the total load carried then decreases. The stretching continues, and the bar breaks at a total load less than the maximum. This load is called the *rupture* or *breaking load,* and the unit stress at rupture is called the *rupture strength* (see Fig. 6.2). Unless specified as actual stresses, all of these four stresses—the elastic limit, the yield point, the ultimate strength, and the rupture strength—are computed by dividing the loads by the original cross-sectional area.
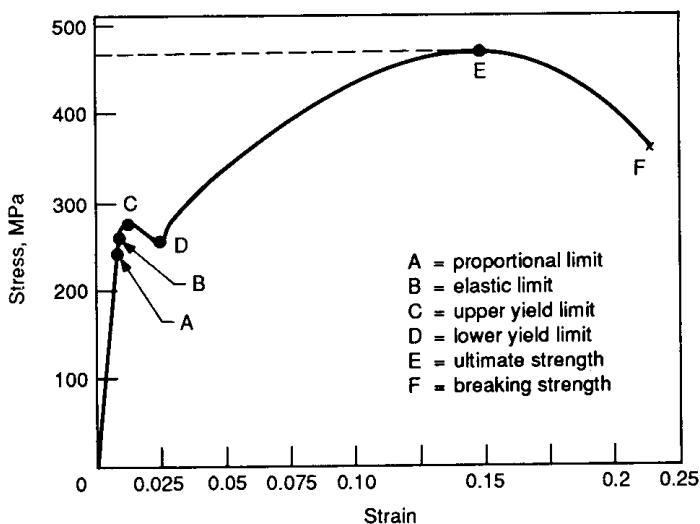
FIGURE 6.2   Stress-strain diagram for mild steel (1 Pa = 1 N/m² = 1.45 × 10⁻⁴ lb$_f$/in²).

5. Bulk Modulus of Elasticity, Reliance, Thermal Stress, and Design Stress.

The *bulk modulus of elasticity K* is the ratio of normal stress, applied to all six faces of a cube, to the change of volume,

$$K = \frac{S}{\Delta V / V_0} \tag{6.6}$$

where V_0 is the original volume and ΔV is the change in the volume equal to $V_0 - V$.

The following relations are useful for calculations (Ref. 2):

$$G = \frac{E}{2(1 + \mu)} \tag{6.7}$$

and

$$\frac{1}{E} = \frac{1}{9K} + \frac{1}{3G} \tag{6.8}$$

Reliance U is the potential energy stored in a deformed body.

Thermal stress is developed if expansion or contraction is prevented; it is equal to

$$S = E \cdot \alpha \cdot \Delta t \tag{6.9}$$

where α is the linear coefficient of thermal expansion and $\Delta t = t_2 - t_1$ is the temperature rise from t_1 to t_2.

The design stress is determined by dividing the applicable material property (yield stress, ultimate stress) by a factor of safety.

6. Plasticity, Ductility, Malleability and Hardness.*

Plasticity. This property is the exact opposite of elasticity. It is the state of a material that has been loaded beyond its elastic state. Under a load beyond that required to cause elastic deformation (the elastic limit), a material possessing the property of plasticity deforms permanently. It takes a permanent set and will not recover when the load is removed. (See Fig. 6.3.)

Ductility. This is the term used when plastic deformation occurs as the result of applying a tensile load. A ductile material combines the properties of plasticity and tenacity (tensile strength) so that it can be stretched or drawn to shape and will retain that shape when the deforming force is removed. For example, in wire drawing the wire is reduced in diameter by drawing it through a die. (See Fig. 6.4.)

Malleability. This is the term used when plastic deformation occurs as the result of applying a compressive load. A malleable material combines the properties of plasticity and compressibility, so that it can be squeezed to shape by such processes as forging, rolling and rivet heading. (See Fig. 6.5.)

Before loading / During loading / After load is removed

FIGURE 6.3

Ductile material — Die — Wire — F — Direction of draw
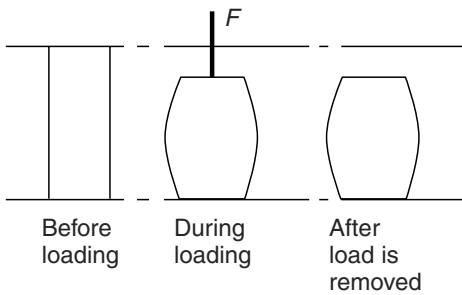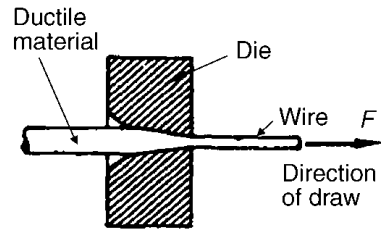
F = applied load (tensile)

FIGURE 6.4

F — 'Set' forms rivet head by compressive loading
Rivet
'Dolly' supports rivet head
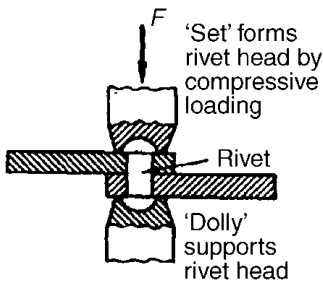
F = applied load (compressive)

FIGURE 6.5

Prescribed load — Hardened ball indenter — Material being tested

Hard material: little indentation

Same prescribed load

Soft material: greater indentation for same load
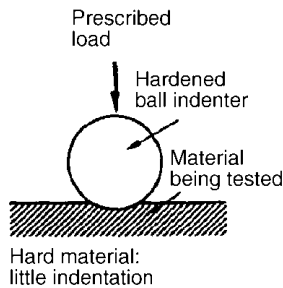
FIGURE 6.6
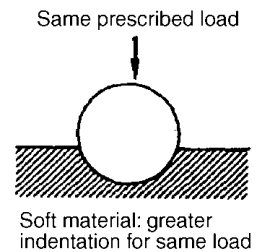
* Refs. 11, 13, 14.

Hardness. This is the ability of a material to withstand scratching (abrasion) or indentation by another hard body. It is an indication of the wear resistance of a material. Processes that increase the hardness of materials also increase their tensile strength. At the same time the toughness of the material is reduced as it becomes more brittle. Hardenability must not he confused with hardness. Hardenability is the ability of a metal to respond to the heat treatment process of quench hardening. To harden it, the hot metal must he chilled at a rate in excess of its critical cooling rate. Since any material cools more quickly at the surface than at the center, there is a limit to the size of bar that can cool quickly enough at its center to achieve uniform hardness throughout. This is the ruling section for the material. The greater its hardenability, the greater will be its ruling section. (See Fig. 6.6.)

DYNAMIC STRESSES

Dynamic stresses occur where the dimension of time is necessary in defining the loads. They include creep, fatigue, and impact stresses.

Creep stresses occur when either the load or the deformation progressively varies with time. They are usually associated with noncyclic phenomena.

Fatigue stresses occur when the cyclic variation of either load or strain is co-incident with respect to time.

Impact stresses occur in regard to loads that are transient in respect to time. The duration of the load application is of the same order of magnitude as the natural period of vibration of the specimen.

For *steady vibration stresses,* the deflection of the bar, or beam, is increased by the dynamic magnification factor K_d:

$$S_{\text{dynamic}} = S_{\text{static}} \cdot K_d \qquad (6.10)$$

and

$$K_d = \frac{1}{1 - (\omega/\omega_n)^2}$$

where ω the frequency of oscillation of the load and ω_n is the natural frequency of the bar determined by

$$\omega_n = \frac{3\,Elg}{L^3 W} \qquad (6.11)$$

where L = length of bar
I = moment of inertia
W = weight of the oscillating load

BEAMS

1. Types of Beams. A beam is a bar or structural member subjected to transverse loads and reactions that tend to bend it. Usually beams are horizontal bars designed to carry vertical loads, but any structural member acts as a beam if bending is induced by external transverse forces.
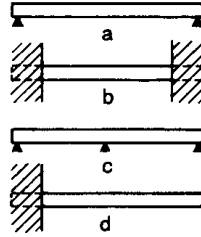
FIGURE 6.7

Beams freely supported at both ends are called *simple beams* (Fig. 6.7*a*). A *fixed beam* is rigidly fixed at both ends or rigidly fixed at one end and simply supported at the other (Fig. 6.7*b*). A *continuous beam* is one that is resting on more than two supports (Fig. 6.7*c*). Beams fixed at one end and unsupported at the other are called *cantilever beams* (Fig. 6.7*d*). Loads are usually concentrated, uniformly distributed over part or all of the length of the beam, or uniformly varying.

2. Shear and Bending Moments in Beams

Vertical Shear.   At any cross section of a beam the resultant of the external vertical forces acting on one side of the section is equal and opposite to the resultant of the external vertical forces on the other side of the section. These forces tend to cause the beam to shear vertically along the section. The value of either resultant, which is a measure of the shearing tendency, is known as the *vertical shear* (V) at the section considered. It is computed by finding the algebraic sum of the vertical forces to the left of the section; i.e., it is equal to the left reaction minus the sum of the vertical downward forces acting between the left support and the section. The vertical shear V is also given as the sum of the transverse shear stresses (S) acting on the section:

$$V = \int S \, dA \tag{6.12}$$

A *shear diagram* is a graphical representation of the vertical shear at all cross sections of the beam.

Bending Moment.   The bending moment, or moment, at any cross section of a beam is the algebraic sum of the moments of the external forces acting on either side of the section. It is considered positive when it causes the beam to bend convex downward, hence causing compression in upper fibers and tension in lower fibers of beam. The shear V is the first derivative of moment with respect to distance x along the beam:

$$V = \frac{dM}{dx} \tag{6.13}$$

Also,

$$M = \int V \, dx \tag{6.14}$$
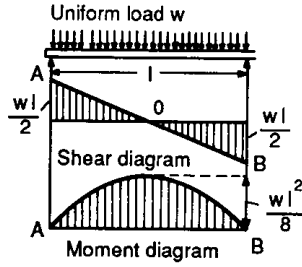
FIGURE 6.8

A *moment diagram* is a line drawn to show the magnitude and character of the bending moment.

Figure 6.8 illustrates a simple beam subjected to a uniform load w per unit length. Then

$$M = R_1 x - wx \cdot \frac{x}{2} = \frac{wlx}{2} - \frac{wx^2}{2}$$

$$V = R_1 - wx = \frac{wl}{2} - wx$$

Also,

$$V = \frac{d}{dx}\left(\frac{wlx}{2} - \frac{wx^2}{2}\right) = \frac{wl}{2} - ws$$

Shear and moment diagrams are also shown in Fig. 6.8. The moment curve is always parabolic under uniformly distributed loads. R_1 = reaction at A.

Table 6.1 gives the reactions, bending moment equations, vertical shear equations, and deflection of some of the more common types of beams.

Flexule Formula. The concave side of a bent beam (Fig. 6.9a and b) is in compression and the convex side in tension. They are divided by the neutral plane of zero stress $A'B'BA$. The intersection of the neutral plane with the face of the beam is in the neutral line or elastic curve AB. The neutral axis NN' is the intersection of the neutral plane with the cross section.

The neutral axis contains the center of gravity of the cross section.

The flexule formula

$$S = \frac{Mc}{I} \tag{6.15}$$

is basic to the design and investigation of beams. It holds only when the maximum horizontal fiber stress S does not exceed the proportional limit of the material. c = distance of that fiber from the neutral axis. The I/c factor, the section modulus, is

TABLE 6.1　Bending Moment, Vertical Shear, and Deflection of Beams of Uniform Cross Section and Various Conditions of Loading

$R_2 = W$

$M_x = -Wx$

$M_{max} = -Wl, (x = l)$

$V_x = -W$

$f = \dfrac{Wl^3}{3EI}$ (max)
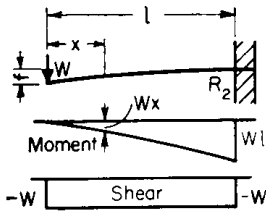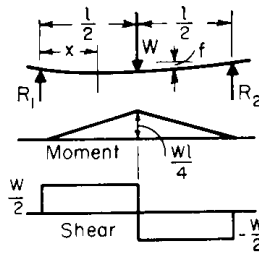
$R_1 = \dfrac{W}{2}, R_2 = \dfrac{W}{2}$

$M_x = \dfrac{Wx}{2}$

$M_{max} = \dfrac{Wl}{4}, \left(x = \dfrac{l}{2}\right)$

$V_x = \pm \dfrac{W}{2}$

$f = \dfrac{W}{EI}\dfrac{l^3}{48}$ (max)
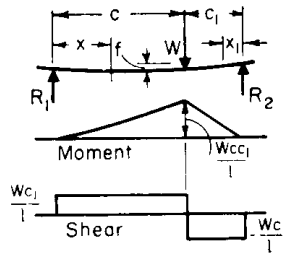
$R_1 = \dfrac{Wc_1}{l}, R_2 = \dfrac{Wc}{l}$

$M_x = \dfrac{Wc_1 x}{l}, \; Mx' = \dfrac{Wcx_1}{l}$

$M_{max} = \dfrac{Wcc_1}{l}, \; (x_1 = c_1 \text{ or } x = c)$

$V_x = \dfrac{Wc_1}{l}, \; V_{x1} = \dfrac{Wc}{l}$

$f = \dfrac{Wc_1}{3EIl}\left[\dfrac{c(l + c_1)}{3}\right]^{3/2}$ (max)

Max f occurs at $x = \sqrt{c(l + c_1)/3}$
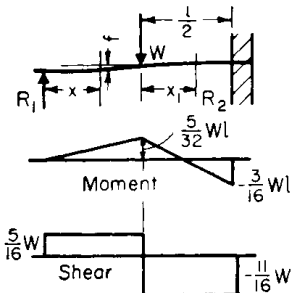
$R_1 = \tfrac{5}{16} W, R_2 = \tfrac{11}{16} W$

$M_x = \tfrac{5}{16} Wx$

$M_{x1} = Wl\left(\dfrac{5}{32} - \dfrac{11}{16}\dfrac{x_1}{l}\right)$

$M_{max} = -\tfrac{3}{16} Wl, (x_1 = \tfrac{1}{2})$

$V_x = \tfrac{5}{16} W, \; V_{x1} = -\tfrac{11}{16} W$

$V_{max} = -\tfrac{11}{16} W,$

$\left(x = \dfrac{l}{2} \text{ to } x = l\right)$

$f = \dfrac{W}{EI}\dfrac{7l^3}{768}$
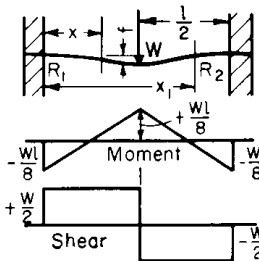
$R_1 = \dfrac{W}{2}, R_2 = \dfrac{W}{2}$

$M_x = \dfrac{Wl}{2}\left(\dfrac{x}{l} - \dfrac{1}{4}\right)$

$M_{x1} = \dfrac{-Wl}{2}\left(\dfrac{x}{l} - \dfrac{3}{4}\right)$

$M_{max} = \dfrac{Wl}{8}\left(x = \dfrac{1}{2}\right)$

$V_x = \dfrac{W}{2}, \; V_{x1} = -\dfrac{W}{2}$

$f = \dfrac{W}{EI}\dfrac{l^3}{192}$ (max)
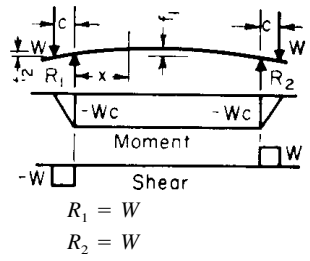
$R_1 = W$

$R_2 = W$

$M_x = -Wc = $ const

V_w to $R_1 = -W$

V_{R1} to $R_2 = 0$

V_{R2} to $W = +W$

$f_1 = \dfrac{Wcl^2}{EI8}$ (max)

$f_2 = \dfrac{Wc^2}{EI3}\left(c + \dfrac{3l}{2}\right)$ (max)

TABLE 6.1 Bending Moment, Vertical Shear, and Deflection of Beams of Uniform Cross Section and Various Conditions of Loading (*Continued*)

$$R_2 = W = wl$$

$$M_x = -\frac{wx^2}{2}$$

$$M_{max} = -\frac{wl^2}{2} \quad (x = l)$$

$$V_x = -wx$$

$$V_{max} = -wl \quad (x = l)$$

$$f = \frac{W}{EI}\frac{l^2}{8} \quad (max)$$
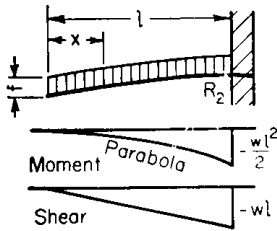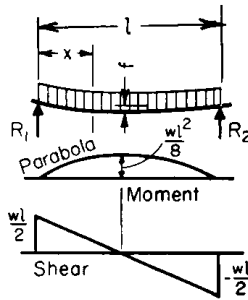
$$R_1 = \frac{W}{2} = \frac{wl}{2}$$

$$R_2 = \frac{W}{2} = \frac{wl}{2}$$

$$M_x = \frac{wx}{2}(l - x)$$

$$M_{max} = \frac{wl^2}{8} \quad (x = \tfrac{1}{2}l)$$

$$V_x = \frac{wl}{2} - wx$$

$$V_{max} = \frac{wl}{2} \quad (x = 0)$$

$$f = \frac{W}{EI}\frac{5l^3}{384} \quad (max)$$
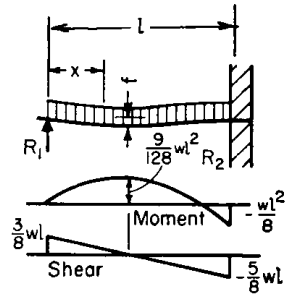
$$R_1 = \tfrac{3}{8}W = \tfrac{3}{8}wl$$

$$R_2 = \tfrac{5}{8}W = \tfrac{5}{8}wl$$

$$M_x = \frac{wx}{2}\left(\frac{3}{4}l - x\right)$$

$$M_{max} = \tfrac{9}{128}wl^2 \quad (x = \tfrac{3}{8}l)$$

$$M_{max} = \frac{wl^2}{8} \quad (x = l)$$

$$V_x = \tfrac{3}{8}wl - wx$$

$$V_{max} = -\tfrac{5}{8}wl$$

$$f = \frac{W}{EI}\frac{l^3}{185} \quad (max)$$
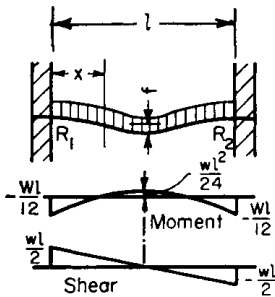
$$R_1 = \frac{W}{2} = \frac{wl}{2}, \ R_2 = \frac{W}{2}$$

$$= \frac{wl}{2}$$

$$M_x = -\frac{wl^2}{2}\left(\frac{1}{16} - \frac{x}{l} + \frac{x^2}{l^2}\right)$$

$$M_{max} = -\tfrac{1}{12}wl^2,$$

$$(x = 0, \text{ or } x = l)$$

$$V_x = \frac{wl}{2} - wx$$

$$V_{max} = \pm\frac{wl}{2}$$

$$f = \frac{W}{EI}\frac{l^2}{384} \quad (max)$$
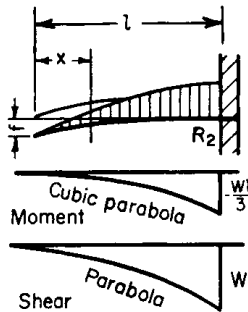
$$R_2 = W = \text{total load}$$

$$M_x = -\frac{Wx^3}{3\,l^2}$$

$$M_{max} = -\frac{Wl}{3}$$

$$V_x = -\frac{Wx^2}{l^2}$$

$$V_{max} = -W$$

$$f = \frac{W}{EI}\frac{l^2}{15} \quad (max)$$
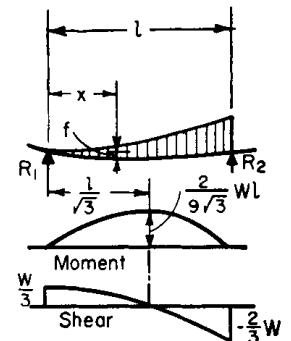
$$R_1 = \tfrac{1}{3}W, \ R_2 = \tfrac{2}{3}W$$

$$M_x = \frac{Wx}{3}\left(1 - \frac{x^2}{l^2}\right)$$

$$M_{max} = \frac{2}{9\sqrt{3}}Wl \quad \left(x = \frac{1}{\sqrt{3}}\right)$$

$$V_x = W\left(\frac{1}{3} - \frac{x^2}{l^2}\right)$$

$$V_{max} = -\tfrac{2}{3}W \quad (x = l)$$

$$f = 0.01304\frac{Wl^2}{EI} \quad (max)$$

TABLE 6.1 Bending Moment, Vertical Shear, and Deflection of Beams of Uniform Cross Section and Various Conditions of Loading (*Continued*)
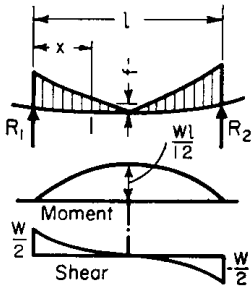
$$R_1 = \frac{W}{2}, R_2 = \frac{W}{2}$$

$$M_x = Wx\left(\frac{1}{2} - \frac{x}{l} + \frac{2x^2}{3l^2}\right)$$

$$M_{max} = \frac{Wl}{12} \ (x = \tfrac{1}{2}l)$$

$$V_x = W\left(\frac{1}{2} - \frac{2x}{l} + \frac{2x^2}{l^2}\right)$$

$$V_{max} = \pm\frac{W}{2} \ (x = 0)$$

$$f = \frac{W}{EI}\frac{3l^3}{320} \ (max)$$
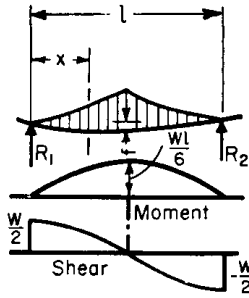
$$R_1 = \frac{W}{2}, R_2 = \frac{W}{2}$$

$$M_x = Wx\left(\frac{1}{2} - \frac{2x^2}{3l^2}\right)$$

$$M_{max} = \frac{Wl}{6} \ (x = \tfrac{1}{2}l)$$

$$V_x = W\left(\tfrac{1}{2} - \frac{2x^2}{l^2}\right)$$

$$V_{max} = \pm\frac{W}{2} \ (x = 0)$$

$$f = \frac{W}{EI}\frac{l^3}{60} \ (max)$$
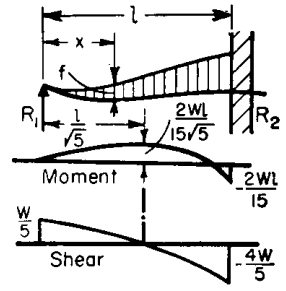
$$R_1 = \frac{W}{5}, R_2 = \frac{4W}{5}$$

$$M_x = Wx\left(\frac{1}{5} - \frac{x^2}{3l^2}\right)$$

$$M_{max} = -\tfrac{2}{15}\,Wl \text{ at support 2}$$

$$V_x = W\left(\frac{1}{5} - \frac{x^2}{l^2}\right)$$

$$V_{max} = -\frac{4W}{5}$$

$$f = \frac{16Wl^2}{1500\sqrt{5}EI}$$

$$= \frac{0.00477Wl^3}{EI} \ (max)$$
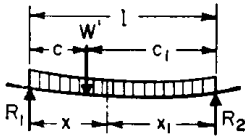
Concentrated load W'

Uniformly dist. load $W = wl$; $c < c_1$

$$R_1 = W'\frac{c_1}{l} + \frac{W}{2}$$

$$R_2 = W'\frac{c}{l} + \frac{W}{2}$$

(a) $\dfrac{W'}{W} < \dfrac{c_1 - c}{2c}$

$$M_{max} = R_2\frac{x_1}{2} = \frac{R_2^2 l}{2W}\left(x_1 = \frac{R_2 l}{W}\right)$$

(b) $\dfrac{W'}{W} > \dfrac{c_1 - c}{2c}$

$$M_{max} = \left(W' + \frac{W}{2}\right)\frac{cc_1}{l} \ (x_1 = c_1)$$

Deflection of beam under W':

$$f = \left(W' + \frac{l^2 + cc_1}{8cc_1}W\right)\frac{c^2c_1^2}{3EIl}$$
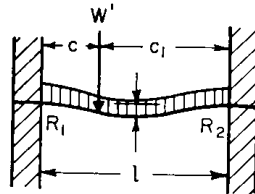
$c < c_1$

$$R_1 = W'\frac{(3c + c_1)c_1^2}{l^3} + \frac{W}{2}$$

$$R_2 = W'\frac{(c + 3c_1)c^2}{l^3} + \frac{W}{2}$$

$$M_{max} = M_1 = W'\frac{cc_1^2}{l^2} + \frac{Wl}{12}$$

Deflection under W'

$$f = \frac{1}{EI}\left(W'\frac{c^3c_1^3}{3l^3} + W\frac{c^2c_1^2}{24l}\right)$$

TABLE 6.1 Bending Moment, Vertical Shear, and Deflection of Beams of Uniform Cross Section and Various Conditions of Loading (*Continued*)
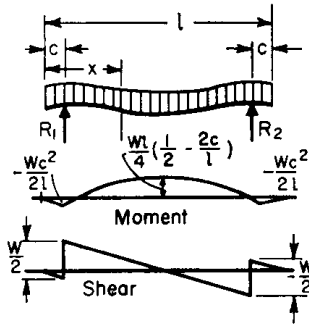
$$R_1 = \frac{W}{2} = \frac{wl}{2}, \quad R_2 = \frac{W}{2} = \frac{wl}{2}$$

$$M_x = \frac{Wx}{2}\left(1 - \frac{c}{x} - \frac{x}{l}\right) \quad (x > c)$$

$$M_x = \frac{Wx^2}{2l} \quad (x \leq c)$$

$$M_{max} = \frac{Wl}{4}\left(\frac{1}{2} - \frac{2c}{l}\right), \quad c \leq \left(\frac{\sqrt{2} - 1}{2}\right)l$$

$$V_x = \frac{W}{2} - wx \quad (x > c)$$

$$V_x = -wx \quad (x \leq c)$$
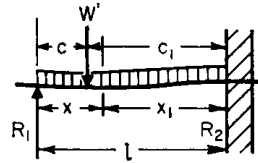
Concentrated load W'

Uniformly dist. load $W = wl$

$$R_1 = W'\frac{c_1^2(3c + 2c_1)}{2l^3} + \tfrac{3}{8}W$$

$$R_2 = W'\frac{(2c^2 + 6cc_1 + 3c_1^2)c}{2l^3} + \tfrac{5}{8}W$$

$$M_2 = W'\frac{cc_1(2c + c_1)}{2l^3} + W\frac{l}{8}$$

$$Mw' = W'\frac{cc_1^2(3c + 2c_1)}{2l^3} + W\frac{(3c_1 - c)c}{8l}$$

$$(a) \quad \frac{W}{W'} < \frac{l^2}{4c_1^2}\frac{5c - 3c_1}{2c_1}$$

$$M_{c\ max} = \frac{R_1^2}{2W}l\left(x = \frac{R_1 l}{W}\right)$$

$$(b) \quad \frac{W'}{W} < \frac{l^2(3c_1 - 5c)}{4c(2c^2 + 6cc_1 + 3c_1^2)}$$

$$M_{c1\ max} = W'c + \frac{(R_1 - W')^2}{2W}l\left(x = \frac{R_1 - W'}{W}l\right)$$

Deflection under W'

$$f = \frac{W'}{EI}\frac{c^2c_1^3(4c + 3c_1)}{12l^3} + \frac{W}{EI}\frac{cc_1^2(3c + c_1)}{48l}$$

Note:

R_1, R_2 = reactions

w = distributed load per longitudinal unit

W = total distributed load

M = bending moment

M_x = local value of bending moment

M_{max} = maximum value of M

V = vertical shear

V_x = local vertical shear (shear at any section)

V_{max} = maximum value of V

f = deflection

E = modulus of elasticity

I = moment of inertia

l = distance between supports

W' = concentrated load

Source: Avallone and Baumeister.[3]
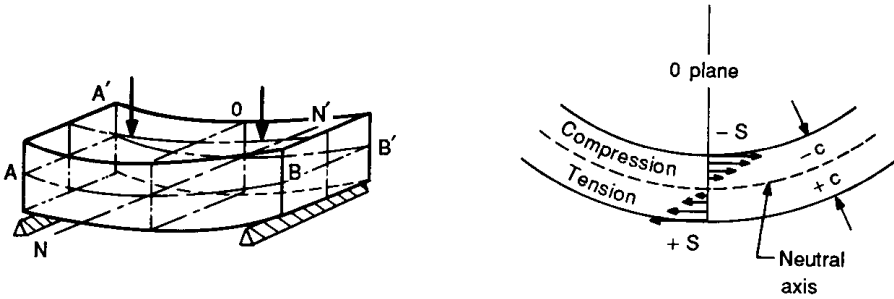
FIGURE 6.9

a measure of the capacity of a section to resist any bending moment M to which it may be subjected. I = moment of inertia of the cross section with respect to its neutral axis.

Values of I and I/c for simple shapes used as beams are given in Table 6.2.

Rolling Loads.* These may change their position on the beam. Figure 6.10 shows a beam with two equal concentrated moving loads (example: two wheels on a crane girder, or the wheels of a truck on a bridge). As the maximum moment occurs where the shear is zero, it is evident from the shear diagram that the maximum moment will occur under the wheel. As $x < a/2$ (Fig. 6.10), then

$$R_1 = P\left(1 - \frac{2x}{l} + \frac{a}{l}\right)$$

$$M_2 = \frac{Pl}{2}\left(1 - \frac{a}{l} + \frac{2x}{l}\frac{a}{l} - \frac{4x^2}{l^2}\right)$$

$$R_2 = P\left(1 + \frac{2x}{l} - \frac{a}{l}\right)$$

$$M_1 = \frac{Pl}{2}\left(1 - \frac{a}{l} - \frac{2a^2}{l^2} + \frac{2x}{l}\frac{3a}{l} - \frac{4x^2}{l^2}\right)$$
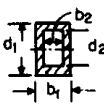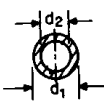
$$M_{2\text{max}} = M_2\left(x = \frac{a}{4}\right)$$

$$M_{1\text{max}} = M_1\left(x = \frac{3a}{4}\right)$$

$$M_{\text{max}} = \frac{Pl}{2}\left(1 - \frac{a}{2l}\right)^2 = \frac{P}{2l}\left(l - \frac{a}{2}\right)^2$$

TABLE 6.2 Properties of Sections of Beams

Section of beam	Moment of inertia I	Section of modulus IC	Radius of gyration r
	$\dfrac{bd^3}{12}$	$\dfrac{bd^3}{6}$	$\dfrac{d}{\sqrt{12}} = 0.289d$
	$\dfrac{b_1d_1^3 - b_2d_2^3}{12}$	$\dfrac{b_1d_1^3 - b_2d_2^3}{6d_1}$	$\sqrt{\dfrac{b_1d_1^3 - b_2d_2^3}{12(b_1d_1 - b_2d_2)}}$
	$\dfrac{\pi d^4}{64}$	$\dfrac{\pi d^3}{32}$	$\dfrac{d}{4}$
	$\dfrac{\pi(d_1^4 - d_2^4)}{64}$	$\dfrac{\pi(d_1^4 - d_2^4)}{32d_1}$	$\dfrac{\sqrt{d_1^2 + d_2^2}}{4}$
	$\dfrac{bd^3}{36}$	$\dfrac{bd^2}{24}$ (min.)	$\dfrac{d}{\sqrt{18}} = 0.236d$

Source: From Avallone and Baumeister.[3]
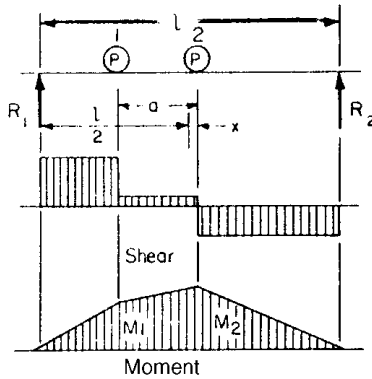
FIGURE 6.10 (*From Avallone and Baumeister.*[3])

COLUMNS*

A column is a bar which is loaded axially in compression. A column is shortened by the compression, and it also tends to deflect laterally, owing partly to the fact that the load cannot be applied symmetrically with respect to the longitudinal axis of the column, and partly to the fact that the material of which the column is made is not perfectly homogeneous. The lateral deflection takes place usually in the direction of the least resisting moment of the section, and ultimate failure is caused by a combination of compression, shearing, and bending stresses. The load that will produce the ultimate failure of a given column is dependent upon the ratio between the length and the lateral resistance of the column. A long column of a given cross section will not support as much load as a shorter column of the same cross section.

If a column has round ends, so that the bending is not restrained, the equation of its elastic curve is

$$EI \frac{d^2y}{dx^2} = -Py \qquad (6.16)$$

when the origin of the coordinate axis is at the top of the column, the positive direction of x being taken downward and the positive direction of y being taken in the direction of the deflection. P = axial load, I = least moment of inertia ($I = Ar^2$) in m⁴ (ft⁴), and E = modulus of elasticity in kg$_f$/m² (lb$_f$/in²). Integrating the above expression twice and determining the constants of integration results in

$$P = \frac{n\pi^2 EI}{l^2} \qquad (6.17)$$

which is Euler's formula for long columns. l = the length of the column.

The coefficient n in Eq. (6.17) accounts for end conditions. When the column is pivoted at both ends, $n = 1$; when one end is fixed and the other rounded, $n = 2$; when both are fixed, $n = 4$; and when one end is fixed with the other free, $n = ¼$. If, under load P, a slight deflection is produced, the column will not return to its original position; if P is decreased, the column will approach its original position; but if P is increased, the deflection will increase until the column fails by bending. For columns with a value of l/r (r = least radius of gyration) less than about 150, Eq. (6.17) gives results distinctly higher than those observed in tests. A theoretical equation for a short column has not been derived. Some empirical formulas for short columns are given in Ref. 3, p. 5-43.

TORSION

A cylindrical bar or shaft which is being twisted about its own axis is said to be *in torsion* (see Fig. 6.11). The twisting moment or torque T is

---

*This section is taken in part from *General Engineering Handbook,* 2d ed., by C. E. O'Rourke. Copyright © 1940. Used by permission of McGraw-Hill, Inc. All rights reserved.
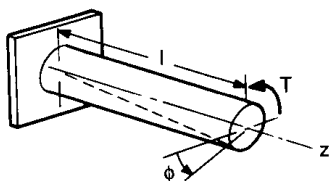
FIGURE 6.11

$$T = \frac{S_s J}{r} \tag{6.18}$$

where S_s = unit shearing stress, r = radius of the shaft, and J = moment of inertia of the cross section with respect to its center. (For a round shaft $J = \pi r^4/2$.)

If ϕ is the angle of twist in radians,

$$T = \frac{GJ\phi}{l} \tag{6.19}$$

where G = shearing modulus of elasticity and l = the length of a bar.

The relationship between the torque in a shaft and the power (P) transmitted by it is given by the equation

$$T = \frac{P}{\omega} = \frac{P}{2\pi n} \tag{6.20}$$

where n = number of revolutions per minute on the shaft.

Table 6.3 gives approximate formulas for the maximum shearing stress and angle of twist in members subjected to torsion.

COMBINED STRESSES*

1. Combined Direct and Flexural Stresses. If a bar is subjected to a direct axial loading and also to transverse loading, the stress at any point is given by the algebraic sum of the direct stress P/A and the bending stress My/I. The maximum and minimum stresses at the outer fibers are given by the equation
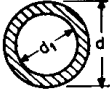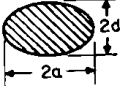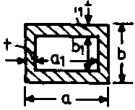
$$S = \frac{P}{A} \pm \frac{Mc}{I} \tag{6.21}$$

in which c is the distance from the neutral axis to the extreme fiber in question.

If a load is applied to a bar in a direction parallel to the axis but at a distance e from it, the single load produces the double effect of direct stress and flexural

---

*Adapted from *General Engineering Handbook,* 2d ed., by C. E. O'Rourke. Copyright © 1940. Used by permission of McGraw-Hill, Inc. All rights reserved.

TABLE 6.3 Approximate Formulas for Maximum Shearing Stress and Angle of Twist in Members Subjected to Torsion

Shape	Maximum unit stress S_s	Angle of twist ϕ
	$\dfrac{16T}{\pi d^2}$	$\dfrac{Tl}{GJ}$
	$\dfrac{16Td}{\pi(d^4 - d_1^4)}$	$\dfrac{32Tl}{\pi(d^4 - d_1^4)G}$
	$\dfrac{2T}{\pi ab^2}$	$\dfrac{T(a^2 + b^2)l}{\pi a^3 b^3 G}$
	$\dfrac{20T}{b^3}$	$\dfrac{46.2Tl}{b^4 G}$
	$\dfrac{T}{2t(a - t)(b - t_1)}$	$\dfrac{Tl(at + bt_1 - t^2 - t_1^2)}{2tt_1(a - t)^2(b - t_1)^2 G}$

Source: From Avallone and Baumeister.[3]

stress. The flexural effect is caused by the moment *Pe,* and the maximum and minimum stresses are given by the equation

$$S = \frac{P}{A} \pm \frac{Pec}{I} \tag{6.22}$$

In the upper chord of a bridge, the compressive load may be applied with enough eccentricity so as just to balance the flexural effect of the weight of the member at the middle point.

2. Combined Shearing and Flexural Stresses. In a beam the direct shearing stress, the induced shearing stress, and the direct flexural stress at any point combine to cause a resultant shearing stress on some diagonal plane which is larger than the direct shearing stress, and a resultant tensile or compressive stress along some diagonal plane which is larger than the direct flexural stress at that point. Let S_s' represent this maximum diagonal unit shearing stress, and let S_t' and S_c', represent the maximum diagonal tensile and compressive unit stresses, respectively. For any point in the beam where the direct flexural stress is tension,

$$S_s' = \sqrt{\left(\frac{S_t}{2}\right)^2 + S_s^2} \quad \text{and} \quad S_t' = \frac{S_t}{2} + S_s' \tag{6.23}$$

For any point in the beam where the direct flexural stress is compression, S_c and

S'_c replace S_t and S'_t, respectively. In a beam the maximum values of S_s and S_t do not occur at the same place.

3. Combined Torsional and Flexural Stresses.   In a shaft which is subjected to both torsion and bending, the maximum values of S_s and S_t (or S_c) occur at the same point; so the values of S'_s and S'_t are greater than the maximum values of S_s and S_t, respectively.

CYLINDERS AND PLATES

1. Thin Cylinder under Internal Pressure.   A cylinder is regarded as thin if $t/d \simeq 0$, i.e., when the thickness of the wall t is small compared with the diameter d. Assuming that the tensile stress across a longitudinal section (Fig. 6.12a) is uniformly distributed over the thickness of the wall,

$$pdl = 2Stl$$

or

$$S = \frac{pd}{2t} \tag{6.24}$$

For tensile stress across a transverse section (Fig. 6.12b)

$$p \frac{\pi d^2}{4} = S_l t \pi d$$

or

$$S_l = \frac{pd}{4t} \tag{6.25}$$

The last equation applies also to the stresses in the walls of a thin *hollow sphere, hemisphere* or *dome*. Here, p = internal pressure, l = length of cylinder, t =

(a)   (b)
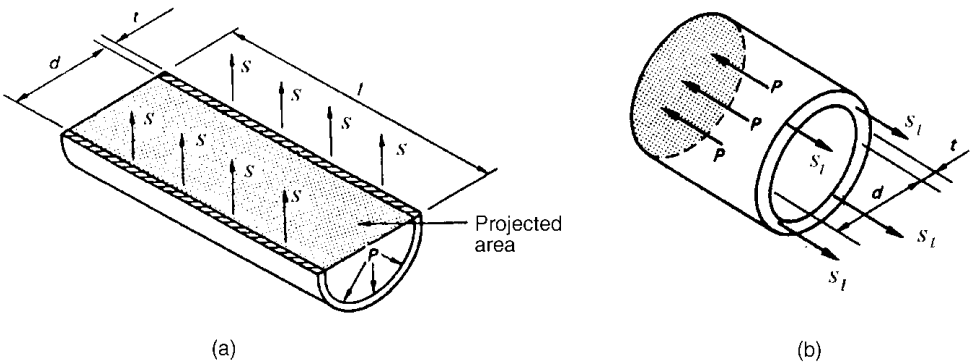
FIGURE 6.12

thickness of wall, d = diameter of cylinder, and S = tensile stress, S_l = longitudinal stress.

2. Circular and Elliptical Flat Plates.   A relation for the maximum stress at the center for a circular flat plate of radius r, uniformly loaded, edge simply supported, is

$$S = \frac{3}{8} \frac{wr^2}{t^2} (3 + \mu) \tag{6.26}$$

and the maximum deflection at the center f is

$$f = \frac{2wr^4}{3Et^3} \tag{6.27}$$

An approximate formula for the maximum stress in elliptical plates, simply supported at the edge (major axis $2a$, minor axis $2b$), is

$$S = \frac{(3a - 2b)}{a} \frac{wb^2}{t^2} \tag{6.28}$$

3. Rectangular and Square Plates.   For a distributed load w, with supports along the four sides, the unit stress is

$$S = \frac{a^2 b^2 w}{2t(a^2 + b^2)} \tag{6.29}$$

where a = long side and b = short side. If $a = b$ (square plate),

$$S = \frac{wa^2}{4t^2} \tag{6.30}$$

In the above equations, w is the uniformly distributed load per unit area and t the thickness of the wall (plate).

NOMENCLATURE

Symbol = definition, SI units (U.S. Customary units)
   A = cross-sectional area, m² (ft²)
   a = distance, m (ft)
   b = distance, m (ft)
   c = distance, m (ft)
   d = diameter, m (ft)
   E = modulus of elasticity, N/m², kg$_f$/cm² (lb$_f$/in²)
   E_s = shearing modulus of elasticity, N/m², kg$_f$/cm² (lb$_f$/in²)
   e = distance. m (ft)
   F = axial force, N (lb$_f$)

f = deflection, m (ft)

G = shearing modulus of elasticity, N/m², kg$_f$/cm² (lb$_f$/in²)

g = acceleration of gravity, m/s² (ft/s²)

I = moment of inertia, m⁴ (ft⁴)

J = moment of inertia of the cross section with respect to its center, m⁴ (ft⁴)

K = bulk modulus of elasticity, N/m², kg$_f$/cm² (lb$_f$/in²)

l = distance, m (ft)

M = bending moment, N · m (lb$_f$ · in)

M_x = local value of bending moment, N · m (lb$_f$ · in)

P = load, N (lb$_f$)

P = axial force, N (lb$_f$)

P = axial load [Eq. (6.17)], N (lb$_f$)

P = power (Eq. (6.20)], W, N · m/s (lb$_f$ · ft/s)

p = pressure, N/m² (lb$_f$/in²)

R = reaction, N (lb$_f$)

r = radius, m (ft)

r = radius of gyration, m (ft)

S = unit stress, N/m², kg$_f$/cm² (lb$_f$/in²)

S_c = unit stress—compression, N/m², kg$_f$/cm² (lb$_f$/in²)

S_s = unit stress—shear, N/m², kg$_f$/cm² (lb$_f$/in²)

S_t = unit stress—tension, N/m², kg$_f$/cm² (lb$_f$/in²)

T = torque (twisting moment), N · m (lb$_f$ · in)

t = wall thickness, m (ft)

U = reliance, N · m, kg$_f$ · cm (lb$_f$ · in)

V = vertical shear [Eq. (6.12)], N, kg$_f$ (lb$_f$)

W = weight, N, kg$_f$ (lb$_f$)

W' = concentrated load, N, kg$_f$ (lb$_f$)

w = load per unit length, N/m, kg$_f$/cm (lb$_f$/in)

w = uniformly distributed load per unit area, N/m² (lb$_f$/in²)

x = distance, m (ft)

y = distance, m (ft)

Greek

α = linear coefficient of thermal expansion, 1/°C (1/°F)

ϕ = angle of twist, rad (deg)

μ = Poisson's ratio (lateral strain/longitudinal strain)

ω = frequency of oscillation, 1/s

ω_n = frequency of natural oscillation, 1/s

δ = strain (unit deformation)

δ_s = shearing strain

Subscripts

max = maximum
1 = position at 1
2 = position at 2

*REFERENCES**

1. C. E. O'Rourke, *General Engineering Handbook,* 2d ed., McGraw-Hill, New York, 1940.

2. Z. D. Jastrzebski, *The Nature and Properties of Engineering Materials,* John Wiley & Sons, New York, 1977.

3. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers,* 9th ed. McGraw-Hill, New York, 1987.

4. N. H. Cook, "Mechanics of Solids," Course 2.01, MIT class notes, 1973–1976.

5. J. P. Den Hartog, *Mechanics,* Dover, New York, 1961.

6. O. W. Eshbach and M. Souders, *Handbook of Engineering Fundamentals,* 3d ed., John Wiley & Sons, New York, 1975.

7. W. Flügge, *Handbook of Engineering Mechanics,* McGraw-Hill, New York, 1962.

8. D. M. Parks, "Solid Mechanics," Course 2.073, MIT class notes, 1973–1976.

9. S. Timoshenko and D. H. Young, *Engineering Mechanics,* 4th ed., McGraw-Hill, New York, 1956.

10. J. H. Williams, Jr., "Applied Elasticity," Course 2.083, MIT class notes, 1973–1976.

11. R. Timings and T. May, *Newnes Mechanical Engineer's Pocket Book,* 2d ed., Newnes, 1997.

12. F. S. Merritt, M. K. Lofin, J. T. Ricketts, *Standard Handbook for Civil Engineers,* 4th ed., McGraw-Hill, 1996.

13. H. Rothbart, *Mechanical Design and Systems Handbook,* 2d ed., McGraw-Hill, 1985.

14. A. Y Ishlinsky, *Classical Mechanics,* Nauka, Moscow, 1987.

---

*Those references listed above but not cited in the text were used for comparison between different data sources, clarification, clarity of presentation, and, most importantly, reader's convenience when further interest in the subject exists.

CHAPTER 7
THERMODYNAMICS

INTRODUCTION

1. Definitions

Thermodynamics. The branch of science that embodies the principles and restrictions of energy transformation in macroscopic systems.

System and Surroundings. A system is taken to be any quantity of matter under consideration (or any object or region associated with the quantity of matter under consideration) selected for study and set apart (imaginary) from everything else, the latter then being called the *surroundings* (Fig. 7.1).

Boundary. The imaginary envelope which encloses the system and separates it from its surroundings is called the *boundary of the system.* With a *closed system* there is no interchange of matter through the boundary between the system and its surroundings. With an *open system* there is such an interchange (Fig. 7.2). An *isolated system* can exchange neither matter nor energy with its surroundings.

Processes. Any change that the system may undergo is known as a *process.* Engineering thermodynamics considers chiefly those processes in which energy transformation occurs by means of changes in the physical state of fluids. The processes are classified into either *reversible* or *irreversible.* Another classification is *nonflow* and *steady flow* processes.

A reversible process is one in which both the system and the surroundings may be returned to their original states. With an irreversible process, this is not possible. All actual processes are irreversible. Among the conditions which contribute to the irreversibility of a process are the following: heat flow from a higher to a lower temperature, mixing of fluids at different temperatures, fluid turbulence, fluid or solid friction, inelastic deformation, etc.

Nonflow processes are those occurring in a container or a space in such a way that the fluid does not flow in or out of the container or space during the process (closed system). An example is the expansion of steam in a cylinder during the period when the valves are closed.

Steady-flow processes are those in which the fluid passes continuously through a region in a steady flow (open system). The steady-flow process or a process which closely approximates steady flow exists in most of the devices and machines employed in engineering practice. Examples are steam engines, turbines, condensers, pumps, boilers, nozzles, valves, and most heat-exchange appliances.

Other common types of processes are defined as follows:

Constant-pressure process, in which the pressure of the fluid is constant throughout the process.
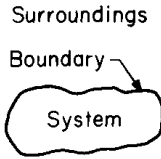
FIGURE 7.1
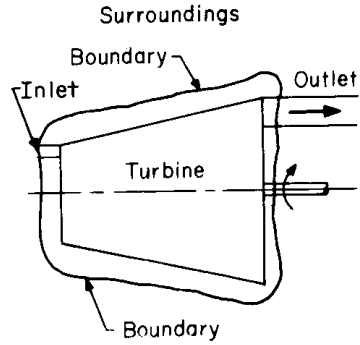
FIGURE 7.2
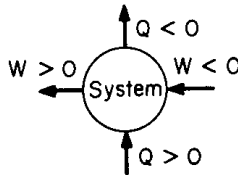
FIGURE 7.3

Constant-volume process, in which the volume of the fluid is constant throughout the process.

Isothermal process, in which the temperature of the fluid is constant throughout the process.

Adiabitic process, in which no heat is added to or removed from the fluid during the process.

Heat. The energy in transit through the system boundary under the influence of a temperature difference or gradient. Its symbol is Q. The quantity of heat is not a property of the system.

Work. Also energy in transit between a system and its surroundings, but resulting from the displacement of an external force acting on the system. Its symbol is W. Like heat, it is not a property of the system. The usual convention with respect to signs for Q and W are shown in Fig. 7.3.

Properties of a System. A system has an identifiable, reproducible state when all its properties are fixed. Properties include the system's internal energy (symbol U), its entropy (symbol S), and its volume (symbol V), all of which are *extensive,* or dependent on the system's size. Other properties, such as temperature (symbol T), pressure (symbol p), entropy per unit mass (symbol s), and internal energy per unit mass (symbol u), are *intensive,* or independent of the system's size.

Properties of a Process. When a system is displaced from an equilibrium state, it undergoes a process during which its properties change until a new equilibrium state is reached. During such a process, the system may interact with its surroundings so as to interchange energy in the forms of heat and work. The amount of energy interchange that occurs during the process is dependent on the specific path

along which a process proceeds. Therefore, heat and work are properties of the process, i.e., Q and W are path functions.

FIRST LAW OF THERMODYNAMICS*

1. Law of Conservation of Energy. The first law of thermodynamics is the law of the conservation of energy: During any process, the total energy of any system and its surroundings is conserved.

For a closed (nonflow) system, the first law is expressed as:

$$Q - W = U_2 - U_1 \tag{7.1}$$

where Q = heat supplied to the system, W = work output, and $U_2 - U_1$ = gain in internal energy.

For an open system (steady-flow process), where the fluid flow rate through a machine or piece of apparatus is constant, the first law is expressed as

$$\dot{Q} - \dot{W} = \dot{m} \left[(h_2 - h_1) + \frac{V_2^2 - V_1^2}{2} + g(z_2 - z_1) \right] \tag{7.2}$$

where $\dot{m}$ = mass flow rate.

For many processes, the last two terms in Eq. (7.2) are often negligible.

Work done on overcoming a fluid pressure is expressed as:

$$W = \int p \, dv \tag{7.3}$$

where p is the pressure effectively applied to the surroundings for doing work and dv represents the change in volume of the system.

For steady-flow processes involving only mechanical effects, the total work done by or on a unit amount of fluid is made up of that done on the two diaphragms (flow cross sections):

$$p_2 v_2 - p_1 v_1 \tag{7.4}$$

and that done on the rest of the surroundings:

$$\int p \, dv - (p_2 v_2 - p_1 v_1) \tag{7.5}$$

By differentiation,

$$p \, dv - d(pv) = -v \, dp \tag{7.6}$$

The net useful work done on the surroundings (often called the *shaft work*) is expressed as:

$$-\int v \, dp \tag{7.7}$$

The net useful or shaft work differs from the total work by $p_2 v_2 - p_1 v_1$.

2. The Joule-Thomson Coefficient. If a fluid is passed adiabatically through a conduit, without doing any net or useful work,

$$h_2 = h_1 \tag{7.8}$$

as velocity and potential effects are negligible. A process of this type is called *Joule-Thomson flow* and the ratio

$$\frac{\partial T}{\partial p} \tag{7.9}$$

for such a flow is the Joule-Thomson coefficient. If a fluid is passed through a nonadiabatic conduit without doing any net or useful work,

$$q = h_2 - h_1 \tag{7.10}$$

assuming $V_2 = V_1$, $z_2 = z_1$. This equation is used to calculate heat balances on different types of flow apparatus (condensers, coolers, and other types of heat exchangers).

THE SECOND LAW OF THERMODYNAMICS

The second law of thermodynamics is related to the limitation of energy conversion by the temperature at which the conversion occurs. It may be shown that the efficiency of all reversible cycles absorbing heat (Q_1) only at a single constant higher temperature T_1 and rejecting heat (Q_2) only at a single constant lower temperature T_2 must be the same. For such cycles

$$\eta = \frac{W}{Q_1} = \frac{T_1 - T_2}{T_1} \tag{7.11}$$

This is called the *Carnot cycle efficiency*.
   Also

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0 \tag{7.12}$$

1. Entropy. The second law of thermodynamics (for open or closed systems) is also expressed as

$$Q_{1-2} = \int_1^2 T \, dS \tag{7.13}$$

where the property S ($S = ms$) is the entropy. (The entropy is a point function, meaning that the change in its value for processes depends only on the end points of the process and not at all upon the particular path taken by the process between the end points.)

Also for any reversible process,

$$\Delta S = S_2 - S_1 = \int_1^2 \frac{dQ_{rev}}{T} \tag{7.14}$$

For any reversible process, the change in entropy of the system and surroundings is zero, whereas for any irreversible process, the net entropy change is positive.

2. Other Relations of General Principles of Thermodynamics

Specific Heat. Specific heat at constant volume:

$$c_v = \left. \frac{\partial u}{\partial T} \right|_v \tag{7.15}$$

Specific heat at constant pressure:

$$c_p = \left. \frac{\partial h}{\partial T} \right|_p \tag{7.16}$$

Mean specific heats for constant pressure and constant volume, respectively, in the temperature range between t_1 and t_2:

$$\bar{c}_p = \bar{c}_p \left|_{t_1}^{t_2}\right. = \frac{\bar{c}_p \left|_0^{t_2}\right. \cdot t_2 - \bar{c}_p \left|_0^{t_1}\right. \cdot t_1}{t_2 - t_1} \tag{7.17}$$

$$\bar{c}_v = \bar{c}_v \left|_{t_1}^{t_2}\right. = \bar{c}_p \left|_{t_1}^{t_2}\right. - R \tag{7.18}$$

Enthalpy

$$h = u + pv \tag{7.19}$$

Free Energy (*Helmholtz Function*)

$$f = u - Ts \tag{7.20}$$

Free Enthalpy (*Gibbs Function*)

$$g = h - Ts \tag{7.21}$$

Availability of System

$$g_0 = h - T_0 s \qquad (7.22)$$

If velocity and potential are not negligible,

$$g_0 = h - T_0 s + \frac{V^2}{2} + gz \qquad (7.23)$$

where T_0 = lowest temperature available for heat discard.

Note. The availability function g_0 is useful for determining thermodynamic efficiencies of a turbine or similar devices, i.e., the ratio of actual work performed during a process to that which theoretically should have been performed.

Maxwell Relations. See Table 7.1.

Other Relations. There are also equations of state:

$$\left. \frac{\partial u}{\partial v} \right|_T = -\left[p - T \left(\frac{\partial p}{\partial T} \right)_v \right] \qquad (7.24)$$

$$\left. \frac{\partial h}{\partial p} \right|_T = \left[v - T \left(\frac{\partial v}{\partial T} \right)_p \right] \qquad (7.25)$$

$$T\, ds = du + pv \qquad (7.26)$$

Note. The last equation expresses the joint statement of first and second laws.

IDEAL GASES*

1. Equation of state. The defining equation of state of an ideal gas is

$$pv = RT \qquad (7.27)$$

or

$$pV = mRT \qquad (7.28)$$

or

$$p = \rho RT \qquad (7.29)$$

If V_m is the volume of one molecular weight of gas,

TABLE 7.1   Maxwell Relations

Function	Differential	Maxwell relation
$\Delta u = q + W$	$du = T\,ds - p\,dv$	$\left(\dfrac{\partial T}{\partial v}\right)_s = -\left(\dfrac{\partial p}{\partial s}\right)_v$
$h = u + pv$	$dh = T\,ds + v\,dp$	$\left(\dfrac{\partial T}{\partial p}\right)_s = \left(\dfrac{\partial v}{\partial s}\right)_p$
$f = u - Ts$	$df = -s\,dT - p\,dv$	$\left(\dfrac{\partial s}{\partial v}\right)_T = \left(\dfrac{\partial p}{\partial T}\right)_v$
$g = h - Ts$	$dg = -s\,dT + v\,dp$	$\left(\dfrac{\partial s}{\partial p}\right)_T = -\left(\dfrac{\partial v}{\partial T}\right)_p$

By holding certain variables constant, a second set of relations is obtained:

Differential	Independent variable held constant	Relation
$du = T\,ds - p\,dv$	s	$\left(\dfrac{\partial u}{\partial v}\right)_s = -p$
	v	$\left(\dfrac{\partial u}{\partial s}\right)_v = T$
$dh = T\,ds + v\,dp$	s	$\left(\dfrac{\partial h}{\partial p}\right)_s = v$
	p	$\left(\dfrac{\partial h}{\partial s}\right)_p = T$
$df = -sdT - p\,dv$	T	$\left(\dfrac{\partial f}{\partial v}\right)_T = -p$
	v	$\left(\dfrac{\partial f}{\partial T}\right)_v = -s$
$dg = -s\,dT + v\,dp$	T	$\left(\dfrac{\partial g}{\partial p}\right)_T = v$
	p	$\left(\dfrac{\partial g}{\partial T}\right)_p = -s$

Source: From Avallone and Baumeister.[4]

$$pV_m = R_m T \qquad (7.30)$$

where
$$R_m = M \cdot R$$
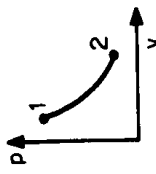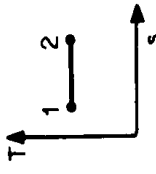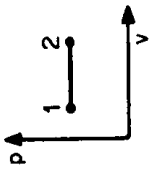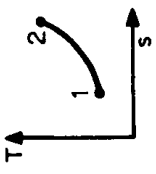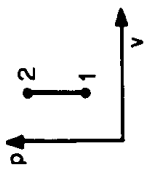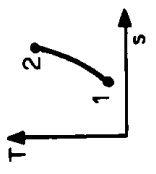
and $R_m = 8314.3$ J/(kmol · K) is the universal gas constant and is the same for all ideal gases in any chosen system of units.

Each change of state of an ideal gas may be represented by the equation

$$p\,v^n = \text{const} \qquad (7.31)$$

See Table 7.2 for various values of n.

TABLE 7.2 Relations for Ideal-Gas Processes

Process	p, v, T relations between states 1 and 2	$w_{1\text{-}2}$, work per unit mass (closed system)—Eq. (7.3); reversible	$w_{1\text{-}2}$, work per unit mass (open system)—Eq. (7.7); reversible	$q_{1\text{-}2}$, heat transferred per unit mass	p-v Diagram	T-S diagram
Isothermal $T = \text{const}$ $n = 1$	$\dfrac{p_2}{p_1} = \dfrac{v_1}{v_2}$	$RT \ln \dfrac{v_2}{v_1}$ $= RT \ln \dfrac{p_1}{p_2}$ $= p_1 v_1 \ln \dfrac{v_2}{v_1}$	$w_{1\text{-}2}$	$w_{1\text{-}2}$		
Isobaric $p = \text{const}$ $n = 0$	$\dfrac{v_2}{v_1} = \dfrac{T_2}{T_1}$	$p(v_2 - v_1) = R(T_2 - T_1)$	0	$\bar{c}_p(T_2 - T_1)$		
Isochoric $v = \text{const}$ $n = \infty$	$\dfrac{p_2}{p_1} = \dfrac{T_2}{T_1}$	0	$v(p_1 - p_2) = R(T_1 - T_2)$	$\bar{c}_v(T_2 - T_1)$		

		$u_1 - u_2$	$h_1 - h_2$	
Isentropic $s = \text{const}$ $n = k$	$\dfrac{p_2}{p_1} = \left(\dfrac{v_1}{v_2}\right)^k$ $\dfrac{v_2}{v_1} = \left(\dfrac{T_1}{T_2}\right)^{1/(k-1)}$	$\begin{aligned} u_1 - u_2 &= \bar{c}_v(T_1 - T_2) \\ &= -\bar{c}_v T_1 \\ &\quad \times \left[\left(\dfrac{p_2}{p_1}\right)^{(k-1)/k} - 1\right] \\ &= -\dfrac{1}{k-1}RT_1 \\ &\quad \times \left[\left(\dfrac{p_2}{p_1}\right)^{(k-1)/k} - 1\right] \end{aligned}$	$\begin{aligned} h_1 - h_2 &= \bar{c}_p(T_1 - T_2) \\ &= -\bar{c}_p T_1 \\ &\quad \times \left[\left(\dfrac{p_2}{p_1}\right)^{(k-1)/k} - 1\right] \\ &= -\dfrac{1}{k-1}RT_1 \\ &\quad \times \left[\left(\dfrac{p_2}{p_1}\right)^{(k-1)/k} - 1\right] \end{aligned}$	0
Politropic $n = \text{const}$	$\dfrac{p_2}{p_1} = \left(\dfrac{v_1}{v_2}\right)^n$ $\dfrac{p_2}{p_1} = \left(\dfrac{T_2}{T_1}\right)^{n/(n-1)}$ $\dfrac{v_2}{v_1} = \left(\dfrac{T_1}{T_2}\right)^{1/(n-1)}$	$\begin{aligned} &\dfrac{1}{n-1}R(T_1 - T_2) \\ &= \dfrac{1}{n-1}(p_1 v_1 - p_2 v_2) \\ &= -\dfrac{1}{n-1}RT \\ &\quad \times \left[\left(\dfrac{p_2}{p_1}\right)^{(n-1)/n} - 1\right] \end{aligned}$	$\begin{aligned} &\dfrac{n}{n-1}R(T_1 - T_2) \\ &= -\dfrac{n}{n-1}RT_1 \\ &\quad \times \left[\left(\dfrac{p_2}{p_1}\right)^{(n-1)/n} - 1\right] \end{aligned}$	$\bar{c}_v \dfrac{n-k}{n-1}(T_2 - T_1)$

Source: From Perry,[2] Avallone and Baumeister,[4] and Gieck.[5]

2. Changes of State of Ideal Gases.   Using the symbols for initial state (p_1, v_1, T_1) and final state (p_2, v_2, T_2), the equation for *internal energy* reads as:

$$u_2 - u_1 = \bar{c}_v(T_2 - T_1) = \frac{p_2 v_2 - p_1 v_1}{k - 1} \tag{7.32}$$

The equation for *enthalpy* reads as:

$$h_2 - h_1 = \bar{c}_p(T_2 - T_1) = \frac{k(p_2 v_2 - p_1 v_1)}{k - 1} \tag{7.33}$$

And the equation for *entropy* reads as:

$$s_2 - s_1 = \bar{c}_v \ln \frac{T_2}{T_1} + R \ln \frac{v_2}{v_1} = \bar{c}_p \ln \frac{T_2}{T_1} - R \ln \frac{p_2}{p_1} = \bar{c}_p \ln \frac{v_2}{v_1} + \bar{c}_v \ln \frac{p_2}{p_1} \tag{7.34}$$

Note that

$$k = \frac{c_p}{c_v} \tag{7.35}$$

3. Processes.   Table 7.2 includes the equations applicable for determining the relationships between states, work, and heat for the simplest processes where an ideal gas is the medium and in which conditions throughout the process are idealized.

   Graphical representation of the change of state of a substance may be done by taking any two of the six variables p, v, T, S, U, and H as independent coordinates. While any pair may be chosen, there are three systems of graphical representation that are widely used:

• The pressure-volume (p-v) diagram
• The T-S diagram
• The H-S diagram (Mollier diagram)

   Figures 7.14 through 7.16 include the diagrams for water.

4. Mixtures of Gases.   There are a number of equations for mixtures, depending on the property of the mixture involved:
   Mass m of a mixture of components m_1, m_2, . . . :

$$m = m_1 + m_2 + \cdots + m_n = \sum_{i=1}^{n} m_i \tag{7.36}$$

   Mass fraction x_i of a mixture:

$$x_i = \frac{m_i}{m} \quad \text{and} \quad \sum_{i=1}^{n} x_i = 1 \tag{7.37}$$

Mole fraction y_i of a mixture:

$$y_i = \frac{n_i}{n} \quad \text{and} \quad \sum_{i=1}^{n} y_i = 1 \tag{7.38}$$

$$n = n_1 + n_2 + \cdots + n_n = \sum_{i=1}^{n} n_i \tag{7.39}$$

Equivalent molecular mass M of a mixture:

$$M_i = \frac{m_i}{n} \quad \text{and} \quad M = \frac{m}{n} \tag{7.40}$$

$$M = \sum_{i=1}^{n} M_i \cdot y_i \quad \text{and} \quad \frac{1}{M} = \sum_{i=1}^{n} \frac{x_i}{M_i} \tag{7.41}$$

and

$$x_i = \frac{M_i}{M} y_i \tag{7.42}$$

Pressure p of the mixture and the partial pressure p_i of its components:

$$p = \sum_{i=1}^{n} p_i \qquad p_i = y_i \cdot p \tag{7.43}$$

Volume fraction ξ_i of a mixture:

$$\xi_i = \frac{\mathbf{V}_i}{\mathbf{V}} = y_i \quad \text{and} \quad \sum_{i=1}^{n} \xi_i = 1 \tag{7.44}$$

Partial volume $\mathbf{V}_i$:

$$\mathbf{V}_i = \frac{m_i R_i T}{p} = \frac{n_i R_m T}{p} \tag{7.45}$$

$$\sum_{i=1}^{n} \mathbf{V}_i = \mathbf{V} \tag{7.46}$$

Internal energy and enthalpy of a mixture:

$$u = \sum_{i=1}^{n} x_i u_i \tag{7.47}$$

$$h = \sum_{i=1}^{n} x_i h_i \tag{7.48}$$

In this relation the temperature of the mixture is given as: For the calculation of u:

$$t = \frac{\overline{c}_{v_1} m_1 t_1 + \overline{c}_{v_2} m_2 t_2 + \cdots + \overline{c}_{v_n} m_n t_n}{\overline{c}_v \cdot m} \tag{7.49}$$

For calculation of h:

$$t = \frac{\bar{c}_{p_1} m_1 t_1 + \bar{c}_{p_2} m_2 t_2 + \cdots + \bar{c}_{p_n} m_n t_n}{\bar{c}_p \cdot m} \tag{7.50}$$

For the mixture:

$$\bar{c}_v = \bar{c}_p - R \tag{7.51}$$

where

$$\bar{c}_p = \sum_{i=1}^{n} x_i \bar{c}_{p_i} \tag{7.52}$$

The *relative humidity* (R.H.) of an ideal-gas mixture is defined as

$$\text{R.H.} = \frac{p_v}{p_{\text{sat}}} \tag{7.53}$$

where p_v is the actual pressure of the vapor phase and p_{sat} is the pressure exerted by the same vapor when saturated at the mixture temperature.

The *absolute humidity* [or *specific humidity* (S.H.)] is the ratio of the mass of vapor in the mixture (m_v) to the mass to the dry gas (m_g):

$$\text{S.H.} = \frac{m_v}{m_g} \tag{7.54}$$

Observations of the concentrations of these two components are usually made in terms of three temperatures:

1. The dew-point temperature (the temperature at which condensation of the vapor takes place if the mixture is cooled at constant pressure)
2. The wet-bulb temperature (the temperature achieved by the mixture if it is saturated by evaporating liquid into it so that the latent heat of vaporization comes from the mixture, thereby depressing its temperature)
3. The dry-bulb temperature (the normal temperature of the mixture)

REAL GASES*

In many cases, especially for gases at high pressures, the ideal-gas law does not adequately represent the relationship between the properties. For that reason it is

---

*This section is based in part on *Engineering Thermodynamics,* 2d ed., by W. C. Reynolds and H. C. Perkins, copyright © 1977; in part on *Engineering Manual,* 3d ed., by R. H. Perry, copyright © 1976; in part on *Perry's Chemical Engineers' Handbook,* 6th ed., by R. H. Perry and D. W. Green, copyright ©1984; and in part on *Marks' Standard Handbook for Mechanical Engineers,* 9th ed., by E. A. Avallone and T. Baumeister III (eds.), copyright © 1987. All material used by permission of McGraw-Hill, Inc. All rights reserved. The section is also based in part on "Heat Transfer Characteristics of Working Fluids for OTEC," by E. N. Ganić, copyright © 1980 by ASME. Used by permission of ASME. All rights reserved.

modified to apply to real gases by the introduction of a correction factor, the compressibility factor C:

$$pV = CnR_mT \tag{7.55}$$

where V = volume of n moles.

Values of C are given in Fig. 7.4 in terms of reduced pressure (p_r) and reduced temperature T_r, where

$$T_r = \frac{T}{T_c} \quad \text{and} \quad p_r = \frac{p}{p_c}$$

Values of critical pressure p_c and critical temperature T_c are shown in Table 7.3.

1. Systems Containing More than One Phase. When several phases exist in equilibrium together, the thermodynamic condition of equilibrium is described in terms of a system property g, or Gibbs function, already given as Eq. (7.21).

In the case where a single component exists in several phases, such as ice and steam, the equilibrium conditions are given as

$$g_I = g_{II} = g_{III} \tag{7.56}$$

where g_I represents the specific Gibbs function for phase I. The different conditions of existence are best represented graphically by a phase diagram and a pressure-volume (p-v) diagram, as illustrated in Fig. 7.5.

2. Vapor Pressure. The vapor pressure or saturation pressure at any given temperature can often be obtained from property tables (see Tables 7.4 and 7.5), or may be approximated by the following relation:
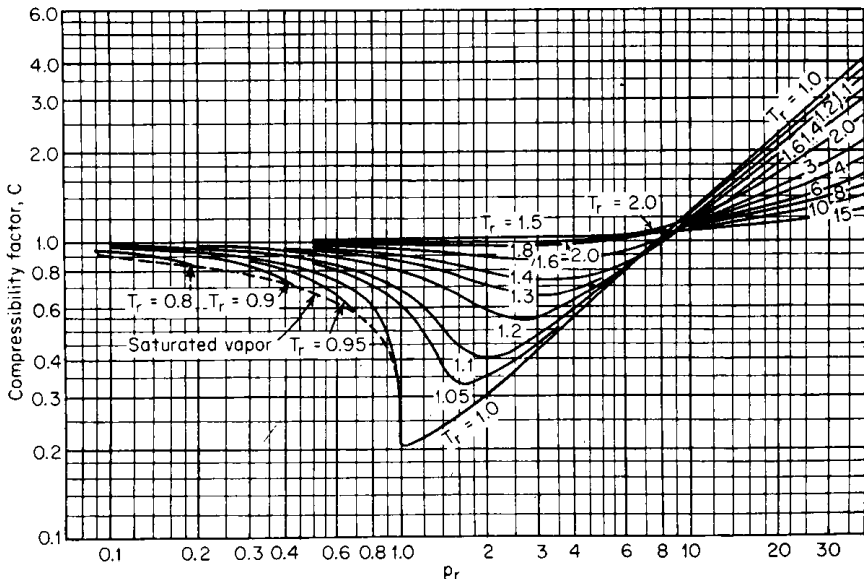
FIGURE 7.4 Compressibility factors of gases. (*From Perry.*[2])

TABLE 7.3 Critical Properties of Gases

Gas	Symbol	Approx. mol wt	Critical pressure,		Critical temperature,	
			psia	(MPa)	°F	(°C)
Acetylene	C_2H_2	26	911	(6.28)	96.3	(35.7)
Air	...	29	546	(3.76)	−220.3	(−140.2)
Ammonia	NH_3	17	1640	(11.30)	270.3	(132.4)
Argon	Ar	40	706	(4.86)	−187.7	(−122)
Benzene	C_6H_6	78	702	(4.84)	551.4	(288.5)
Butane	C_4H_{10}	58	530	(3.65)	307.4	(153)
Carbon dioxide	CO_2	44	1073	(7.39)	88.0	(31.1)
Carbon monoxide	CO	28	515	(3.55)	−220.3	(−140.2)
Dichlorodifluoromethane (R12)	CCl_2F_2	121	597	(4.11)	233.6	(112)
Ethane	C_2H_6	30	718	(4.95)	90.0	(32.2)
Ethylene	C_2H_4	28	748	(5.15)	49.3	(9.6)
Helium	He	4	33	(0.22)	−450.2	(−267.8)
Heptane	C_7H_{16}	100	394	(2.71)	517.1	(269.5)
Hydrogen	H_2	2	188	(1.29)	−399.8	(−239.8)
Methane	CH_4	16	674	(4.64)	−116.5	(−82.5)
Nitrogen	N_2	28	493	(3.39)	−232.8	(−147.1)
Oxygen	O_2	32	731	(5.04)	−181.8	(−118.7)
Propane	C_3H_8	44	632	(4.35)	206.3	(96.8)
Water	H_2O	18	3106	(21.41)	705.5	(374.2)

Source: From Perry.[2]

7.14
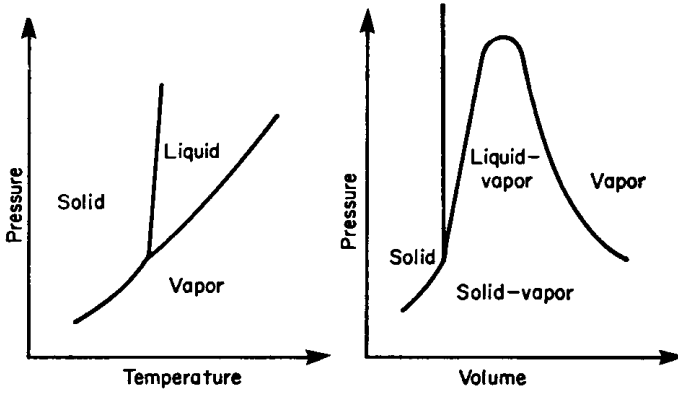
FIGURE 7.5 Phase and pressure-volume diagrams.

$$\ln p = \frac{A}{T} + B \tag{7.57}$$

where A and B are constants whose values depend on the substance.

3. Clapeyron Equation. The Clapeyron equation is an important relationship useful in calculations relating to the constant-pressure evaporation of pure substances. In such cases the equation may be written as

$$v_{fg} = \frac{h_{fg}}{T} \frac{1}{dp/dT} \tag{7.58}$$

4. Properties of Mixtures of Liquid and Vapor. The properties of a unit mass of a mixture of liquid and vapor of quality x (mass fraction) are given by the following expressions:

$$v = v_f + x \cdot v_{fg} \tag{7.59}$$

$$h = h_f + x \cdot h_{fg} \tag{7.60}$$

$$u = u_f + x \cdot u_{fg} \tag{7.61}$$

$$s = s_f + x \cdot s_{fg} \tag{7.62}$$

where $v_{fg} = v_g - v_f$ = increase of volume during vaporization

$h_{fg} = h_g - h_f$ = heat of vaporization (heat required to vaporize unit mass of liquid at constant pressure and temperature)

$u_{fg} = u_g - u_f$ = increase of internal energy during vaporization

$s_{fg} = s_g - s_f = \dfrac{h_{fg}}{T}$ = increase of entropy during vaporization

$p\,v_{fg}$ = work performed during vaporization

Note that the subscripts g and f are related to the gas (vapor) and liquid states, respectively. Tables 7.4 through 7.6 and Figs. 7.14 through 7.16 are useful for calculating these properties.

TABLE 7.4 Properties of Saturated Water and Water Vapor-Temperature Table

P, MPa	T, °C	Volume, m³/kg		Internal energy, kJ/kg		Enthalpy, kJ/kg			Entropy, kJ/(kg · K)		
		v_f	v_g	u_f	u_g	h_f	h_{fg}	h_g	s_f	s_{fg}	s_g
0.000611	0.01	0.001000	206.1	0.0	2375.3	0.0	2501.3	2501.3	0.0000	9.1571	9.1571
0.0008	3.8	0.001000	159.7	15.8	2380.5	15.8	2492.5	2508.3	0.0575	9.0007	9.0582
0.001	7.0	0.001000	129.2	29.3	2385.0	29.3	2484.9	2514.2	0.1059	8.8706	8.9765
0.0012	9.7	0.001000	108.7	40.6	2388.7	40.6	2478.5	2519.1	0.1460	8.7639	8.9099
0.0014	12.0	0.001001	93.92	50.3	2391.9	50.3	2473.1	2523.4	0.1802	8.6736	8.8538
0.0016	14.0	0.001001	82.76	58.9	2394.7	58.9	2468.2	2527.1	0.2101	8.5952	8.8053
0.0018	15.8	0.001001	74.03	66.5	2397.2	66.5	2464.0	2530.5	0.2367	8.5259	8.7626
0.002	17.5	0.001001	67.00	73.5	2399.5	73.5	2460.0	2533.5	0.2606	8.4639	8.7245
0.003	24.1	0.001003	45.67	101.0	2408.5	101.0	2444.5	2545.5	0.3544	8.2240	8.5784
0.004	29.0	0.001004	34.80	121.4	2415.2	121.4	2433.0	2554.4	0.4225	8.0529	8.4754
0.006	36.2	0.001006	23.74	151.5	2424.9	151.5	2415.9	2567.4	0.5208	7.8104	8.3312
0.008	41.5	0.001008	18.10	173.9	2432.1	173.9	2403.1	2577.0	0.5924	7.6371	8.2295
0.01	45.8	0.001010	14.67	191.8	2437.9	191.8	2392.8	2584.6	0.6491	7.5019	8.1510
0.012	49.4	0.001012	12.36	206.9	2442.7	206.9	2384.1	2591.0	0.6961	7.3910	8.0871
0.014	52.6	0.001013	10.69	220.0	2446.9	220.0	2376.6	2596.6	0.7365	7.2968	8.0333
0.016	55.3	0.001015	9.433	231.5	2450.5	231.5	2369.9	2601.4	0.7719	7.2149	7.9868
0.018	57.8	0.001016	8.445	241.9	2453.8	241.9	2363.9	2605.8	0.8034	7.1425	7.9459
0.02	60.1	0.001017	7.649	251.4	2456.7	251.4	2358.3	2609.7	0.8319	7.0774	7.9093
0.03	69.1	0.001022	5.229	289.2	2468.4	289.2	2336.1	2625.3	0.9439	6.8256	7.7695
0.04	75.9	0.001026	3.993	317.5	2477.0	317.6	2319.1	2636.7	1.0260	6.6449	7.6709
0.06	85.9	0.001033	2.732	359.8	2489.6	359.8	2293.7	2653.5	1.1455	6.3873	7.5328
0.08	93.5	0.001039	2.087	391.6	2498.8	391.6	2274.1	2665.7	1.2331	6.2023	7.4354
0.1	99.6	0.001043	1.694	417.3	2506.1	417.4	2258.1	2675.5	1.3029	6.0573	7.3602
0.12	104.8	0.001047	1.428	439.2	2512.1	439.3	2244.2	2683.5	1.3611	5.9378	7.2989
0.14	109.3	0.001051	1.237	458.2	2517.3	458.4	2232.0	2690.4	1.4112	5.8360	7.2472
0.16	113.3	0.001054	1.091	475.2	2521.8	475.3	2221.2	2696.5	1.4553	5.7472	7.2025
0.18	116.9	0.001058	0.9775	490.5	2525.9	490.7	2211.1	2701.8	1.4948	5.6683	7.1631
0.2	120.2	0.001061	0.8857	504.5	2529.5	504.7	2201.9	2706.6	1.5305	5.5975	7.1280

0.3	133.5	0.001073	0.6058	561.1	2543.6	561.5	2163.8	2725.3	1.6722	5.3205	6.9927
0.4	143.6	0.001084	0.4625	604.3	2553.6	604.7	2133.8	2738.5	1.7770	5.1197	6.8967
0.6	158.9	0.001101	0.3157	669.9	2567.4	670.6	2086.2	2756.8	1.9316	4.8293	6.7609
0.8	170.4	0.001115	0.2404	720.2	2576.8	721.1	2048.0	2769.1	2.0466	4.6170	6.6636
1	179.9	0.001127	0.1944	761.7	2583.6	762.8	2015.3	2778.1	2.1391	4.4482	6.5873
1.2	188.0	0.001139	0.1633	797.3	2588.8	798.6	1986.2	2784.8	2.2170	4.3072	6.5242
1.4	195.1	0.001149	0.1408	828.7	2592.8	830.3	1959.7	2790.0	2.2847	4.1854	6.4701
1.6	201.4	0.001159	0.1238	856.9	2596.0	858.8	1935.2	2794.0	2.3446	4.0780	6.4226
1.8	207.2	0.001168	0.1104	882.7	2598.4	884.8	1912.3	2797.1	2.3986	3.9816	6.3802
2	212.4	0.001177	0.09963	906.4	2600.3	908.8	1890.7	2799.5	2.4478	3.8939	6.3417
3	233.9	0.001216	0.06668	1004.8	2604.1	1008.4	1795.7	2804.1	2.6462	3.5416	6.1878
4	250.4	0.001252	0.04978	1082.3	2602.3	1087.3	1714.1	2801.4	2.7970	3.2739	6.0709
6	275.6	0.001319	0.03244	1205.4	2589.7	1213.3	1571.0	2784.3	3.0273	2.8627	5.8900
8	295.1	0.001384	0.02352	1305.6	2569.8	1316.6	1441.4	2758.0	3.2075	2.5365	5.7440
9	303.4	0.001418	0.02048	1350.5	2557.8	1363.3	1378.8	2742.1	3.2865	2.3916	5.6781
10	311.1	0.001452	0.01803	1393.0	2544.4	1407.6	1317.1	2724.7	3.3603	2.2546	5.6149
12	324.8	0.001527	0.01426	1472.9	2513.7	1491.3	1193.6	2684.9	3.4970	1.9963	5.4933
14	336.8	0.001611	0.01149	1548.6	2476.8	1571.1	1066.5	2637.6	3.6240	1.7486	5.3726
16	347.4	0.001711	0.009307	1622.7	2431.8	1650.0	930.7	2580.7	3.7468	1.4996	5.2464
18	357.1	0.001840	0.007491	1698.9	2374.4	1732.0	777.2	2509.2	3.8722	1.2332	5.1054
20	365.8	0.002036	0.005836	1785.6	2293.2	1826.3	583.7	2410.0	4.0146	0.9135	4.9281
22.088	374.136	0.003155	0.003155	2029.6	2029.6	2099.3	0.0	2099.3	4.4305	0.0000	4.4305

Source: From Reynolds and Perkins.[1]

7.17

TABLE 7.5 Properties of Saturated Water and Water Vapor-Pressure Table

T, °C	P, MPa	Volume, m³/kg		Internal energy, kJ/kg		Enthalpy, kJ/kg			Entropy, kJ/(kg · K)		
		v_f	v_g	u_f	u_g	h_f	h_{fg}	h_g	s_f	s_{fg}	s_g
0.010	0.0006113	0.001000	206.1	0.0	2375.3	0.0	2501.3	2501.3	0.0000	9.1571	9.1571
2	0.0007056	0.001000	179.9	8.4	2378.1	8.4	2496.6	2505.0	0.0305	9.0738	9.1043
5	0.0008721	0.001000	147.1	21.0	2382.2	21.0	2489.5	2510.5	0.0761	8.9505	9.0266
10	0.001228	0.001000	106.4	42.0	2389.2	42.0	2477.7	2519.7	0.1510	8.7506	8.9016
15	0.001705	0.001001	77.93	63.0	2396.0	63.0	2465.9	2528.9	0.2244	8.5578	8.7822
20	0.002338	0.001002	57.79	83.9	2402.9	83.9	2454.2	2538.1	0.2965	8.3715	8.6680
25	0.003169	0.001003	43.36	104.9	2409.8	104.9	2442.3	2547.2	0.3672	8.1916	8.5588
30	0.004246	0.001004	32.90	125.8	2416.6	125.8	2430.4	2556.2	0.4367	8.0174	8.4541
35	0.005628	0.001006	25.22	146.7	2423.4	146.7	2418.6	2565.3	0.5051	7.8488	8.3539
40	0.007383	0.001008	19.52	167.5	2430.1	167.5	2406.8	2574.3	0.5723	7.6855	8.2578
45	0.009593	0.001010	15.26	188.4	2436.8	188.4	2394.8	2583.2	0.6385	7.5271	8.1656
50	0.01235	0.001012	12.03	209.3	2443.5	209.3	2382.8	2592.1	0.7036	7.3735	8.0771
55	0.01576	0.001015	9.569	230.2	2450.1	230.2	2370.7	2600.9	0.7678	7.2243	7.9921
60	0.01994	0.001017	7.671	251.1	2456.6	251.1	2358.5	2609.6	0.8310	7.0794	7.9104
65	0.02503	0.001020	6.197	272.0	2463.1	272.0	2346.2	2618.2	0.8934	6.9384	7.8318
70	0.03119	0.001023	5.042	292.9	2469.5	293.0	2333.8	2626.8	0.9549	6.8012	7.7561
75	0.03858	0.001026	4.131	313.9	2475.9	313.9	2321.4	2635.3	1.0155	6.6678	7.6833
80	0.04739	0.001029	3.407	334.8	2482.2	334.9	2308.8	2643.7	1.0754	6.5376	7.6130
85	0.05783	0.001032	2.828	355.8	2488.4	355.9	2296.0	2651.9	1.1344	6.4109	7.5453
90	0.07013	0.001036	2.361	376.8	2494.5	376.9	2283.2	2660.1	1.1927	6.2872	7.4799
95	0.08455	0.001040	1.982	397.9	2500.6	397.9	2270.2	2668.1	1.2503	6.1664	7.4167
100	0.1013	0.001044	1.673	418.9	2506.5	419.0	2257.0	2676.0	1.3071	6.0486	7.3557
110	0.1433	0.001052	1.210	461.1	2518.1	461.3	2230.2	2691.5	1.4188	5.8207	7.2395
120	0.1985	0.001060	0.8919	503.5	2529.2	503.7	2202.6	2706.3	1.5280	5.6024	7.1304
130	0.2701	0.001070	0.6685	546.0	2539.9	546.3	2174.2	2720.5	1.6348	5.3929	7.0277

140	0.3613	0.001080	0.5089	588.7	2550.0	589.1	2144.8	2733.9	1.7395	5.1912	6.9307
150	0.4758	0.001090	0.3928	631.7	2559.5	632.2	2114.2	2746.4	1.8422	4.9965	6.8387
160	0.6178	0.001102	0.3071	674.9	2568.4	675.5	2082.6	2758.1	1.9431	4.8079	6.7510
170	0.7916	0.001114	0.2428	718.3	2576.5	719.2	2049.5	2768.7	2.0423	4.6249	6.6672
180	1.002	0.001127	0.1941	762.1	2583.7	763.2	2015.0	2778.2	2.1400	4.4466	6.5866
190	1.254	0.001141	0.1565	806.2	2590.0	807.5	1978.8	2786.4	2.2363	4.2724	6.5087
200	1.554	0.001156	0.1274	850.6	2595.3	852.4	1940.8	2793.2	2.3313	4.1018	6.4331
210	1.906	0.001173	0.1044	895.5	2599.4	897.7	1900.8	2798.5	2.4253	3.9340	6.3593
220	2.318	0.001190	0.08620	940.9	2602.4	943.6	1858.5	2802.1	2.5183	3.7686	6.2869
230	2.795	0.001209	0.07159	986.7	2603.9	990.1	1813.9	2804.0	2.6105	3.6050	6.2155
240	3.344	0.001229	0.05977	1033.2	2604.0	1037.3	1766.5	2803.8	2.7021	3.4425	6.1446
250	3.973	0.001251	0.05013	1080.4	2602.4	1085.3	1716.2	2801.5	2.7933	3.2805	6.0738
260	4.688	0.001276	0.04221	1128.4	2599.0	1134.4	1662.5	2796.9	2.8844	3.1184	6.0028
270	5.498	0.001302	0.03565	1177.3	2593.7	1184.5	1605.2	2789.7	2.9757	2.9553	5.9310
280	6.411	0.001332	0.03017	1227.4	2586.1	1236.0	1543.6	2779.6	3.0674	2.7905	5.8579
290	7.436	0.001366	0.02557	1278.9	2576.0	1289.0	1477.2	2766.2	3.1600	2.6230	5.7830
300	8.580	0.001404	0.02168	1332.0	2563.0	1344.0	1405.0	2749.0	3.2540	2.4513	5.7053
310	9.856	0.001447	0.01835	1387.0	2546.4	1401.3	1326.0	2727.3	3.3500	2.2739	5.6239
320	11.27	0.001499	0.01549	1444.6	2525.5	1461.4	1238.7	2700.1	3.4487	2.0883	5.5370
330	12.84	0.001561	0.01300	1505.2	2499.0	1525.3	1140.6	2665.9	3.5514	1.8911	5.4425
340	14.59	0.001638	0.01080	1570.3	2464.6	1594.2	1027.9	2622.1	3.6601	1.6765	5.3366
350	16.51	0.001740	0.008815	1641.8	2418.5	1670.6	893.4	2564.0	3.7784	1.4338	5.2122
360	18.65	0.001892	0.006947	1725.2	2351.6	1760.5	720.7	2481.2	3.9154	1.1382	5.0536
370	21.03	0.002213	0.004931	1844.0	2229.0	1890.5	442.2	2332.7	4.1114	0.6876	4.7990
374.136	22.088	0.003155	0.003155	2029.6	2029.6	2099.3	0.0	2099.3	4.4305	0.0000	4.4305

Source: From Reynolds and Perkins.[1]

7.19

TABLE 7.6 Properties of Superheated Water Vapor

P, MPa (T_{sat}, °C)		50	100	150	200	250	300	350	400	500	600	700	800	900
														Temperature, °C
0.002 (17.5)	v, m³/kg	74.52	86.08	97.63	109.2	120.7	132.3	143.8	155.3	178.4	201.5	224.6	247.6	270.7
	u, kJ/kg	2445.2	2516.3	2588.3	2661.6	2736.2	2812.2	2889.8	2969.0	3132.3	3302.5	3479.7	3663.9	3855.1
	h, kJ/kg	2594.3	2688.4	2783.6	2879.9	2977.6	3076.7	3177.4	3279.6	3489.1	3705.5	3928.8	4159.1	4396.5
	s, kJ/(kg·K)	8.9227	9.1936	9.4328	9.6479	9.8442	10.0251	10.1935	10.3513	10.6414	10.9044	11.1465	11.3718	11.5832
0.005 (32.9)	v, m³/kg	29.78	34.42	39.04	43.66	48.28	52.90	57.51	62.13	71.36	80.59	89.82	99.05	108.3
	u, kJ/kg	2444.7	2516.0	2588.1	2661.4	2736.1	2812.2	2889.8	2968.9	3132.3	3302.5	3479.6	3663.9	3855.0
	h, kJ/kg	2593.6	2688.1	2783.3	2879.8	2977.5	3076.6	3177.3	3279.6	3489.1	3705.4	3928.8	4159.1	4396.5
	s, kJ/(kg·K)	8.4982	8.7699	9.0095	9.2248	9.4212	9.6022	9.7706	9.9284	10.2185	10.4815	10.7236	10.9489	11.1603
0.01 (45.8)	v, m³/kg	14.87	17.20	19.51	21.83	24.14	26.45	28.75	31.06	35.68	40.29	44.91	49.53	54.14
	u, kJ/kg	2443.9	2515.5	2587.9	2661.3	2736.0	2812.1	2889.7	2968.9	3132.3	3302.5	3479.6	3663.8	3855.0
	h, kJ/kg	2592.6	2687.5	2783.0	2879.5	2977.3	3076.5	3177.2	3279.5	3489.0	3705.4	3928.7	4159.1	4396.4
	s, kJ/(kg·K)	8.1757	8.4487	8.6890	8.9046	9.1010	9.2821	9.4506	9.6084	9.8985	10.1616	10.4037	10.6290	10.8404
0.02 (60.1)	v, m³/kg		8.585	9.748	10.91	12.06	13.22	14.37	15.53	17.84	20.15	22.45	24.76	27.07
	u, kJ/kg		2514.5	2587.3	2660.9	2735.7	2811.9	2889.5	2968.8	3132.2	3302.4	3479.6	3663.8	3855.0
	h, kJ/kg		2686.2	2782.3	2879.1	2977.0	3076.3	3177.0	3279.4	3488.9	3705.3	3928.7	4159.1	4396.4
	s, kJ/(kg·K)		8.1263	8.3678	8.5839	8.7807	8.9619	9.1304	9.2884	9.5785	9.8417	10.0838	10.3091	10.5205
0.05 (81.3)	v, m³/kg		3.418	3.889	4.356	4.820	5.284	5.747	6.209	7.134	8.057	8.981	9.904	10.83
	u, kJ/kg		2511.6	2585.6	2659.8	2735.0	2811.3	2889.1	2968.4	3131.9	3302.2	3479.5	3663.7	3854.9
	h, kJ/kg		2682.5	2780.1	2877.6	2976.0	3075.5	3176.4	3278.9	3488.6	3705.1	3928.5	4158.9	4396.3
	s, kJ/(kg·K)		7.6955	7.9409	8.1588	8.3564	8.5380	8.7069	8.8650	9.1554	9.4186	9.6608	9.8861	10.0975
0.07 (89.9)	v, m³/kg		2.434	2.773	3.108	3.441	3.772	4.103	4.434	5.095	5.755	6.415	7.074	7.734
	u, kJ/kg		2509.6	2584.5	2659.1	2734.5	2811.0	2888.8	2968.2	3131.8	3302.1	3479.4	3663.6	3854.9
	h, kJ/kg		2680.0	2778.6	2876.7	2975.3	3075.0	3176.1	3278.6	3488.4	3704.9	3928.4	4158.8	4396.2
	s, kJ/(kg·K)		7.5349	7.7829	8.0020	8.2001	8.3821	8.5511	8.7094	8.9999	9.2632	9.5054	9.7307	9.9422
0.1 (99.6)	v, m³/kg		1.696	1.936	2.172	2.406	2.639	2.871	3.103	3.565	4.028	4.490	4.952	5.414
	u, kJ/kg		2506.6	2582.7	2658.0	2733.7	2810.4	2888.4	2967.8	3131.5	3301.9	3479.2	3663.5	3854.8
	h, kJ/kg		2676.2	2776.4	2875.3	2974.3	3074.3	3175.5	3278.1	3488.1	3704.7	3928.2	4158.7	4396.1
	s, kJ/(kg·K)		7.3622	7.6142	7.8351	8.0341	8.2165	8.3858	8.5442	8.8350	9.0984	9.3406	9.5660	9.7775

Temperature, °C

P, MPa (T_{sat}, °C)		150	200	250	300	350	400	450	500	550	600	700	800	900
0.15 (111.4)	v, m³/kg	1.285	1.444	1.601	1.757	1.912	2.067	2.222	2.376	2.530	2.685	2.993	3.301	3.609
	u, kJ/kg	2579.8	2656.2	2732.5	2809.5	2887.7	2967.3	3048.4	3131.1	3215.6	3301.6	3479.0	3663.4	3854.6
	h, kJ/kg	2772.6	2872.9	2972.7	3073.0	3174.5	3277.3	3381.7	3487.6	3595.1	3704.3	3927.9	4158.5	4395.9
	s, kJ/(kg·K)	7.4201	7.6441	7.8446	8.0278	8.1975	8.3562	8.5057	8.6473	8.7821	8.9109	9.1533	9.3787	9.5903
0.2 (120.2)	v, m³/kg	0.9596	1.080	1.199	1.316	1.433	1.549	1.665	1.781	1.897	2.013	2.244	2.475	2.706
	u, kJ/kg	2576.9	2654.4	2731.2	2808.6	2886.9	2966.7	3047.9	3130.7	3215.2	3301.4	3478.8	3663.2	3854.5
	h, kJ/kg	2768.8	2870.5	2971.0	3071.8	3173.5	3276.5	3381.0	3487.0	3594.7	3704.0	3927.7	4158.2	4395.8
	s, kJ/(kg·K)	7.2803	7.5074	7.7094	7.8934	8.0636	8.2226	8.3723	8.5140	8.6489	8.7778	9.0203	9.2458	9.4574
0.4 (143.6)	v, m³/kg	0.4708	0.5342	0.5951	0.6548	0.7139	0.7726	0.8311	0.8893	0.9475	1.006	1.121	1.237	1.353
	u, kJ/kg	2564.5	2646.8	2726.1	2804.8	2884.0	2964.4	3046.0	3129.2	3213.9	3300.2	3477.9	3662.5	3853.9
	h, kJ/kg	2752.8	2860.5	2964.2	3066.7	3169.6	3273.4	3378.4	3484.9	3592.9	3702.4	3926.5	4157.4	4395.1
	s, kJ/(kg·K)	6.9307	7.1714	7.3797	7.5670	7.7390	7.8992	8.0497	8.1921	8.3274	8.4566	8.6995	8.9253	9.1370
0.6 (158.9)	v, m³/kg		0.3520	0.3938	0.4344	0.4742	0.5137	0.5529	0.5920	0.6309	0.6697	0.7472	0.8245	0.9017
	u, kJ/kg		2638.9	2720.9	2801.0	2881.1	2962.0	3044.1	3127.6	3212.5	3299.1	3477.1	3661.8	3853.3
	h, kJ/kg		2850.1	2957.2	3061.6	3165.7	3270.2	3375.9	3482.7	3591.1	3700.9	3925.4	4156.5	4394.4
	s, kJ/(kg·K)		6.9673	7.1824	7.3732	7.5472	7.7086	7.8600	8.0029	8.1386	8.2682	8.5115	8.7375	8.9494
0.8 (170.4)	v, m³/kg		0.2608	0.2931	0.3241	0.3544	0.3843	0.4139	0.4433	0.4726	0.5018	0.5601	0.6181	0.6761
	u, kJ/kg		2630.6	2715.5	2797.1	2878.2	2959.7	3042.2	3125.9	3211.2	3297.9	3476.2	3661.1	3852.8
	h, kJ/kg		2839.2	2950.0	3056.4	3161.7	3267.1	3373.3	3480.6	3589.3	3699.4	3924.3	4155.7	4393.6
	s, kJ/(kg·K)		6.8167	7.0392	7.2336	7.4097	7.5723	7.7245	7.8680	8.0042	8.1341	8.3779	8.6041	8.8161
1 (179.9)	v, m³/kg		0.2060	0.2327	0.2579	0.2825	0.3066	0.3304	0.3541	0.3776	0.4011	0.4478	0.4943	0.5407
	u, kJ/kg		2621.9	2709.9	2793.2	2875.2	2957.3	3040.2	3124.3	3209.8	3296.8	3475.4	3660.5	3852.2
	h, kJ/kg		2827.9	2942.6	3051.2	3157.7	3263.9	3370.7	3478.4	3587.5	3697.9	3923.1	4154.8	4392.9
	s, kJ/(kg·K)		6.6948	6.9255	7.1237	7.3019	7.4658	7.6188	7.7630	7.8996	8.0298	8.2740	8.5005	8.7127
1.5 (198.3)	v, m³/kg		0.1325	0.1520	0.1697	0.1866	0.2030	0.2192	0.2352	0.2510	0.2668	0.2981	0.3292	0.3603
	u, kJ/kg		2598.1	2695.3	2783.1	2867.6	2951.3	3035.3	3120.3	3206.4	3293.9	3473.2	3658.7	3850.8
	h, kJ/kg		2796.8	2923.2	3037.6	3147.4	3255.8	3364.1	3473.0	3582.9	3694.0	3920.3	4152.6	4391.2
	s, kJ/(kg·K)		6.4554	6.7098	6.9187	7.1025	7.2697	7.4249	7.5706	7.7083	7.8393	8.0846	8.3118	8.5243

7.21

TABLE 7.6 Properties of Superheated Water Vapor (*Continued*)

P, MPa (T_{sat}, °C)		Temperature, °C												
		250	300	350	400	450	500	550	600	650	700	750	800	900
2 (212.4)	v, m³/kg	0.1114	0.1255	0.1386	0.1512	0.1635	0.1757	0.1877	0.1996	0.2114	0.2232	0.2350	0.2467	0.2700
	u, kJ/kg	2679.6	2772.6	2859.8	2945.2	3030.4	3116.2	3203.0	3290.9	3380.2	3471.0	3563.2	3657.0	3849.3
	h, kJ/kg	2902.5	3023.5	3137.0	3247.6	3357.5	3467.6	3578.3	3690.1	3803.1	3917.5	4033.2	4150.4	4389.4
	s, kJ/(kg·K)	6.5461	6.7672	6.9571	7.1279	7.2853	7.4325	7.5713	7.7032	7.8290	7.9496	8.0656	8.1774	8.3903
3 (233.9)	v, m³/kg	0.07058	0.08114	0.09053	0.09936	0.1079	0.1162	0.1244	0.1324	0.1404	0.1484	0.1563	0.1641	0.1798
	u, kJ/kg	2644.0	2750.0	2843.7	2932.7	3020.4	3107.9	3196.0	3285.0	3375.2	3466.6	3559.4	3653.6	3846.5
	h, kJ/kg	2855.8	2993.5	3115.3	3230.8	3344.0	3456.5	3569.1	3682.3	3796.5	3911.7	4028.2	4146.0	4385.9
	s, kJ/(kg·K)	6.2880	6.5398	6.7436	6.9220	7.0842	7.2346	7.3757	7.5093	7.6364	7.7580	7.8747	7.9871	8.2008
4 (250.4)	v, m³/kg		0.05884	0.06645	0.07341	0.08003	0.08643	0.09269	0.09885	0.1049	0.1109	0.1169	0.1229	0.1347
	u, kJ/kg		2725.3	2826.6	2919.9	3010.1	3099.5	3189.0	3279.1	3370.1	3462.1	3555.5	3650.1	3843.6
	h, kJ/kg		2960.7	3092.4	3213.5	3330.2	3445.2	3559.7	3674.4	3789.8	3905.9	4023.2	4141.6	4382.3
	s, kJ/(kg·K)		6.3622	6.5828	6.7698	6.9371	7.0908	7.2343	7.3696	7.4981	7.6206	7.7381	7.8511	8.0655
6 (275.6)	v, m³/kg		0.03616	0.04223	0.04739	0.05214	0.05665	0.06101	0.06525	0.06942	0.07352	0.07758	0.08160	0.08958
	u, kJ/kg		2667.2	2789.6	2892.8	2988.9	3082.2	3174.6	3266.9	3359.6	3453.2	3547.6	3643.1	3837.8
	h, kJ/kg		2884.2	3043.0	3177.2	3301.8	3422.1	3540.6	3658.4	3776.2	3894.3	4013.1	4132.7	4375.3
	s, kJ/(kg·K)		6.0682	6.3342	6.5415	6.7201	6.8811	7.0296	7.1685	7.2996	7.4242	7.5433	7.6575	7.8735
8 (295.1)	v, m³/kg		0.02426	0.02995	0.03432	0.03817	0.04175	0.04516	0.04845	0.05166	0.05481	0.05791	0.06097	0.06702
	u, kJ/kg		2590.9	2747.7	2863.8	2966.7	3064.3	3159.8	3254.4	3349.0	3444.0	3539.6	3636.1	3832.1
	h, kJ/kg		2785.0	2987.3	3138.3	3272.0	3398.3	3521.0	3642.0	3762.3	3882.5	4002.9	4123.8	4368.3
	s, kJ/(kg·K)		5.7914	6.1309	6.3642	6.5559	6.7248	6.8786	7.0214	7.1553	7.2821	7.4027	7.5182	7.7359
10 (311.1)	v, m³/kg			0.02242	0.02641	0.02975	0.03279	0.03564	0.03837	0.04101	0.04358	0.04611	0.04859	0.05349
	u, kJ/kg			2699.2	2832.4	2943.3	3045.8	3144.5	3241.7	3338.2	3434.7	3531.5	3629.0	3826.3
	h, kJ/kg			2923.4	3096.5	3240.8	3373.6	3500.9	3625.3	3748.3	3870.5	3992.6	4114.9	4361.2
	s, kJ/(kg·K)			5.9451	6.2127	6.4197	6.5974	6.7569	6.9037	7.0406	7.1696	7.2919	7.4086	7.6280
12 (324.8)	v, m³/kg			0.01721	0.02108	0.02412	0.02680	0.02929	0.03164	0.03390	0.03610	0.03824	0.04034	0.04447
	u, kJ/kg			2641.1	2798.3	2918.8	3026.6	3128.9	3228.7	3327.2	3425.3	3523.4	3621.8	3820.6
	h, kJ/kg			2847.6	3051.2	3208.2	3348.2	3480.3	3608.3	3734.0	3858.4	3982.3	4105.9	4354.2
	s, kJ/(kg·K)			5.7604	6.0754	6.3006	6.4879	6.6535	6.8045	6.9445	7.0757	7.1998	7.3178	7.5390

Temperature, °C

P, MPa (T_sat, °C)		400	450	500	550	600	650	700	750	800	850	900	950	1000
15 (342.2)	v, m³/kg	0.01565	0.01845	0.02080	0.02293	0.02491	0.02680	0.02861	0.03037	0.03210	0.03379	0.03546	0.03711	0.03875
	u, kJ/kg	2740.7	2879.5	2996.5	3104.7	3208.6	3310.4	3410.9	3511.0	3611.0	3711.2	3811.9	3913.2	4015.4
	h, kJ/kg	2975.4	3156.2	3308.5	3448.6	3582.3	3712.3	3840.1	3966.6	4092.4	4218.0	4343.8	4469.9	4596.6
	s, kJ/(kg·K)	5.8819	6.1412	6.3451	6.5207	6.6784	6.8232	6.9580	7.0848	7.2048	7.3192	7.4288	7.5340	7.6356
20 (365.8)	v, m³/kg	0.00994	0.01270	0.01477	0.01656	0.01818	0.1969	0.02113	0.02251	0.02385	0.02516	0.02645	0.02771	0.02897
	u, kJ/kg	2619.2	2806.2	2942.8	3062.3	3174.0	3281.5	3386.5	3490.0	3592.7	3695.1	3797.4	3900.0	4003.1
	h, kJ/kg	2818.1	3060.1	3238.2	3393.4	3537.6	3675.3	3809.1	3940.3	4069.8	4198.3	4326.4	4454.3	4582.5
	s, kJ/(kg·K)	5.5548	5.9025	6.1409	6.3356	6.5056	6.6591	6.8002	6.9317	7.0553	7.1723	7.2839	7.3907	7.4933
22.088 (374.136)	v, m³/kg	0.00818	0.01104	0.01305	0.01475	0.01627	0.01768	0.01901	0.02029	0.02152	0.02272	0.02389	0.02505	0.02619
	u, kJ/kg	2552.9	2772.1	2919.0	3043.9	3159.1	3269.1	3376.1	3481.1	3585.0	3688.3	3791.4	3894.5	3998.0
	h, kJ/kg	2733.7	3015.9	3207.2	3369.6	3518.4	3659.6	3796.0	3929.2	4060.3	4190.1	4319.1	4447.9	4576.6
	s, kJ/(kg·K)	5.4013	5.8072	6.0634	6.2670	6.4426	6.5998	6.7437	6.8772	7.0024	7.1206	7.2330	7.3404	7.4436
30	v, m³/kg	0.00279	0.00674	0.00868	0.01017	0.01145	0.01260	0.01366	0.01466	0.01562	0.01655	0.01745	0.01833	0.01920
	u, kJ/kg	2067.3	2619.3	2820.7	2970.3	3100.5	3221.0	3335.8	3447.0	3555.6	3662.6	3768.5	3873.8	3978.8
	h, kJ/kg	2151.0	2821.4	3081.0	3275.4	3443.9	3598.9	3745.7	3886.9	4024.3	4159.0	4291.9	4423.6	4554.7
	s, kJ/(kg·K)	4.4736	5.4432	5.7912	6.0350	6.2339	6.4066	6.5614	6.7030	6.8341	6.9568	7.0726	7.1825	7.2875
40	v, m³/kg	0.00191	0.00369	0.00562	0.00698	0.00809	0.00906	0.00994	0.01076	0.01152	0.01226	0.01296	0.01365	0.01432
	u, kJ/kg	1854.5	2365.1	2678.4	2869.7	3022.6	3158.0	3283.6	3402.9	3517.9	3629.8	3739.4	3847.5	3954.6
	h, kJ/kg	1930.8	2512.8	2903.3	3149.1	3346.4	3520.6	3681.3	3833.1	3978.8	4120.0	4257.9	4393.6	4527.6
	s, kJ/(kg·K)	4.1143	4.9467	5.4707	5.7793	6.0122	6.2063	6.3759	6.5281	6.6671	6.7957	6.9158	7.0291	7.1365
60	v, m³/kg	0.00163	0.00208	0.00296	0.00396	0.00483	0.00560	0.00627	0.00689	0.00746	0.00800	0.00851	0.00900	0.00948
	u, kJ/kg	1745.3	2053.9	2390.5	2658.8	2861.1	3028.8	3177.2	3313.6	3441.6	3563.6	3681.0	3795.0	3906.4
	h, kJ/kg	1843.4	2179.0	2567.9	2896.2	3151.2	3364.5	3553.6	3726.8	3889.1	4043.3	4191.5	4335.0	4475.2
	s, kJ/(kg·K)	3.9325	4.4128	4.9329	5.3449	5.6460	5.8838	6.0832	6.2569	6.4118	6.5523	6.6814	6.8012	6.9135
80	v, m³/kg	0.00152	0.00177	0.00219	0.00276	0.00339	0.00398	0.00452	0.00502	0.00548	0.00591	0.00632	0.00671	0.00709
	u, kJ/kg	1687.0	1944.9	2218.9	2483.9	2711.8	2904.7	3073.2	3225.3	3365.7	3497.3	3622.3	3742.1	3857.8
	h, kJ/kg	1808.3	2086.9	2393.9	2704.9	2982.7	3222.8	3434.7	3626.6	3803.8	3970.1	4127.9	4279.1	4425.2
	s, kJ/(kg·K)	3.8338	4.2328	4.6432	5.0331	5.3609	5.6284	5.8521	6.0445	6.2137	6.3652	6.5026	6.6289	6.7459

Source: From Reynolds and Perkins.[1]

7.23

TABLE 7.7 Properties of Air

a) Approximate Composition of Air

Gas	Molecular weight	% volume	% mass
Oxygen	32	21	23
Nitrogen	28	79	77

b) General Properties of Air (at 300 K, 1 bar)

Mean molecular weight	$M = 28.96$
Specific heat at constant pressure	$c_p = 1.005$ kJ kg^{-1}
Specific heat at constant volume	$c_v = 0.718$ kJ kg^{-1} K^{-1}
Density	$\rho = 1.183$ kg m^{-3}
Dynamic viscosity	$\mu = 1.853 \times 10^{-5}$ Ns m^{-2}
Kinematic viscosity	$v = 1.566 \times 10^{-5}$ m^2 s^{-1}
Thermal conductivity	$k = 0.0261$ W m^{-1} K^{-1}
Thermal diffusivity	$\alpha = 2203$ m^2 s^{-1}
Prandtl number	$P_r = 0.711$

Source: Revised and adopted from Refs. 1–3.

5. The Phase Rule. For equilibrium states, not all variables which describe the state of the system are independent, and fixing a limited number of them automatically establishes the others. This number of independent variables (pressure, temperature, mole fractions, etc.) is given by the phase rule and is called the *number of degrees of freedom of the system.* The following relationship (the actual phase rule) applies:

$$F = 2 - \pi + m - r \qquad (7.63)$$

where π = the number of phases
m = the number of chemical species
r = the number of independent chemical reactions
F = the number of degrees of freedom

For example, for the liquid-vapor system $F = 1$ when $\pi = 2$, $m = 1$, and $r = 0$, i.e., at a given pressure two phases coexist only at one temperature.

6. Systems with Flow of Compressible Fluids

Flow through Orifices and Nozzles. As a compressible fluid passes through a nozzle, its pressure drops and simultaneously its velocity increases. By assuming that flow is adiabatic, it is possible to calculate from the properties of the fluid the required area for the cross section of the nozzle at any point. The smallest cross section of the nozzle is called its *throat,* and the pressure at the throat is the critical flow pressure p_m. If the nozzle is cut off at the throat with no diverging section and the pressure at the discharge end is progressively decreased, with fixed inlet pressure, the amount of fluid passing increases until the discharge pressure equals the critical pressure, but a further decrease in discharge pressure does not result in increased flow. For gases $p_m/p_1 \simeq 0.53$, for saturated steam $p_m/p_1 \simeq 0.575$, and for moderately superheated steam $p_m/p_1 \simeq 0.55$.
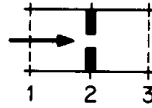
FIGURE 7.6

For orifice computation (see Fig. 7.6), the general relation from the first law gives

$$\frac{V_2^2 - V_1^2}{2} = h_1 - h_2 \tag{7.64}$$

where section 2 is at the orifice and section 3 is beyond the orifice. Then

$$V_2 = \frac{C\sqrt{2(h_1 - h_2)}}{\sqrt{1 - (A_2/A_1)^2(v_1/v_2)^2}} \tag{7.65}$$

where coefficient of discharge $C \approx 0.50$ to 0.60 (for steam nozzles this may be as high as 0.95). The volume flow is $V_2 A_2$ and mass flow is $V_2 A_2 \rho$.

For an ideal gas, assuming reversible adiabatic expansion through the orifice,

$$V_2 = \frac{C\sqrt{2p_1 v_1 [k/(k-1)][1 - (p_2/p_1)^{k-1/k}]}}{\sqrt{1 - (A_2/A_1)^2(p_2/p_1)^{2/k}}} \tag{7.66}$$

Throttling. When a fluid flows from a region of higher pressure into a region of lower pressure through a valve or similar constricted passage, it is said to be *throttled* or *withdrawn*. Equation (7.64) is applicable to the throttling process and since V_2 and V_1 are practically equal,

$$h_1 = h_2 \tag{7.67}$$

i.e., in a throttling process there is no change in enthalpy.

For a mixture of liquid and vapor, the equation of throttling is

$$h_{f_1} + x_1 \cdot h_{fg_1} = h_{f_2} + x_2 \cdot h_{fg_2} \tag{7.68}$$

POWER CYCLES*

1. Heat Engine. According to the first law of thermodynamics, when a system undergoes a complete cycle, the net heat supplied is equal to the net work done. This is based on the conservation of energy principle, which follows from the

---

observation of natural events. The second law of thermodynamics, which is also a natural law, indicates that, although the net heat supplied in a cycle is equal to the net work done, the gross heat supplied must be greater than the net work done [see Eq. (7.11)]; some heat must always be rejected by the system. To explain and analyze the second law and analyze power cycles, this chapter next discusses heat engine.

A heat engine is a system operating in a complete cycle and developing net work from a supply of heat. The second law implies that a source of heat supply and the rejection of heat are both necessary, since some heat must always be rejected by the system. Therefore, it is impossible for a heat engine to produce net work in a complete cycle if it exchanges heat only with bodies (reservoirs) at a single fixed temperature. A diagrammatic representation of a heat engine is shown in Fig. 7.7. The first and second laws apply equally well to cycles working in the reverse direction of those applicable to a heat engine. In the case of a reversed cycle, net work is done on the system which is equal to the net heat rejected by the system. Such cycles occur in heat pumps and refrigerators. A diagrammatic representation of a heat pump (or refrigerator) is shown in Fig. 7.8.

2. Ideal Gas Cycles.   The analysis of real power cycles can often be approximated by idealized cycles, using ideal gases as working fluids. Several such approximations are of interest.

Carnot Cycle.   This cycle consists of four processes, two isothermal and two isentropic, as shown in Fig. 7.9. A detailed analysis of the cycle shows that

$$q_{\text{net}} = w_{\text{net}} = (T_3 - T_4)R \ln \frac{p_2}{p_3}$$

$$q_{1\text{-}2} = q_{3\text{-}4} = 0$$

$$q_{2\text{-}3} = w_{2\text{-}3} = p_2 v_2 \ln \frac{p_2}{p_3}$$

$$w_{1\text{-}2} = \bar{c}_v(T_2 - T_1)$$

$$w_{3\text{-}4} = \bar{c}_v(T_4 - T_3)$$

$$u_3 - u_2 = 0 \qquad u_1 - u_4 = 0$$

$$q_{1\text{-}4} = w_{4\text{-}1} = p_4 v_4 \ln \frac{p_4}{p_1}$$

$$\eta = \frac{T_1 - T_2}{T_1} = \frac{T_3 - T_4}{T_3}$$

Otto Cycle.   This cycle consists of two isentropic and two constant-volume processes, as shown in Fig. 7.10, and is often used as a representation of a spark-ignition engine.
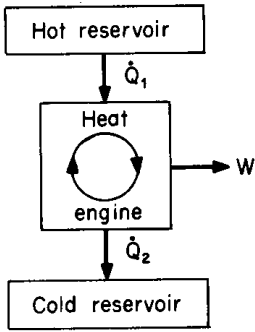
FIGURE 7.7

FIGURE 7.8

(a)

(b)
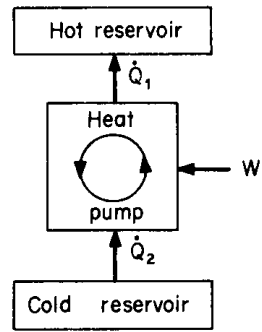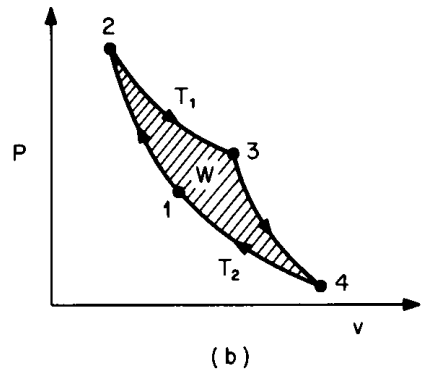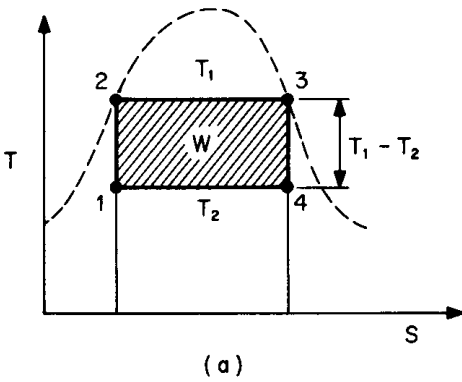
FIGURE 7.9 The Carnot cycle. (*a*) T-S diagram, (*b*) p-v diagram.
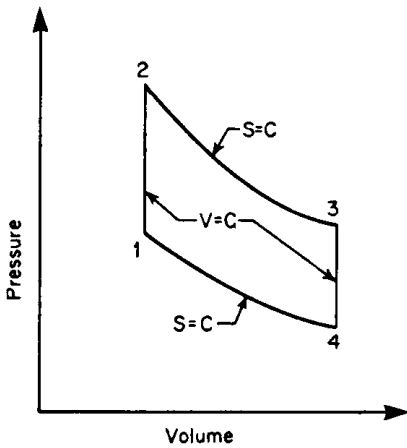
FIGURE 7.10 Otto cycle.
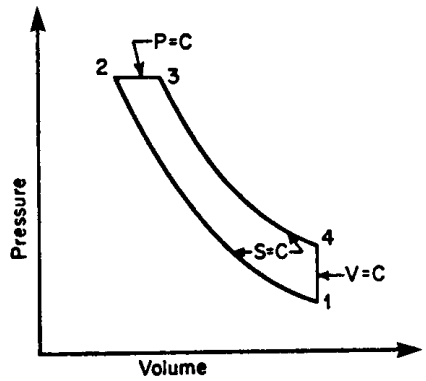
FIGURE 7.11 Diesel cycle.

Diesel Cycle. This cycle consists of two isentropic, one constant-volume, and one constant-pressure process, as shown in Fig. 7.11. It is representative of a diesel engine.

Brayton Cycle. This cycle consists of two isentropic and two constant-pressure processes. It is representative of the gas turbine, and is shown in Fig. 7.12.

Stirling Cycle. The stirling cycle consists of two isothermal and two constant-volume processes. Practical examples of this cycle have been developed recently, and depend regeneration to achieve practical efficiencies. Figure 7.13 shows this cycle. Practical cycles use the heat rejected in process 4-1 to partly regenerate the gas during process 2-3. Cycle analysis must depend on the degree of regeneration.
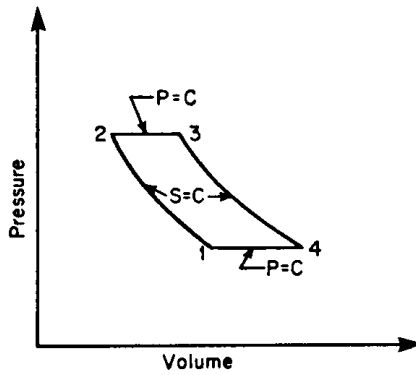
FIGURE 7.12   Brayton cycle.
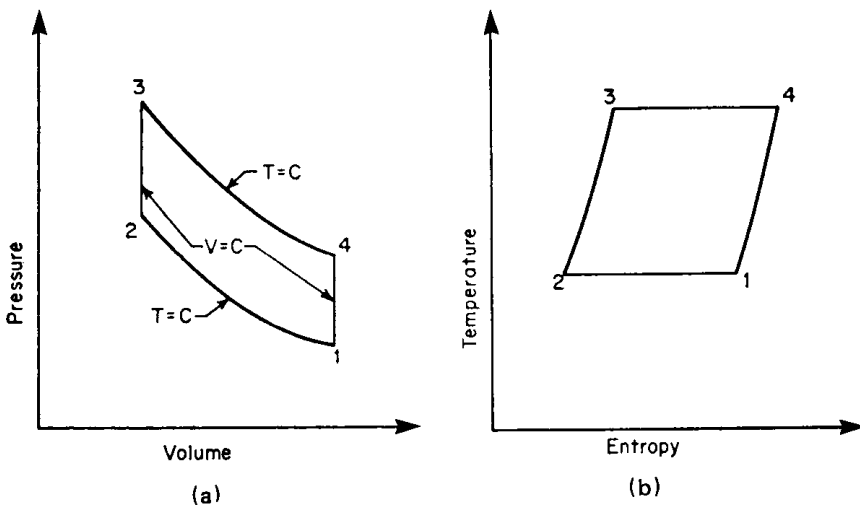
FIGURE 7.13   Stirling cycle. (*a*) p-v diagram; (*b*) T-S diagram.

3. Vapor Cycles.   Analyses of vapor cycles are done using Figs. 7.14 through 7.16 and Tables 7.4 through 7.6.

Rankine Cycle.   The Rankine cycle is the ideal representation for the vapor power cycle. It consists of five processes—two isothermal, two isentropic, and one constant-pressure. The cycle is shown in Fig. 7.17 and the corresponding apparatus diagram in Fig. 7.18.
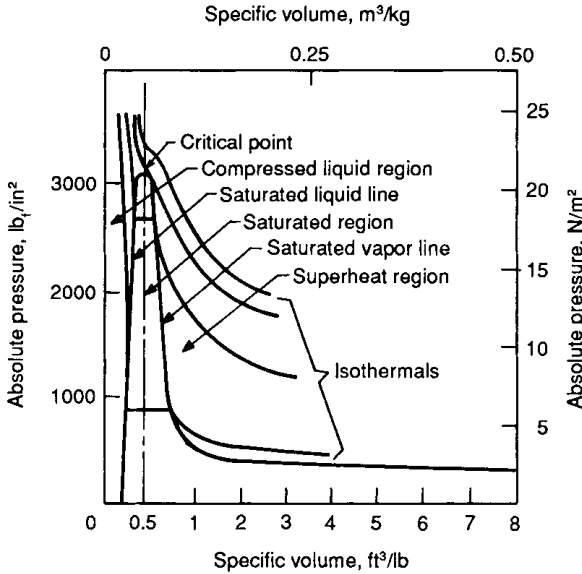
FIGURE 7.14   p-v diagram, water and water vapor.
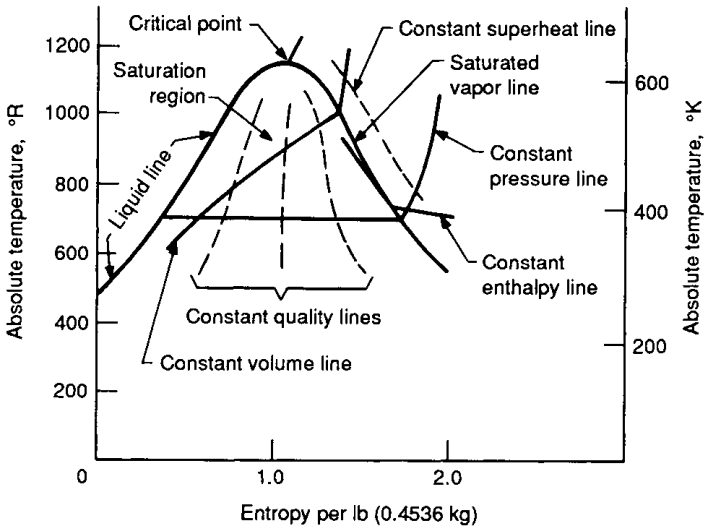
FIGURE 7.15   T-S diagram, water and water vapor.
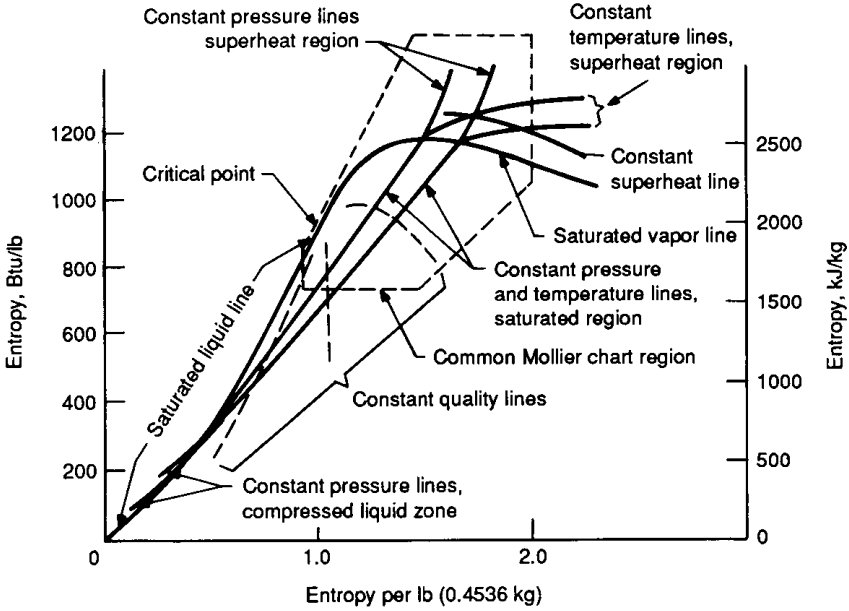
FIGURE 7.16 H-S diagram, water and water vapor.
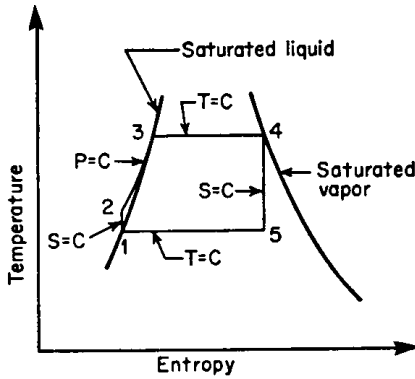
FIGURE 7.17 Rankine cycle.
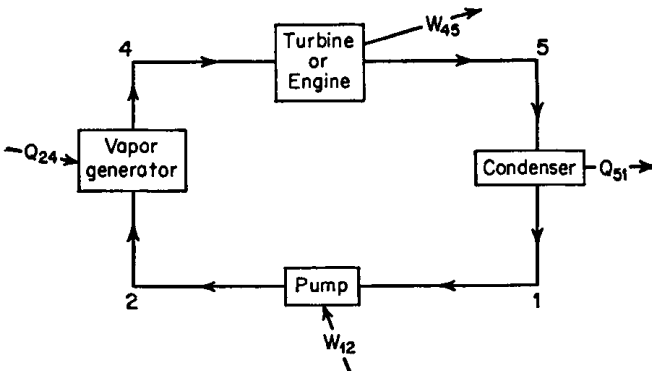
FIGURE 7.18 Vapor power system.

The results of a cycle analysis show

$$q_{2\text{-}3} = h_4 - h_2$$

$$q_{5\text{-}1} = h_1 - h_5$$

$$w_{4\text{-}5} = h_5 - h_4$$

$$w_{1\text{-}2} = h_1 - h_2$$

$$\eta = \frac{(h_5 - h_4) + w_{1\text{-}2}}{h_4 - h_2}$$

Compression Refrigeration Cycle. The idealized compression refrigeration cycle consists of two constant-pressure processes, an isentropic process, and an irreversible throttling process. The cycle is shown in Fig. 7.19 and the corresponding apparatus in Fig. 7.20.

Note that C in Figs. 7.10 through 7.13 and 7.17 through 7.20 represent the constant property line.

4. Working Fluid Selection. Different fluids are used as the "working fluid" in various power cycles and other energy conversion cycles. The most desirable properties of a working fluid are:

1. Large enthalpy of evaporation—to minimize the mass-flow rate for a given power output
2. High critical temperature—to permit evaporation at a high temperature
3. Low saturation pressures at the maximum temperatures—to minimize pressure vessel and piping costs
4. Pressure just above 1 atm at condensing temperature—to eliminate air leakage problems
5. Rapidly diverging pressure lines on the *h-s* diagram—to minimize reheating
6. No degrading aspects—to prevent corrosion and clogging
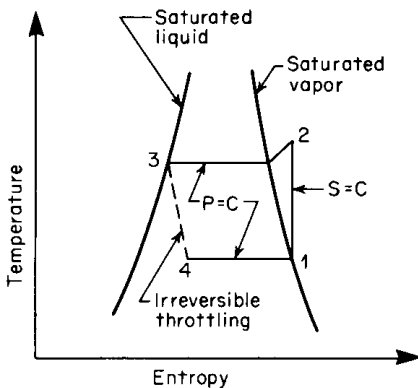
FIGURE 7.19 Compression-refrigeration cycle.
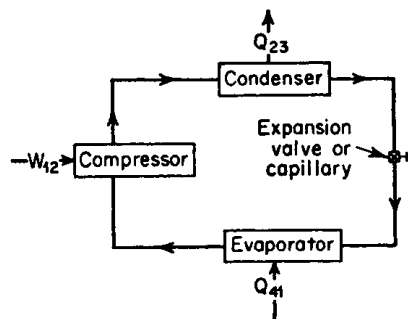
FIGURE 7.20 Compression-refrigeration system.

7. No hazardous features—to prevent toxicity and flammability

8. Low cost and ready availability

Note that water is rather good only in regard to items 1, 5, 7, and 8. Other fluids are better for other situations. However, water remains a predominant choice for ground-based power plants.

5. Summary of Engineering Examples of Energy Conversions

As shown, energy is the capacity, or ability, to do work. There are several forms of energy, and these include:

1. Heat or thermal

2. Chemical

3. Mechanical

4. Electrical

5. Nuclear

6. Light

7. Sound

Energy may be converted from one form to another. As shown in previous sections, the principle of conservation of energy states that the total amount of energy remains the same in such conversions, i.e., energy can not be created or destroyed.

Some examples of energy conversions:

1. Heat energy is converted to mechanical energy by a steam engine.

2. Heat energy is converted to electrical energy by a solar cell.

3. Heat energy is converted to chemical energy by living plants.

4. Heat energy is converted to electrical energy by a thermocouple.

5. Chemical energy is converted to heat energy by burning fuels.

6. Chemical energy is converted to electrical energy by batteries.

7. Mechanical energy is converted to electrical energy by a generator.

8. Mechanical energy is converted to heat energy by friction.

9. Electrical energy is converted to mechanical energy by a motor.

10. Electrical energy is converted to heat energy by an electric fire.

11. Electrical energy is converted to light energy by a light bulb.

12. Electrical energy is converted to chemical energy by electrolysis.

13. Sound energy is converted to electrical energy by a microphone.

Efficiency of energy conversion is defined as the ratio of the useful output energy to the input energy. Hence

$$\text{efficiency, } \eta = \frac{\text{useful output energy}}{\text{input energy}}$$

Efficiency, therefore, has no units and is often stated as a percentage. A "perfect machine" would have an efficiency of 100%. However, all machines have an efficiency lower than this due to friction and other losses.

NOMENCLATURE

Symbol = definition, SI units (U.S. Customary units)

c_v = specific heat at constant volume, J/kg · K (Btu/lb · °F)

c_p = specific heat at constant pressure, J/kg · K (Btu/lb · °F)

g = acceleration due to gravity, m/s² (ft/s²)

h = enthalpy per unit mass, J/kg (Btu/lb)

H = enthalpy, J (Btu)

$k = c_p/c_v$

m = mass, kg (lb)

$\dot{m}$ = mass flow rate, kg/s (lb/s)

p = pressure, N/m², Pa (lb$_f$/ft²)

Q = heat, J (Btu)

$\dot{Q}$ = heat rate, J/s (Btu/s)

q = heat per unit mass, J/kg (Btu/lb)

R = gas constant, J/kg · K (Btu/lb · °R)

R_m = universal gas constant, J/mol · K (Btu/mol · °R)

s = entropy per unit mass J/kg · K (Btu/lb · °R)

S = entropy, J/K (Btu/°R)

T = temperature, K (°R)

t = temperature, °C (°F)

u = internal energy per unit mass, J/kg (Btu/lb)

U = internal energy, J (Btu)

v = specific volume, m³/kg (ft³/lb)

$\mathbf{V}$ = volume, m³ (ft³)

$\mathbf{V}_m$ = molar volume, m³/mol (ft³/mol)

V = velocity, m/s (ft/s)

W = work, J (Btu)

w = work per unit mass, J/kg (Btu/lb)

W = work per unit time, J/s (Btu/s)

Greek

ρ = density, kg/m^3 (lb/ft^3)

Δ = finite change

η = thermal efficiency

Subscripts

g = gas (vapor)

g = dry gas (in "Clapeyron Equation" subsection)

f = liquid

fg = $f - g$, or change from liquid to gas

p = at constant p

v = at constant v

s = at constant s

v = vapor (in "Clapeyron Equation" subsection)

T = at constant T

rev = reversible

sat = saturation

1 = state 1

2 = state 2

1-2 = change from 1 to 2

Superscript

$^-$ = mean value

REFERENCES*

1. W. C. Reynolds and H. C. Perkins, *Engineering Thermodynamics,* 2d ed., McGraw-Hill, New York, 1977.

2. R. H. Perry, *Engineering Manual,* 3d ed., McGraw-Hill, New York, 1976.

3. R. H. Perry and D. W. Green, *Perry's Chemical Engineers' Handbook,* 6th ed., McGraw-Hill, New York, 1984.

4. E. A. Avallone and T. Baumeister III (eds.), *Mark's Standard Handbook for Mechanical Engineers,* 9th ed., McGraw-Hill, New York, 1987.

5. E. N. Ganić, "Heat Transfer Characteristics of Working Fluids for OTEC," *ASME Publication HTD,* vol. 12, pp. 55–66, 1980.

*Those references listed above but not cited in the text were used for comparison between different data sources, clarification, clarity of presentation, and, most important, reader's convenience when further interest in subject exists. Sketches on Figures 7.5, 7.10 to 7.13, 7.17 to 7.20 are from *Perry*.[2]

6. K. Gieck, *Engineering Formulas,* 5th ed., McGraw-Hill, New York, 1986.

7. E. G. Cravalo and J. L. Smith, Jr., *Engineering Thermodynamics,* Pitman, Boston, 1981.

8. T. D. Eastop and A. McConkey, *Applied Thermodynamics for Engineering Technologists,* 3d ed., Longman, London, 1978.

9. O. W. Eshbach and M. Sounder, *Handbook of Engineering Fundamentals,* 3d ed., John Wiley & Sons, New York, 1975.

10. V. M. Faires and C. M. Simmang, *Thermodynamics,* 6th ed., Macmillan, New York, 1978.

11. H. Gartmann, *De Laval Engineering Handbook,* 3d ed., McGraw-Hill, New York, 1970.

12. J. H. Keenan, *Thermodynamics,* MIT Press, Cambridge, Mass., 1970.

13. C. E. O'Rourke, *General Engineering Handbook,* 2d ed., McGraw-Hill, New York, 1940.

14. A. Shapiro, *Dynamics and Thermodynamics of Compressible Fluid Flow,* Ronald Press, New York, 1953.

15. M. W. Zemansky, *Heat and Thermodynamics,* 5th ed., McGraw-Hill, New York, 1968.

16. M. W. Zemansky, M. M. Abbott, and H. C. Van Ness, *Basic Engineering Thermodynamics,* 2d ed., McGraw-Hill, New York, 1975.

17. J. Bird, *Newnes Engineering Science Pocket Book*, 3d ed. Newnes, 2001.

18. J. F. Silver and J. E. Nydahl, *Introduction to Engineering Thermodynamics*, West Publishing Co., 1977.

CHAPTER 8
MECHANICS OF FLUIDS

NATURE OF FLUIDS

Fluid is a substance that deforms continuously when subjected to a shear stress, no matter how small that shear stress may be. A *shear force* is the force component tangent to a surface, and this force divided by the area of the surface is the average shear stress over the area. Shear stress at a point is a limiting value of shear force to area as the area is reduced to the point.

The relation between the shear stress τ and the rate of angular deformation for one-dimensional flow of a fluid du/dy is given as

$$\tau = \mu \frac{du}{dy} \tag{8.1}$$

where the proportionality factor μ is the viscosity of fluid.

The viscosity of a gas increases with temperature, but the viscosity of a liquid decreases with temperature. For ordinary pressures, viscosity is independent of pressure and depends on temperature only. For very great pressures, gases and most liquids have shown erratic variations of viscosity with pressure. The viscosity μ is frequently referred to as the *absolute viscosity* or the *dynamic viscosity* to avoid confusing it with the *kinematic viscosity* ν, which is the ratio of viscosity to mass density:

$$\nu = \frac{\mu}{\rho} \tag{8.2}$$

Fluids may be classified as newtonian or nonnewtonian. In a *newtonian fluid,* there is a linear relation between the magnitude of applied shear stress and the resulting rate of deformation [μ constant in Eq. (8.1)], as shown in Fig. 8.1. In a *nonnewtonian fluid,* there is a nonlinear relation between the magnitude of applied shear stress and the rate of angular deformation.

An *ideal plastic* has a definite yield stress and a constant linear relation of τ to du/dy. A *thixotropic* substance, such as printer's ink, has a viscosity that is dependent on the immediately prior angular deformation of the substance and has a tendency to take a set when at rest. Gases and liquids tend to be newtonian fluids, whereas thick, long-chained hydrocarbons may be nonnewtonian.

For purposes of analysis, the assumption is frequently made that a fluid is nonviscous (inviscid). With zero viscosity, the shear stress is always zero, regardless
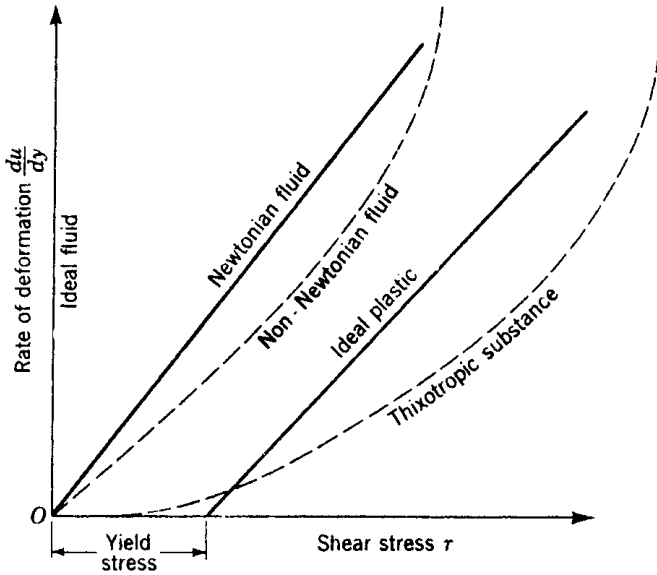
FIGURE 8.1 Deformation characteristics of substances. (*From Streeter and Wylie,*[1] *p. 5.*)

of the motion of the liquid. If the fluid is also considered to be incompressible, it is then called an *ideal fluid* and plots as the ordinate in Fig. 8.1.

The properties of density and viscosity play principal roles in open- and closed-channel flow and in flow around immersed objects. Surface-tension effects are important in the formation of droplets, in the flow of small jets, and in situations where liquid-gas-solid or liquid-liquid-solid interfaces occur, as well as in the formation of capillary waves. Values of density and surface tension for different fluids are given in Chap. 2.

Capillary attraction is caused by surface tension and by the relative value of adhesion between liquid and solid to cohesion of the liquid. A liquid that wets the solid has greater adhesion than cohesion. The action of surface tension in this case is to cause the liquid to rise within a small vertical tube that is partially immersed in it. For liquids that do not wet the solid, surface tension tends to depress the meniscus in a small vertical tube. When the contact angle between liquid and solids is known, the capillary rise can be computed for an assumed shape of meniscus. Figure 8.2 shows the capillary rise for water and mercury in circular glass tubes in air.

If the interface is curved, a mechanical balance shows that there is a pressure difference across the interface, the pressure being higher on the concave side. This is illustrated in Fig. 8.3. The pressure increase in the interior of a spherical droplet balances a ring of surface-tension force

$$\pi r^2 \, \Delta p = 2 \pi r \sigma \tag{8.3}$$

or
$$\Delta p = \frac{2\sigma}{r} \tag{8.4}$$

An important surface effect is also the contact angle θ that appears when a liquid interface intersects with a solid surface, as in Fig. 8.4. The force balance would
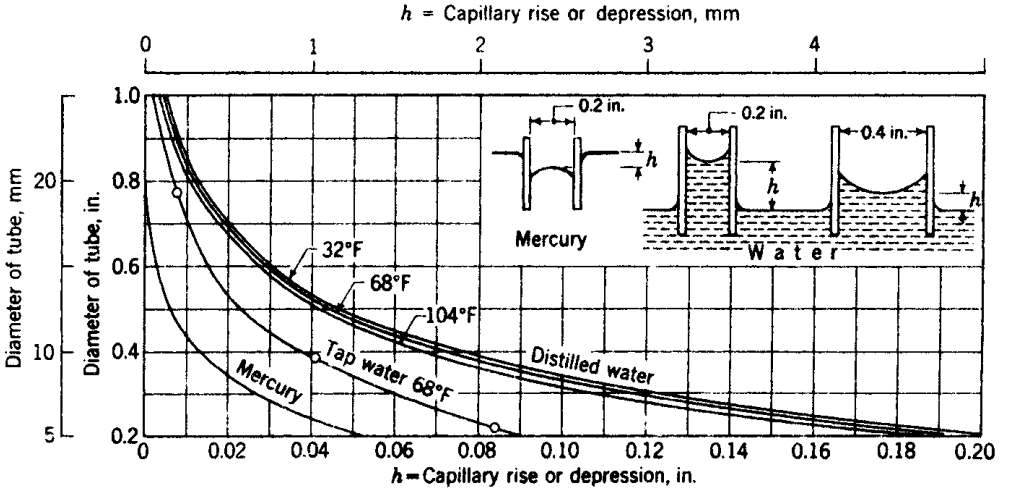
FIGURE 8.2 Capillarity in circular glass tubes. (*From Streeter and Wylie,*[1] *p. 18.*)
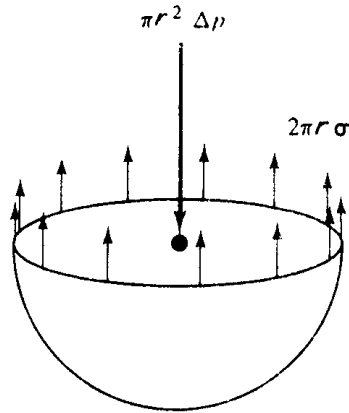
FIGURE 8.3 Pressure change across a curved interface resulting from surface tension. (*From White,*[2] *p. 32.*)
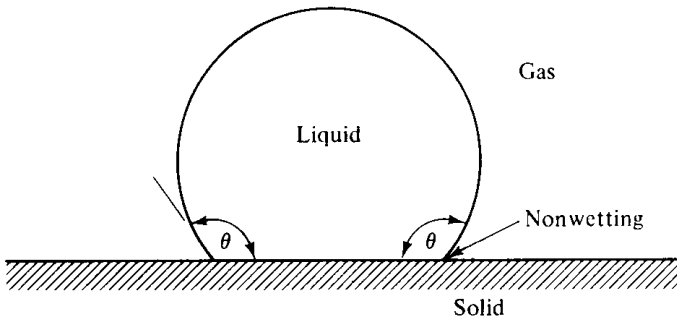
FIGURE 8.4 Contact-angle effects at a liquid-gas-solid interface. (*From White,*[2] *p. 33.*)

then involve both σ and θ. If the contact angle is less than 90°, the liquid is said to *wet* the solid; if $\theta > 90°$, the liquid is termed *nonwetting*.

The *specific weight* γ of a substance is its weight per unit volume. It changes with location, that is,

$$\gamma = \rho g \tag{8.5}$$

depending on gravity. It is a convenient property when dealing with fluid statics or with liquids with a free surface.

Specific gravity (sp gr) of a substance is the ratio of its weight to the weight of an equal volume of water at standard conditions. It also may be expressed as a ratio of a substance's density or specific weight to that of water.

FLUID STATICS

1. Introduction.   Fluid statics basically treats pressure and its variations throughout a fluid and pressure forces on finite surfaces. The average pressure is calculated by dividing the normal force pushing against a plane area by the area. The pressure at a point is the limit of the ratio of normal force to area as the area approaches zero size at the point. At a point, a fluid at rest has the same pressure in all directions.

The fluid static law of variation of pressure is given by the relation

$$\nabla p = \rho g \tag{8.6}$$

In component form (assuming that coordinate z is "up"), this equation becomes

$$\frac{\partial p}{\partial x} = 0 \qquad \frac{\partial p}{\partial y} = 0 \qquad \frac{\partial p}{\partial z} = -\rho g$$

Since pressure is a function of z, only

$$\frac{dp}{z} = -\rho g \tag{8.7}$$

or

$$p_2 - p_1 = -\int_1^2 \rho g \, dz \tag{8.8}$$

Equation (8.8) is the solution to the hydrostatic problem.

2. Hydrostatic Pressure in Liquids.   Assuming constant density in liquid hydrostatic calculation, Eq. (8.8) becomes

$$p_2 - p_1 = -\rho g (z_2 - z_1) \tag{8.9}$$

or

$$z_1 - z_2 = \frac{p_2}{\rho g} - \frac{p_1}{\rho g} \tag{8.10}$$

The quantity $p/\rho g$ is a length called the *pressure head* of the liquid.

For oceans and atmospheres, the coordinate system is usually chosen as in Fig. 8.5 with $z = 0$ at the free surface, where p equals the surface atmospheric pressure p_a.
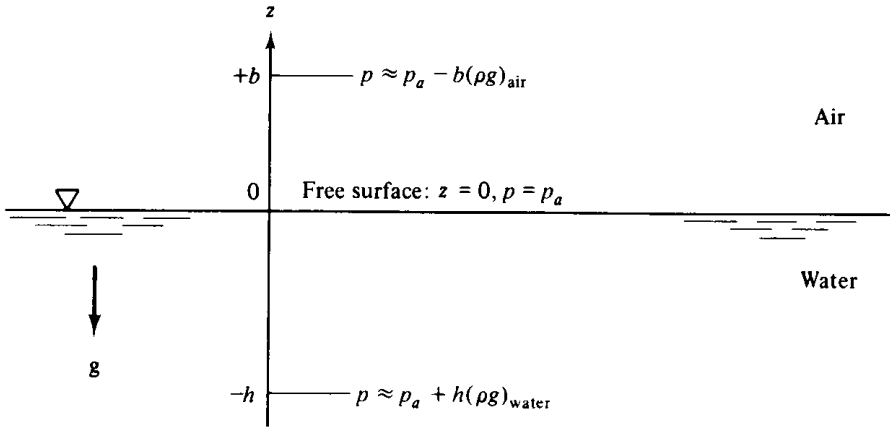
FIGURE 8.5 Hydrostatic-pressure distribution in oceans and atmospheres. (*From White,*[2] *p. 63.*)

3. Hydrostatic Pressure in Gases. Gases are compressible, with density proportional to pressure. It is sufficiently accurate to introduce the perfect-gas law

$$p = \rho RT$$

in Eq. (8.7). Then,

$$\int_1^2 \frac{dp}{p} = \ln \frac{p_2}{p_1} = - \frac{g}{R} \int_1^2 \frac{dz}{T} \tag{8.11}$$

One common approximation is the isothermal atmosphere, where $T = T_0$,

$$p_2 = p_1 \exp \left[- \frac{g(z_2 - z_1)}{RT_0} \right] \tag{8.12}$$

4. Scales of Pressure. Pressure may be expressed with reference to any arbitrary datum. The usual data are absolute zero and local atmospheric pressure. When a pressure is expressed as a difference between its value and a complete vacuum, it is called an *absolute pressure.* When it is expressed as a difference between its value and the local atmospheric pressure, it is called a *gage pressure.*

Figure 8.6 illustrates the data and the relations of the common units of pressure measurements. A pressure expressed in terms of the length of a column of liquid ($p = \gamma h$) is equivalent to the force per unit area at the base of the column.

5. Pressure-Sensing Devices. The two principal devices using liquids are the *barometer* and the *manometer.* The barometer senses absolute pressures, and the manometer senses pressure differential.

Manometers are a direct application of the basic equation of fluid statics. For a *U-tube manometer* (Fig. 8.7),

$$p_1 - p_2 = (\gamma_m - \gamma_f)h \tag{8.13}$$

For a *well* or *cistern-type manometer* (Fig. 8.8),
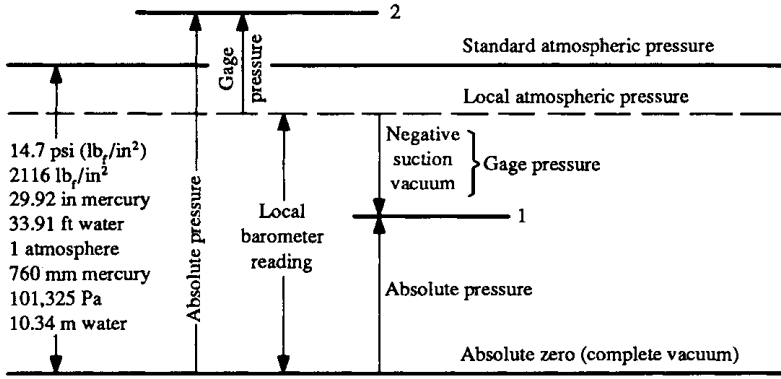
FIGURE 8.6 Units and scales for pressure measurements. (*From Streeter and Wylie,*[1] *p. 30.*)
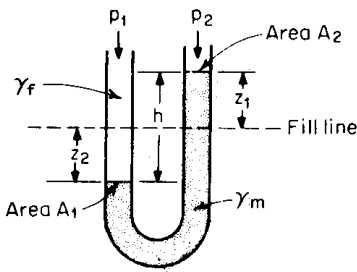
FIGURE 8.7 U-tube manometer. (*From Avallone and Baumeister,*[3] *p. 3-39.*)
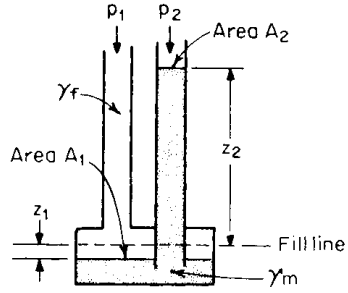
FIGURE 8.8 Well or cistern-type manometer. (*From Avallone and Baumeister,*[3] *p. 3-39.*)

$$p_1 - p_2 = (\gamma_m - \gamma_f)(z_2) \left(1 + \frac{A_2}{A_1} \right) \tag{8.14}$$

For an *inclined manometer* (Fig. 8.9),

$$p_1 - p_2 = (\gamma_m - \gamma_f) \left(\frac{A_2}{A_1} + \sin \theta \right) R \tag{8.15}$$
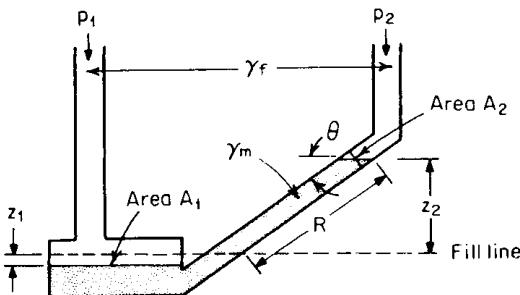
FIGURE 8.9 Inclined manometer. (*From Avallone and Baumeister,*[3] *p. 3-39.*)

6. Hydrostatic Forces. The force exerted by a liquid (i.e., caused by the weight of the fluid) on a plane submerged surface (Fig. 8.10) is given by

$$F = \int_A p \, dA = \gamma \int_A \sin \theta y \, dA = \gamma \sin \theta \bar{y} A = \gamma \bar{h} A = p_G A \qquad (8.16)$$

where $\qquad \bar{y} \sin \theta = \bar{h} \qquad$ and $\qquad p_G = \gamma \bar{h}$

i.e., the force F is the product of the area and the pressure at its centroid G.

The line of action of the resultant force F has its piercing point in the surface at the point called the *pressure center* (p.c.) with coordinates (x_p, y_p) given as

$$x_p = \frac{1}{\bar{y}A} \int_A xy \, dA \qquad (8.17)$$

$$y_p = \frac{1}{\bar{y}A} \int_A y^2 \, dA \qquad (8.18)$$

Note that the pressure center (p.c.) is always below the centroid of the surface.

The hydrostatic force acting on the curved surface AB is resolved into the horizontal component F_H and the vertical component F_V (Fig. 8.11). Note that F_V is equal to the weight of the fluid above the area AB. F_H is equal to the hydrostatic pressure force acting on the projection of the considered surface AB on the plane perpendicular to F_H. Calculation of F_V and F_H is accomplished also by Eq. (8.16).

7. Buoyant Force. The resultant force exerted on a body by a static fluid in which it is submerged or floating is called the *buoyant force*. The buoyant force always acts vertically upward. There can be no horizontal force of the resultant because the projection of the submerged body or submerged portion of the floating body on a vertical plane is always zero.
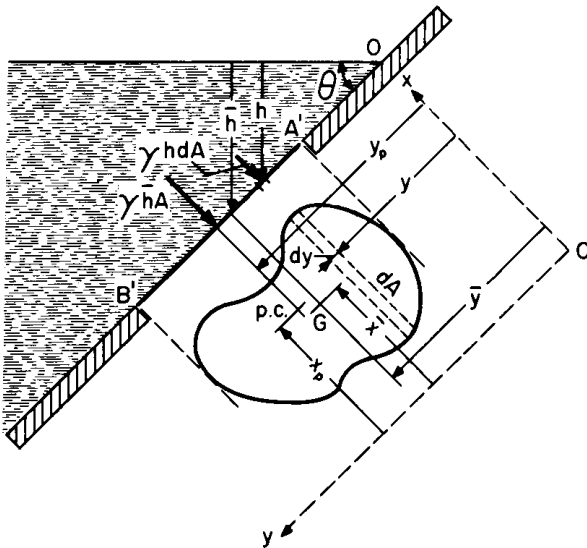
FIGURE 8.10 Notation for force of liquid on one side of a plane inclined area. (*From Streeter and Wylie,*[1] *p. 40.*)
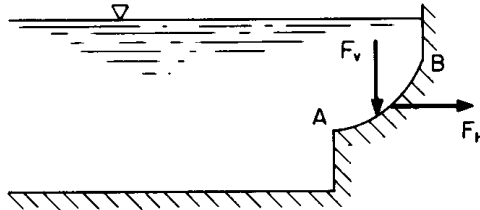
FIGURE 8.11   Forces on a curved surface.

For the system in Fig. 8.12, the buoyancy force F is equal to the weight of the displaced fluids of densities ρ_l and ρ_g, that is,

$$F = g\rho_l V + g\rho_g V' \tag{8.19}$$

If the fluid of density ρ_g is a gas, then

$$F \simeq g\rho_l V \tag{8.20}$$

If $\rho_l > \rho_m$ (where ρ_m is the density of the body), the body will float.
If $\rho_l = \rho_m$, the body will remain suspended.
If $\rho_l < \rho_m$, the body will sink.

FLUID-FLOW CHARACTERISTICS

Flow may be classified in many ways, such as turbulent, laminar, real, ideal, uniform, nonuniform, steady, unsteady, reversible, or irreversible.

1. Turbulent Flow.   Turbulent flow situations are most prevalent in engineering practice. In *turbulent flow,* the fluid particles move in very irregular paths, causing an exchange of momentum from one portion of the fluid to another. The fluid particles can range in size from very small (e.g., a few thousand molecules) to very large (e.g., thousands of cubic feet in a large swirl in a river or an atmospheric gust). Also, in turbulent flow, the momentum losses vary about 1.7 to 2 powers of the velocity; in laminar flow, they vary about the first power of the velocity.
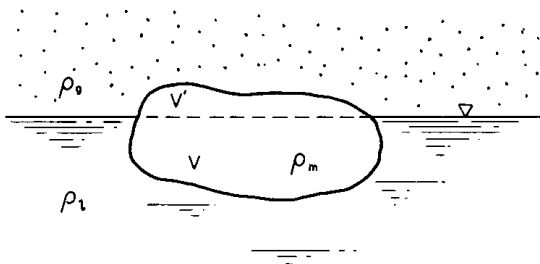
FIGURE 8.12   Buoyant force on the body.

2. Laminar Flow. In *laminar flow,* fluid particles move along smooth paths in *laminas,* or layers, with one layer gliding smoothly over an adjacent layer. Laminar flow is governed by Newton's law of viscosity [Eq. (8.1)] or extensions of it to three-dimensional flow, which relates shear stress to rate of angular deformation. Laminar flow is not stable in situations involving combinations of low viscosity, high velocity, or large flow passages and breaks down into turbulent flow. Figure 8.13 illustrates how an unstable laminar flow may turn into a turbulent flow. Figure 8.13*a* shows that a local disturbance causes an increase in the velocity of particle *A.* This increase in velocity will cause the pressure at *A* to fall below that at *B.* The pressure difference will further increase the velocity difference across the surface of discontinuity, as shown in Fig. 8.13*b.* The result is the formation of eddies at the surface of discontinuity and the initiation of turbulent flow.

The turbulent mixing also results in a more rapid transfer of momentum between different layers of fluid. Thus the velocity distribution in a turbulent flow is more uniform than that in a laminar flow, as shown in Fig. 8.14.

An equation similar in form to Newton's law of viscosity may be written for turbulent flow:

$$\tau = \eta \frac{du}{dy} \qquad (8.21)$$

The eddy viscosity η, however, is not a fluid property alone; it depends on the fluid motion and the density. In many practical flow situations, both viscosity and turbulence contribute to the shear stress:
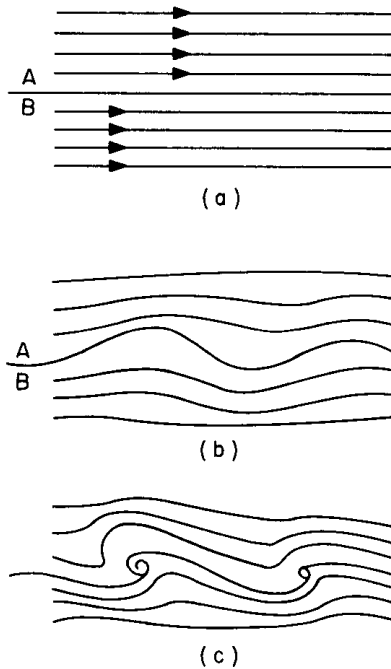
FIGURE 8.13 Eddies formed in an unstable laminar flow. (*From Probstein.*[4])

(a)

(b)
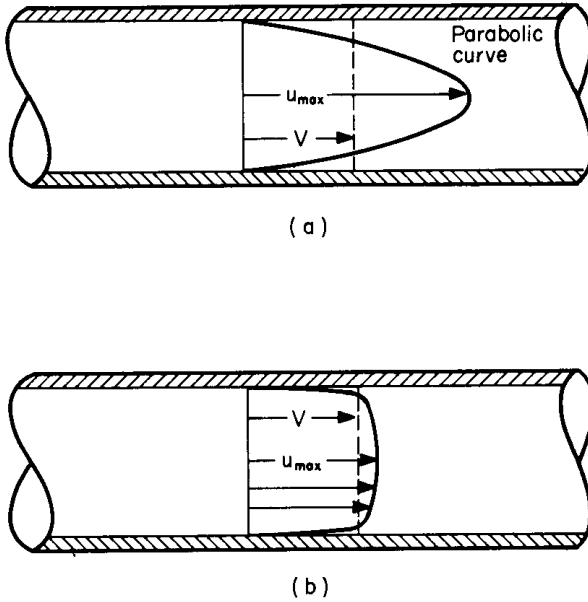
FIGURE 8.14   Comparison of velocity distribution for (*a*) laminar and (*b*) turbulent pipe flow. (*From Probstein.*[4])

$$\tau = (\mu + \eta)\,\frac{du}{dy} \qquad\qquad (8.22)$$

and experiments are required to determine this type of flow.

3. Boundary Layer.   An *ideal fluid* is frictionless (inviscid) and incompressible and should not be confused with a perfect (ideal) gas. The assumption of an ideal fluid is helpful in analyzing flow situations involving large expanses of fluids, as in the motion of an airplane or a submarine. A frictionless fluid is nonviscous, and its flow processes are reversible.

The layer of fluid in the immediate neighborhood of an actual flow boundary that has had its velocity relative to the boundary affected by viscous shear is called the *boundary layer.* Boundary layers may be laminar or turbulent (Fig. 8.15) depending generally on their length, the viscosity, the velocity of the flow near them, and the boundary roughness.

The *boundary-layer thickness* δ may be defined as the region in which the fluid velocity changes from its free-stream, or inviscid-flow, value to zero at the body surface (Fig. 8.16). Developing boundary layers in the entrance of a duct flow are shown in Fig. 8.17.

A particularly interesting phenomenon connected with transition in the boundary layer occurs with blunt bodies, e.g., a sphere or a circular cylinder. In the region of adverse pressure gradient (Fig. 8.18), the boundary layer separates from the surface. At this location, the shear stress goes to zero, and beyond this point there is a reversal of flow in the vicinity of the wall, as shown in Fig. 8.18. In this separation region, analysis of the viscous flow is very difficult, and emphasis is placed on the use of experimental methods.
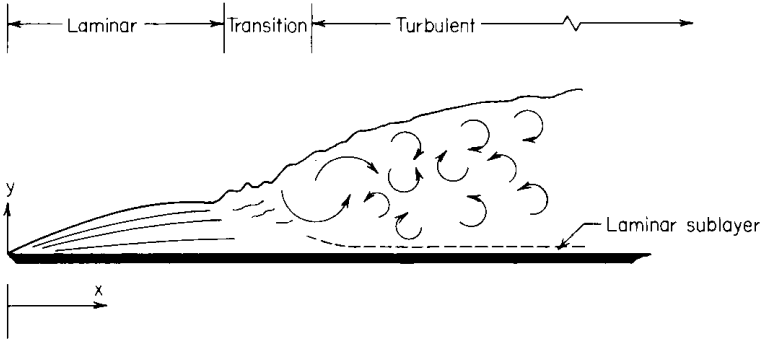
FIGURE 8.15 Laminar, transition, and turbulent boundary-layer-flow regimes in flow over a flat plate. (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 1-9.*)
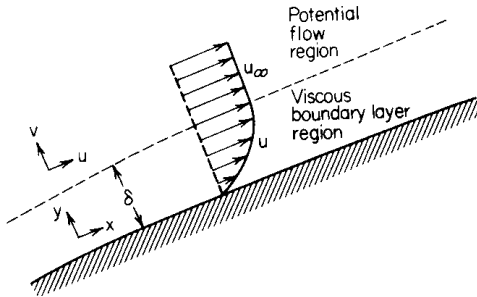
FIGURE 8.16 Boundary-layer flow past an external surface. (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 1-8.*)

4. Steady and Unsteady Flow. *Steady flow* occurs when conditions at any point in the fluid do not change with time. This can be expressed as $\partial V / \partial t = 0$, in which space ($x$, y, z coordinates of the point) is held constant. Likewise, in steady flow there is no change in density ρ, pressure p, or temperature T with time at any point. Thus

$$\frac{\partial \rho}{\partial t} = 0 \qquad \frac{\partial p}{\partial t} = 0 \qquad \frac{\partial T}{\partial t} = 0 \qquad (8.23)$$

In turbulent flow, a very efficient mixing takes place; i.e., macroscopic chunks of fluid move across streamlines and transport energy and mass as well as momentum vigorously. The most essential feature of a turbulent flow is the fact that at a given point in it, the flow property X (e.g., velocity component, pressure, temperature, etc.) is not constant with time but exhibits very irregular, high-frequency fluctuations (Fig. 8.19). At any instant, X may be represented as the sum of a time-mean value $\overline{X}$ and a fluctuating component X'. The average is taken over a time that is large compared with the period of typical fluctuation, and if $\overline{X}$ is independent of time, the time-mean flow is said to be *steady*.
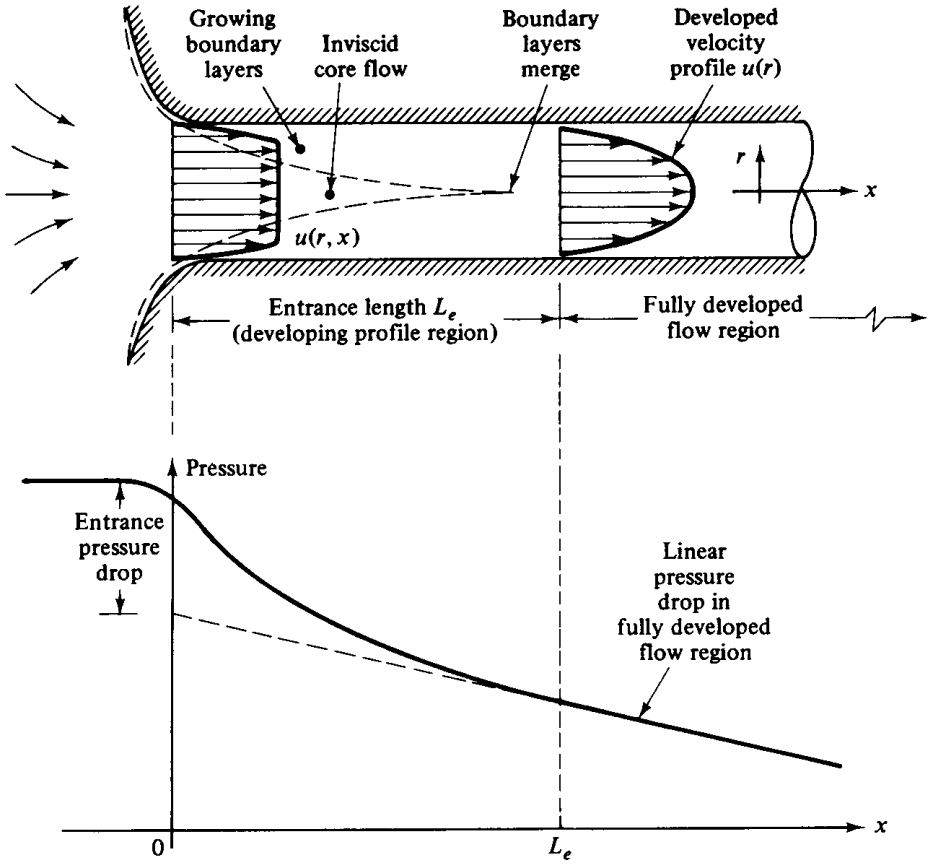
FIGURE 8.17 Developing velocity profiles and pressure changes in the entrance of a duct flow. (*From White,*[2] *p. 312.*)
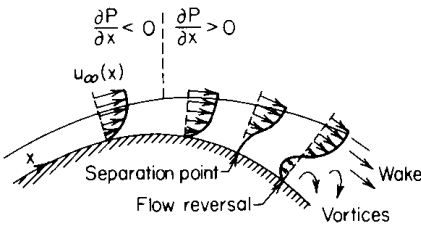
FIGURE 8.18 Velocity profile associated with separation on a circular cylinder in cross flow. (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 1-12.*)
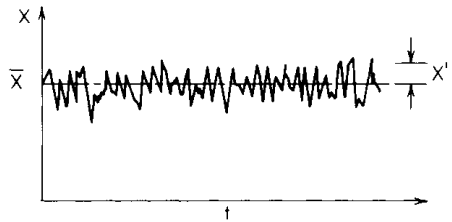
FIGURE 8.19 Property variation with time at some point in steady turbulent flow. (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 1-10.*)

The flow is *unsteady* when conditions at any point change with time, that is, $\partial V/\partial t \neq 0$. Water being pumped through a fixed system at a constant rate is an example of steady flow. Water being pumped through a fixed system at an increasing rate is an example of unsteady flow.

Uniform flow occurs when, at every point, the velocity vector is identically the same (in magnitude and direction) for any given instant.

One-dimensional flow neglects variations or changes in velocity, pressure, etc. transverse to the main flow direction. Conditions at a cross section are expressed in terms of average values of velocity, density, and other properties. Flow through a pipe, for example, usually may be characterized as one dimensional.

5. Streamlines and Stream Tubes. Velocity is a vector, and hence it has both magnitude and direction. A *streamline* is a line that gives the direction of the velocity of a fluid particle at each point in the flow stream. When streamlines are connected by a closed curve in steady flow, they will form a boundary through which the fluid particles cannot pass. The space between the streamlines becomes a *stream tube* (Fig. 8.20).

The stream-tube concept broadens the application of fluid-flow principles; for example, it allows treating the flow inside a pipe and the flow around an object with the same laws.

Velocity V along a streamline is a function of distance s and time t, or $V = f(s, t)$. Also,

$$dV = \frac{\partial V}{\partial s} \, ds + \frac{\partial V}{\partial t} \, dt$$

Acceleration dV/dt may be obtained for the last equation:

(a)   (b)
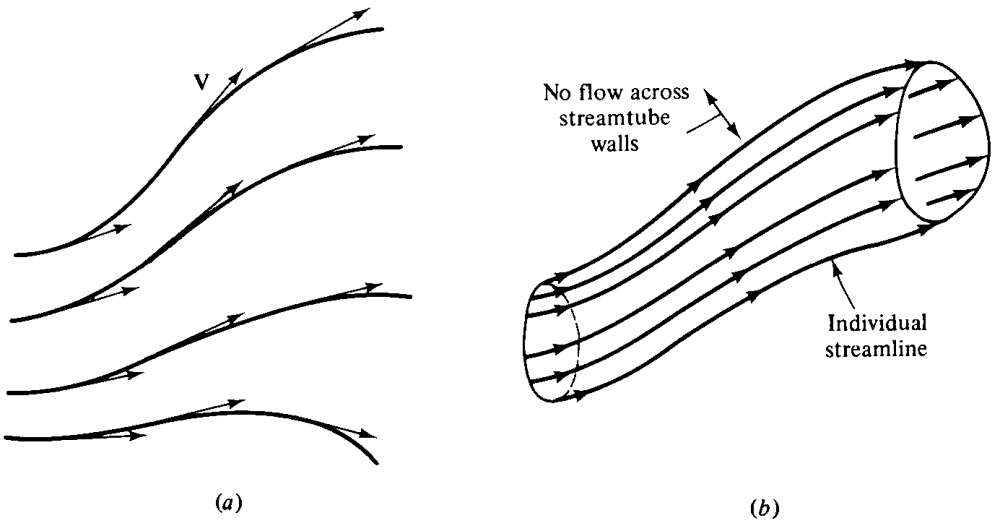
FIGURE 8.20   The most common method of flow-pattern presentation: (*a*) streamlines are everywhere tangent to the local velocity vector; (*b*) a stream tube is formed by a closed collection of streamlines. (*From White,*[2] *p. 40.*)

$$\frac{dV}{dt} = \frac{\partial V}{\partial s}\frac{ds}{dt} + \frac{\partial V}{\partial t} \tag{8.24}$$

For steady flow, $\partial V/\partial t = 0$.

FLUID DYNAMICS

This section includes basic relations for steady one-dimensional flow.

1. Continuity Equation. (See Fig. 8.21.)

$$\rho_1 A_1 V_1 = \rho A V = \rho_2 A_2 V_2 = \dot{m} \tag{8.25}$$

where $\dot{m}$ is mass flow rate, which is constant across every section of the tube. Equation (8.25) is a law of conservation of mass.

The continuity equation also may be written as

$$\frac{dA}{A} = -\frac{dV}{V} - \frac{d\rho}{\rho} \tag{8.26}$$

For incompressible fluids, $d\rho = 0$, so

$$\frac{dA}{A} = -\frac{dV}{V} \tag{8.27}$$

From this equation,

If the area increases, the velocity decreases.

If the area is constant, the velocity is constant.

There are no critical values.

For the frictionless flow of compressible fluids, it can be shown that

$$\frac{dA}{A} = -\frac{dV}{V}\left[1 - \left(\frac{V}{c}\right)^2\right] \tag{8.28}$$

where c is the velocity of sound. From Eq. (8.28),

In *subsonic* velocity, $V < c$. If the area increases, the velocity decreases, same as for incompressible flow.
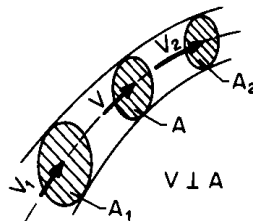
FIGURE 8.21 Steady-flow tube.

In *sonic* velocity, $V = c$. Sonic velocity can exist only where the change in area is zero, i.e., at the end of a convergent passage or at the exit of a constant-area duct.

In *supersonic* velocity, $V > c$. If area increases, the velocity increases, the reverse of incompressible flow. Also, supersonic velocity can exist only in the expanding portion of a passage after a contraction where sonic velocity existed.

2. Bernoulli's Equation (Law of Conservation of Energy)

1. For *nonviscous fluid* (no friction),

$$\frac{p_1}{\rho} + gz_1 + \frac{V_1^2}{2} = \frac{p_2}{\rho} + gz_2 + \frac{V_2^2}{2} \tag{8.29}$$

where
$$\begin{aligned} p/\rho &= \text{pressure energy per unit mass} \\ gz &= \text{potential energy per unit mass} \\ V^2/2 &= \text{kinetic energy per unit mass} \end{aligned}$$

Figure 8.22 gives energy relations for Eq. (8.29). The energy grade line at any point is $\Sigma(p/\rho g + V^2/2g + z)$ and the hydraulic grade line is $\Sigma(p/\rho g + z)$. In
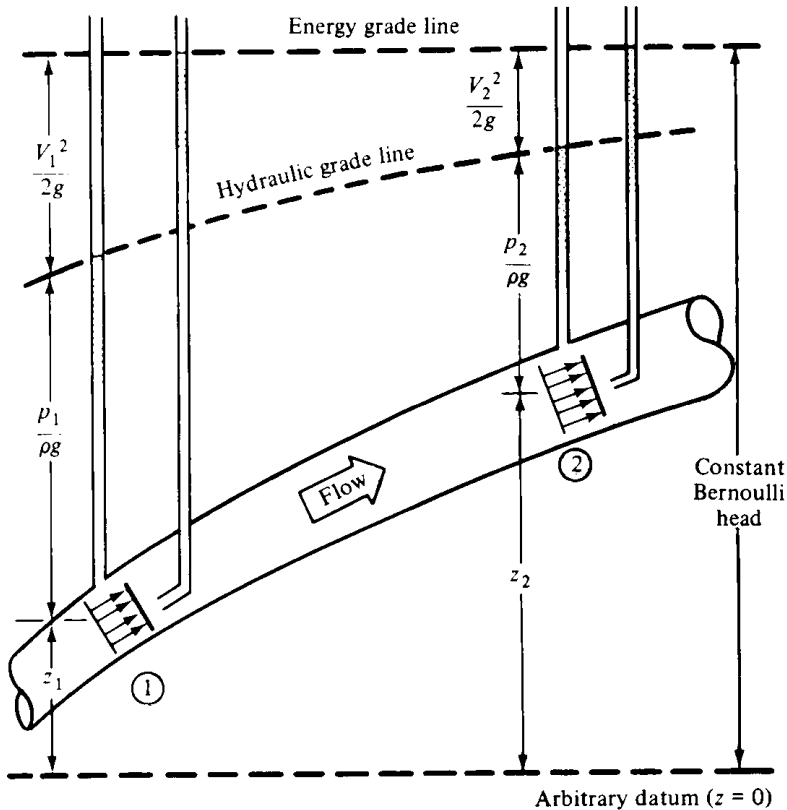
FIGURE 8.22   Hydraulic and energy grade lines for frictionless flow in a duct. (*From White,[2] p. 164.*)

hydraulic practice, each type of energy is referred to as a *head*. The *static pressure head* is $p/\rho g$. The *velocity head* is $V^2/2g$, and the *potential head* is z.

2. For *real fluid* (including losses),

$$\frac{p_1}{\rho} + gz_1 + \frac{V_1^2}{2} = \frac{p_2}{\rho} + gz_2 + \frac{V_2^2}{2} + \bar{h}_{1-2} \tag{8.30}$$

where $\bar{h}_{1-2}$ = resistance losses along path from 1 to 2 per unit mass. The energy loss between sections $h_{1-2} = \bar{h}_{1-2}/g$ is called the *friction head*.

Note: The power P of a hydraulic machine is

$$P = \dot{m}w_{1-2} \tag{8.31}$$

and work per unit mass is

$$w_{1-2} = \frac{1}{\rho}(p_2 - p_1) + g(z_2 - z_1) + \frac{1}{\rho}(V_2^2 - V_1^2) - \bar{h}_{1-2} \tag{8.32}$$

For pumps, $w_{1-2} < 0$. For hydraulic machines, $w_{1-2} > 0$.

EXAMPLE 8.1 Liquid flows from a large tank, open to the atmosphere, through a small, well-rounded aperture into the atmosphere, as shown in Fig. 8.23a. Neglecting losses (e.g., assume frictionless fluid), find an expression for the velocity of the efflux from the tank. Assume that the pressure at 2, the aperture exit, is atmospheric.
From Eq. (8.29),

$$\frac{p_1 - p_2}{\rho} + g(z_1 - z_2) = gh = \frac{V_2^2 - V_1^2}{2}$$

For this case, $p_1 = p_2 = p_a$ and V_1^2 is approximately zero, since the cross-sectional area of the tank A_1 is much greater than the aperture A_2. Therefore,

$$V_2 = \sqrt{2gh}$$

Next, place a 90° elbow at the tank exit (Fig. 8.23b). Determine how high the water jet will reach. Again, assume frictionless flow, and take aperture 2 to be distance h below the water surface.

(a)   (b)
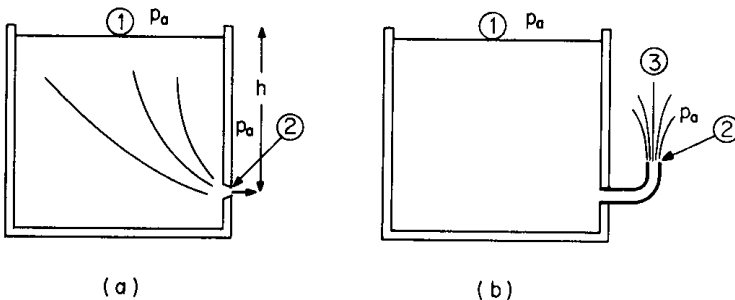
FIGURE 8.23   *(Adapted from Probstein[4] and Sonin.[6])*

The velocity at the exit remains unchanged in magnitude and is given by (for a frictionless elbow) the last equation. Next, use Bernoulli's equation between points 2 and 3. At point 3, the position of maximum elevation of the jet, the velocity of the jet V_3 is zero. Thus

$$\frac{p_3 - p_2}{\rho} + g(z_3 - z_2) = gz = \frac{V_2^2 - V_3^2}{2}$$

Hence $p_2 = p_3 = p_a$; therefore, $gz = V_2^2/2$. With $V_2^2 = 2gh$, it follows that $gz = gh$ and $z = h$. Thus, if there are no losses, the water jet reaches the initial level of the water in the tank. With losses, the maximum elevation of the water jet will be below the water level in the tank.

EXAMPLE 8.2 The siphon in Fig. 8.24 is filled with water and is discharging at 150 L/s (1 m³ = 1000 L). Find the losses from point 1 to point 3 in terms of velocity head. Find the pressure at point 2 if two-thirds of the losses occur between points 1 and 2. From Eq. (8.30) for points 1 and 3,

$$\frac{V_1^2}{2g} + \frac{p_1}{\rho g} + z_1 = \frac{V_3^2}{2g} + \frac{p_3}{\rho g} + z_3 + h_{1-2}$$

or for $p_1 = p_3 = p_a$, $V_1 \simeq 0$, and $z_1 - z_3 = 1.5$ m,

$$1.5 = \frac{V_3^2}{2g} + h_{1-2}$$

So
$$V_3 = \frac{Q}{A} = \frac{150}{(3.14)(0.1^2)(1000)} = 4.77 \text{ m/s}$$

Substituting for V_3 in the last equation,
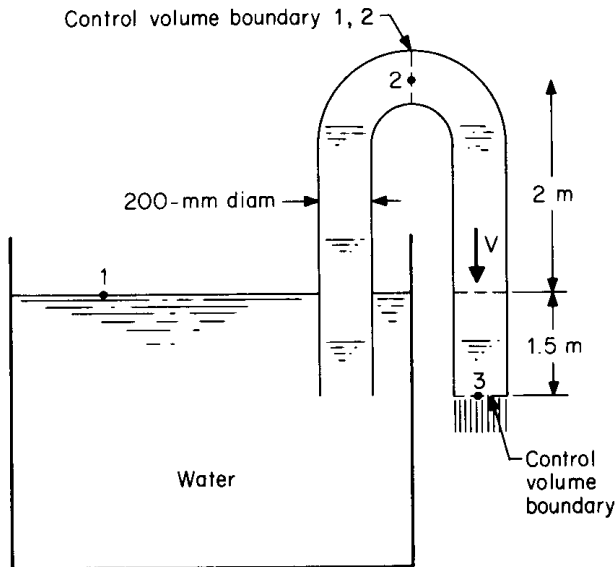
FIGURE 8.24 Siphon. (*From Streeter and Wylie,*[1] *p. 114.*)

$$h_{1-2} = 0.34 \text{ m} \cdot \text{N/N}$$

Pressure at point 2 is

$$0 = 1.16 + \frac{p_2}{\rho g} + 2 + \frac{2}{3} h_{1-2}$$

or $\qquad p_2 = -33.2 \text{ kPa}$

3. Linear Momentum Equation. For a fluid flowing through a stationary reference volume, the following linear momentum equation is valid:

$$\Sigma F = \dot{m}(V_2 - V_1) \qquad (8.33)$$

This equation is derived from Newton's second law and is used to calculate the resultant force exerted on a solid boundary by a fluid stream. Actually, ΣF is the sum of the forces acting on the fluid contained in the reference volume. These can be volume forces (e.g., weight), pressure forces, and friction forces.

V_2 is the exit velocity of the fluid leaving the reference control volume, V_1 is the entrance velocity of the fluid entering the reference volume, and $\dot{m}$ is the mass flow per unit time. In selecting the arbitrary control volume, it is generally advantageous to take the surface normal to velocity whenever it cuts across the flow.

As a scalar equation for any direction, such as the x direction, Eq. (8.33) becomes

$$\Sigma F_x = \dot{m}(V_{x_2} - V_{x_1}) \qquad (8.34)$$

These equations are restricted to steady flow in the simple forms given above.

EXAMPLE 8.3 We wish to evaluate the force coming onto the reducing elbow shown in Fig. 8.25 as a result of an internal steady flow of fluid. The flow may be assumed to be one-dimensional. The average values of the flow characteristics at the inlet and outlet are known, as is the geometry of the reducer. A control volume chosen at the interior of the reducer will enable us to relate known quantities at the inlet and outlet with the force **R** on the fluid from the reducer wall.

The reaction to this latter force is the quantity to be computed. This is shown in the diagram, where the control volume has been separately illustrated. All the
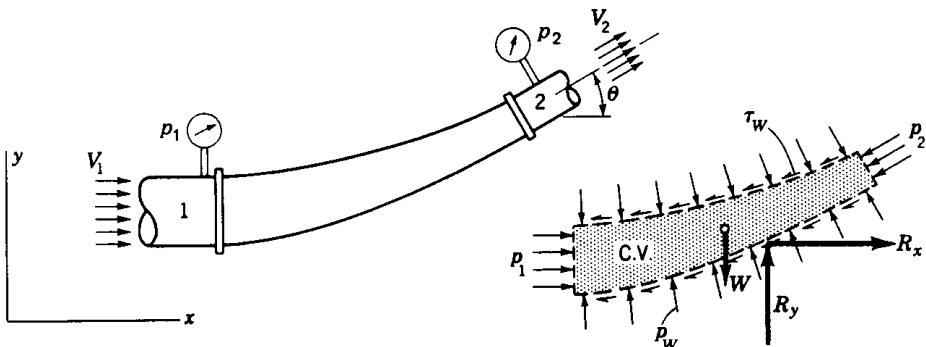
FIGURE 8.25 *(From Shames,[7] p. 98.)*

forces acting on the fluid in the control volume at any time t have been designed. The surface forces include the effects of pressures p_1 and p_2 at the entrance and exit of the reducer as well as distributions of normal and shear stresses p_w and τ_w, whose resultant force is **R**. The body force is simply the weight of the fluid inside the control volume at time t and is indicated in the diagram as W. The x and y components of the resultant force on the fluid may be expressed as

$$\Sigma F_x = p_1 A_1 - p_2 A_2 \cos \theta + R_x$$

$$\Sigma F_y = -p_2 A_2 \sin \theta - W + R_y$$

where R_x and R_y are the net force components of the reducer wall on the fluid. R_x and R_y, being unknown, have been selected positive.

The continuity equation for this control volume, meanwhile, is

$$\rho_1 V_1 A_1 = \rho_2 V_2 A_2 = \dot{m}$$

Now, substituting the preceding results into the momentum equations in x and y directions, we get

$$p_1 A_1 - p_2 A_2 \cos \theta + R_x = \dot{m}(V_2 \cos \theta - V_1) - p_2 A_2 \sin \theta - W + R_y$$

$$= \dot{m}(V_2 \sin \theta)$$

One may now solve for R_x and R_y. Changing the sign of these results will then give the force components on the elbow from the fluid.

EXAMPLE 8.4   In Fig. 8.26, the fluid is deflected through the angle θ by the fixed vane. To find the force components F_x and F_y exerted on the fluid by the vane, the momentum equation is applied to the free body of fluid between sections 1 and 2,

$$-F_x = \dot{m}(V_0 \cos \theta - V_0)$$

or
$$F_x = \dot{m}V_0(1 - \cos \theta)$$

In the y direction (neglecting W),

$$F_y = \dot{m}V_0 \sin \theta$$

where $\dot{m} = Q\rho$ (where Q is volume flow rate).

For a vane moving in the x direction with velocity u (Fig. 8.27),

$$F_x = \dot{m}(V_0 - u)(1 - \cos \theta)$$

$$F_y = \dot{m}(V_0 - u) \sin \theta$$

The force acting on the vanes in the direction of motion is equal and opposite to F_x and yields the power given up to the wheel when multiplied by u (a series of vanes is mounted on a wheel such that a vane always intercepts the jet):

$$P = \dot{m}(V_0 - u)u(1 - \cos \theta)$$

Note that P has maximum value for $\theta = 180°$ and $u = V_0^2/2$. The impulse turbines are designed based on these relations.

EXAMPLE 8.5 Find the output power P for the aircraft jet engine sketched in Fig. 8.26a if:

V = jet velocity relative to aircraft
U = aircraft velocity
$\dot{m}$ = mass flow rate of air
$\dot{m}_f$ = mass flow rate of fuel

Since thrust $T = \dot{m}U - (\dot{m} + \dot{m}_f)V$, output power $P = TU = \dot{m}U^2 - (\dot{m} + \dot{m}_f)UV$.

EXAMPLE 8.6.* Analyze the water jet boat when a) the water enters the side of the boat and b) water enters the front of the boat (see sketch in Fig. 8.26b).
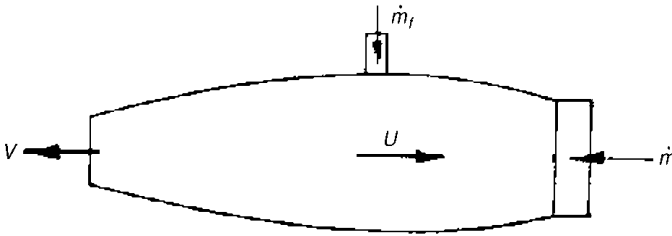
FIGURE 8.26a   Sketch of aircraft jet engine.
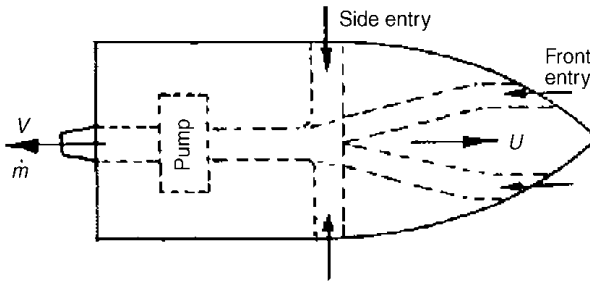
FIGURE 8.26b   Sketch of the water jet boat.
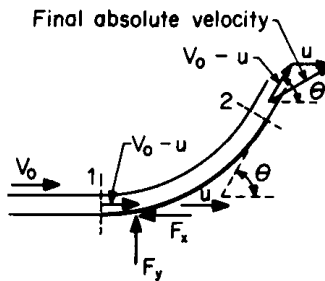
FIGURE 8.27   Deflection of jet by a moving vane. (*Adapted from Probstein[4] and Sonin.[6]*)

---

*Source: Refs. 4, 18.

If

V = jet velocity relative to boat
U = boat velocity

let
$$r = \frac{U}{V}$$

$\dot{m}$ = mass flow rate of jet

a. If *water enters side of boat,*

$$\text{thrust } F = \dot{m}V(1 - r)$$

and
$$\text{pump water power } P = \dot{m}\,\frac{V^2}{2}$$

$$\text{Efficiency } \eta = 2r(1 - r); \qquad \eta_{\text{max}} = 0.5, \text{ at } r = 0.5.$$

b. If *water enters from front of boat,*

$$\text{thrust } F = \dot{m}V(1 - r)$$

and
$$\text{pump power } P = \dot{m}\,\frac{(V^2 - U^2)}{2} = \dot{m}\,\frac{V^2}{2}(1 - r^2)$$

$$\text{Efficiency } \eta = \frac{2r}{(1 + r)}$$

$\eta = 0.667$, for $r = 0.5$.

$\eta = 1.0$, for $r = 1.0$.

Output power for both cases

$$P_\text{o} = \dot{m}V^2 r(1 - r) \quad \text{and} \quad P_{\text{o max}} = \dot{m}\,\frac{V^2}{4}, \text{ at } r = 0.5.$$

4. Moment-of-Momentum Equation. The *moment of momentum,* or *angular momentum,* is useful in analyzing the flow in steady rotating channels, such as in a centrifugal pump or a reaction turbine. Since momentum is a vector quantity, its moment about an axis may be determined analogous to the moment of a force about the axis. The component of the momentum in the radial direction contributes nothing to the moment of momentum. The equation is

$$T = \dot{m}(V_{t_2}r_2 - V_{t_1}r_1) \tag{8.35}$$

where T is the resultant torque on the fluid, $\dot{m}$ is mass flow having its angular momentum changed, and V_t is the tangential component of velocity, with subscript 2 referring to the final condition and subscript 1 referring to the initial condition. Expressing the tangential velocities in terms of the absolute velocities (see Fig. 8.28) at entrance and exit, we have
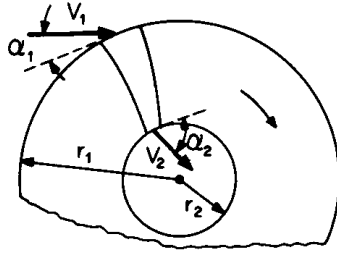
FIGURE 8.28 Turbine runner (rotating chan-
nel). (*Adapted from Probstein*[4] *and Sonin.*[6])

$$T = \dot{m}(V_2 r_2 \cos \alpha_2 - V_1 r_1 \cos \alpha_1) \tag{8.36}$$

The power obtained from the fluid is then

$$P = T \cdot \omega \tag{8.37}$$

where ω is angular velocity.

BOUNDARY-LAYER FLOWS

1. Laminar Flow. The *boundary-layer thickness* is defined as the locus of points where the velocity u parallel to the plate reaches 99 percent of the external velocity U. The boundary-layer thickness for incompressible flow over smooth, flat plates (Fig. 8.29) may be calculated from the following relation:

$$\frac{\delta}{x} = \frac{5.0}{\mathrm{Re}_x} \tag{8.38}$$

where $\mathrm{Re}_x = \rho U x / \mu$ is called the *local Reynolds number* of the flow along the flat plate.

The *fluid shear stress* at the surface of the plate is

$$\tau_w = \mu \left(\frac{\partial u}{\partial y} \right)_{y=0} \tag{8.39}$$

or

$$\tau_x = 0.332 \mu \frac{U}{x} \sqrt{\mathrm{Re}_x} \tag{8.40}$$

The local *skin-friction coefficient* C_{f_x} is defined by

$$C_{f_x} = \frac{\tau_w}{\frac{1}{2} \rho U^2} \tag{8.41}$$

and for laminar boundary-layer flow over a flat plate, it is

$$C_{f_x} = \frac{0.664}{\sqrt{\mathrm{Re}_x}} \tag{8.42}$$

The wall shear stress varies with x, and in order to obtain the total force exerted
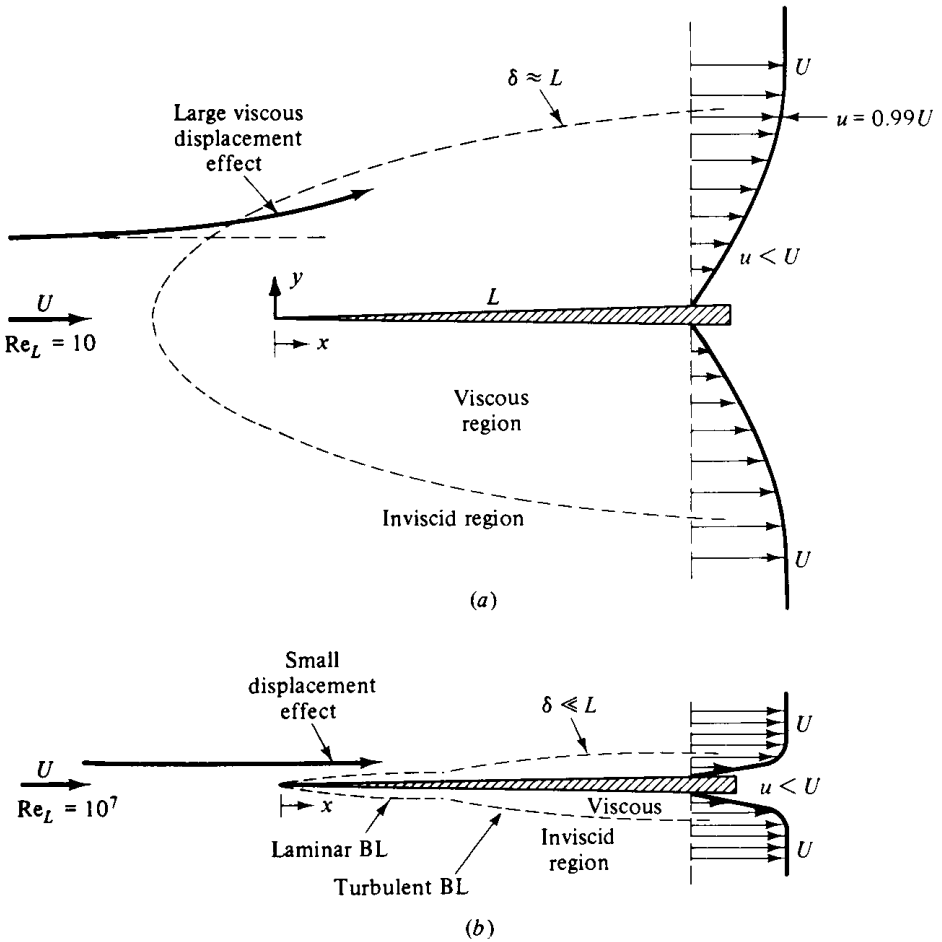
FIGURE 8.29 Comparison of flow past a sharp, flat plate at low and high Reynolds numbers: (*a*) laminar, low Reynolds number flow; (*b*) high Reynolds number flow. (*From White,*[2] *p. 401.*)

by the fluid passing over the plate, we must integrate. The *total drag force* F_D on a plate of length L is

$$F_D = \int_0^L \tau_w(x)b \ dx = \frac{0.664\rho U^2 bL}{\sqrt{\mathrm{Re}_L}} \tag{8.43}$$

where $\mathrm{Re}_L = \rho UL/\nu$ and b is the width of the plate.

The skin-friction drag coefficient C_{D_f} is defined by

$$C_{D_f} = \frac{F_D}{\frac{1}{2}\rho U^2 bL} = \frac{1.328}{\sqrt{\mathrm{Re}_L}} \tag{8.44}$$

Note that C_{f_x} is a local skin-friction coefficient and C_{D_f} is a coefficient expressing the integrated drag over a length L of a surface.

2. Turbulent Flow. A critical Re_x for transition of laminar to turbulent boundary layer is approximately 5×10^5. The turbulent boundary thickness is

$$\frac{\delta}{x} \simeq \frac{0.37}{(Re_x)^{1/5}} \tag{8.45}$$

The effect of turbulence is to bring about a more uniform profile than would be obtained for the laminar boundary layer (Fig. 8.29). It has been found that a profile based on

$$\frac{u}{U} = \left(\frac{y}{\delta}\right)^{1/7} \tag{8.46}$$

gives a very good correlation with experimental data over a wide range of turbulent Reynolds numbers.

The $\frac{1}{7}$-power law is not valid in the immediate vicinity of the wall, since in this region there exists the laminar sublayer in which turbulence is damped out by the wall. It is therefore necessary to rely on the experimental data to obtain an expression for τ_w. An equation that yields good agreement with data for

$$5 \times 10^5 < Re_x < 10^7$$

is the Blasius resistance formula, that is,

$$\frac{\tau_w}{\rho U^2} = 0.0225 \left(\frac{\nu}{U\delta}\right)^{1/4} \tag{8.47}$$

From Eqs. (8.45) and (8.47),

$$\tau_w = \frac{0.029 \rho U^2}{(Re_x)^{1/5}} \tag{8.48}$$

Also, analogously to Eqs. (8.42), (8.43), and (8.44),

$$C_{f_x} = \frac{0.058}{(Re_x)^{1/5}} \tag{8.49}$$

$$F_D = \frac{0.036 \rho U^2 bL}{(Re_L)^{1/5}} \tag{8.50}$$

$$C_{D_f} = \frac{0.074}{(Re_L)^{1/5}} \tag{8.51}$$

Figure 8.30 includes values for C_{D_f} as function of Re_L.

3. Flow Over Immersed Bodies. In analyzing uniform flow past a body surface, the engineer is most often interested in the resultant fluid forces acting on the body. Such forces can be divided into two components: *lift force,* acting upward and normal to the approach flow, and *drag force,* acting in the same direction as the approach flow, as shown in Fig. 8.31. For uniform flow past a finite body, such as

FIGURE 8.30 The drag law for smooth plates. (*From Streeter and Wylie,[1] p. 217.*)
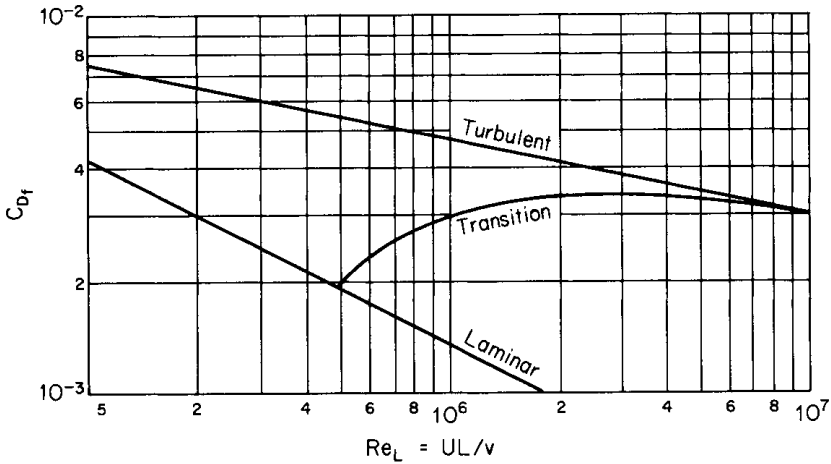
Laminar $C_{D_f} = \dfrac{1.328}{\sqrt{Re_L}}$, transition $C_{D_f} = \dfrac{0.074}{Re_L^{1/5}} - \dfrac{1700}{Re_L}$, turbulent $C_{D_f} = \dfrac{0.074}{Re_L^{1/5}}$
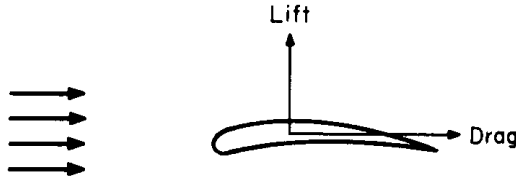
FIGURE 8.31 Resultant fluid forces acting on the body. (*Adapted from Probstein[4] and Sonin.[6]*)

an airfoil or cylinder, the boundary layer is subjected to both positive and negative pressure gradients (Fig. 8.18), and separation of the flow from the body surface is a result of a positive pressure gradient acting on the boundary layer.

When analyzing a flow over immersed bodies, it is assumed that the Reynolds numbers are large enough so that the flow can be divided into two parts: a viscous region, consisting of a thin boundary layer, separation region, and wake, and a nonviscous region, where the effects of viscosity can be neglected, as shown in Fig. 8.32.

Pressure distribution on the surface of a cylinder immersed in uniform steady flow is shown in Fig. 8.33. Owing to boundary-layer separation, the average pressure on the rear half of the cylinder is less than that on the front half. On the rear half of the cylinder, the pressure on the cylinder surface does not recover to the full stagnation pressure as in the case of potential flow (Fig. 8.33). Thus, by integrating the pressure forces over the cylinder surface for the real-fluid case, we obtain a pressure drag on the cylinder. It is obvious that pressure drag can be reduced by shaping or streamlining the body so as to place the point of boundary-layer separation in the vicinity of the trailing edge of the body. In the two preceding sections, relations for skin-friction drag are given. The total drag on a body is due
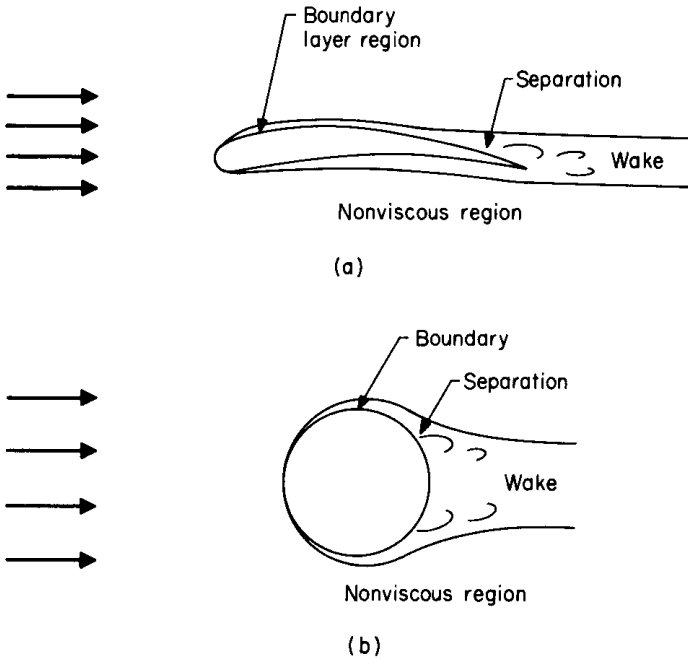
FIGURE 8.32 (*a*) Real fluid flow about an airfoil. (*b*) Real fluid flow about a circular cylinder. (*Adapted from Probstein[4] and Sonin.[6]*)

to the sum of skin-friction drag ("surface" drag) and pressure drag ("shape" drag). For uniform flow over a flat plate aligned with the flow direction (Fig. 8.34*a*) with no pressure gradients, the entire drag is due to skin friction. For a flat plate normal to the flow direction (Fig. 8.34*b*), the entire drag is pressure drag. For a circular cylinder (Fig. 8.33), over 90 percent of the drag is pressure drag, with only a small fraction due to skin friction.

The drag of a bluff body is expressed in terms of a nondimensional parameter C_D, called the *drag coefficient:*

$$C_D = \frac{F_D}{\frac{1}{2}\rho U^2 A} \tag{8.52}$$

where A is the projected frontal area of the bluff body normal to the flow direction. Values of the drag coefficient for circular cylinders and spheres are given in Figs. 8.35 and 8.36. Table 8.1 includes C_D values for several shapes.

An analytical solution is available for flow about a sphere at Reynolds numbers less than 1 in which the inertial forces can be neglected in comparison to the viscous forces. This analysis, initially derived by Stokes, showed that for $Re_D < 1$,

$$C_D = \frac{24}{Re_D} \tag{8.53}$$

Equation (8.53) is useful for the design of settling basins for separating small solid particles from fluids ($Re_D = \rho U D / \mu$).

A typical pressure distribution for a symmetrical airfoil is given in Fig. 8.37. As the angle of attack α of the foil is increased from zero, the fluid moving over the
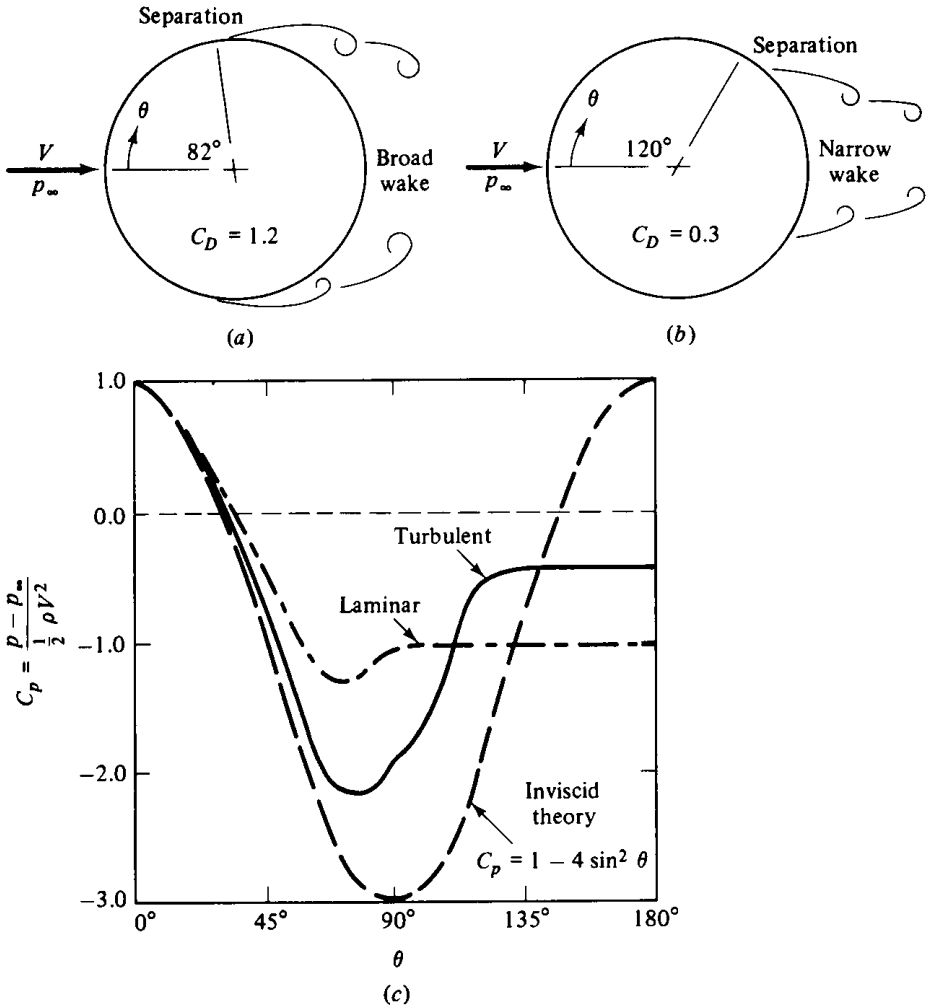
FIGURE 8.33 Flow past a circular cylinder: (*a*) laminar separation; (*b*) turbulent separation; (*c*) theoretical and actual surface-pressure distributions. (*From White*[2] *p. 431.*)
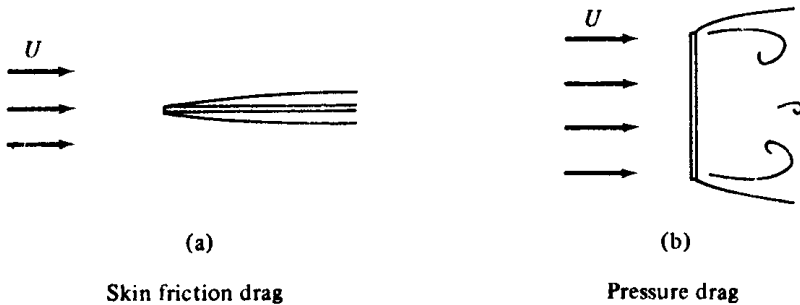
FIGURE 8.34 (*a*) Flat plate aligned with flow. (*b*) Flat plate normal to flow. (*Adapted from Probstein*[4] *and Sonin.*[6])
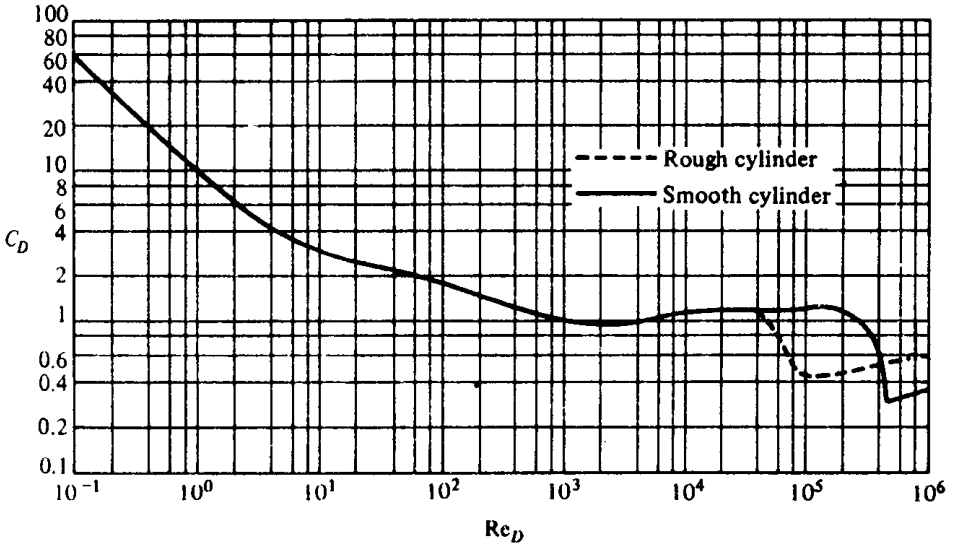
FIGURE 8.35   Drag coefficients of circular cylinders. (*Adapted from Schlichting,*[11] *p. 17.*)
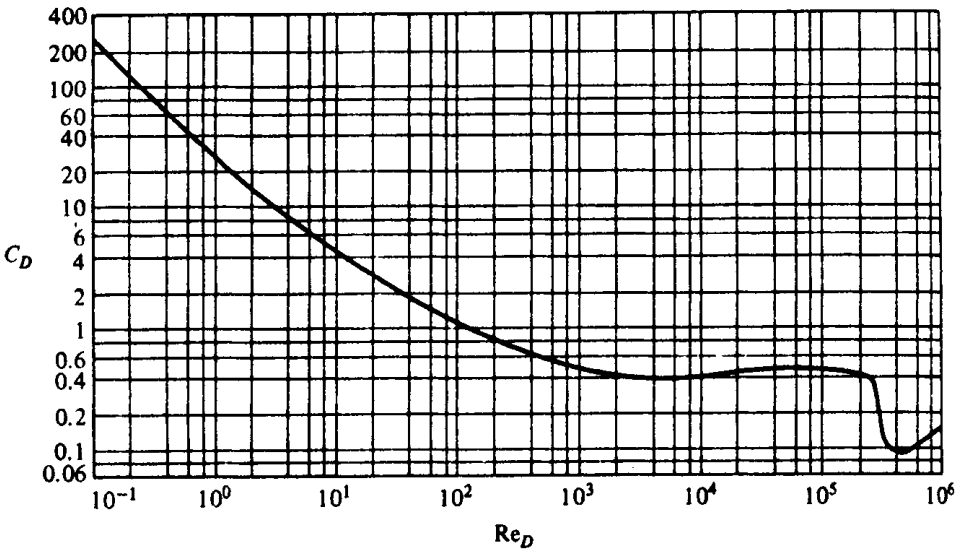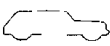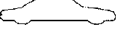
FIGURE 8.36   Drag coefficients of spheres. (*Adapted from Schlichting,*[11] *p. 17.*)

top surface must accelerate more rapidly, yielding a more negative pressure coefficient C_p. The fluid traveling over the lower surface undergoes a much more gradual acceleration. The resultant difference in pressure between the upper and lower surfaces yields a positive lift force on the foil, generally expressed in terms of a lift coefficient C_L with

$$C_L = \frac{F_L}{\frac{1}{2}\rho U^2 A} \tag{8.54}$$

For small angles of attack, the lift coefficient varies linearly with the angle of

TABLE 8.1 Drag Coefficients for Several Shapes

2-dimensional shape		Reynolds number	C_D
Circular cylinder	○	10^4 to 10^5	1.2
Semitubular	(	4×10^4	1.2
Semitubular	)	4×10^4	2.3
Square cylinder	□	3.5×10^4	2.0
Flat plate	▯	$10^4 \times 10^4$	1.98
Elliptical cylinder	⬭ 2:1	10^5	0.46
Elliptical cylinder	⬭ 8:1	2×10^5	0.20

3-dimensional shape		Reynolds number	C_D
Sphere	○	10^4 to 10^5	0.47
Hemisphere	◖	10^4 to 10^5	0.42
Hemisphere	◗	10^4 to 10^5	1.17
Cube	□	10^4 to 10^5	1.05
Cube	◇	10^4 to 10^5	0.80
Rectangular plate with $\dfrac{\text{Length}}{\text{Width}} = 5$	▯	10^3 to 10^5	1.20
High drag car		$\sim 10^5$	0.55
Medium drag car		$\sim 10^5$	0.45
Low drag car		$\sim 10^5$	0.30

Sources: Adapted from Refs. 1, 2, 7 to 11 and 18.

attack, with $C_L = 0$ for $\alpha = 0$ for a symmetrical airfoil. Figure 8.38 gives typical values of lift and drag coefficients for an airfoil.

Note on Aerodynamic Drag of Automobile. The resistance of a vehicle to motion is made up of aerodynamic drag and rolling resistance. In addition, so called "gradient force" is to be added.

The aerodynamic drag force:
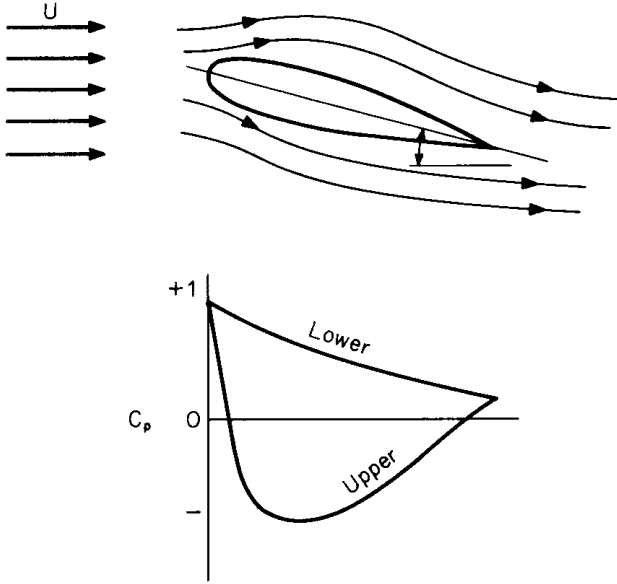
FIGURE 8.37 Flow past an airfoil. (*Adapted from Probstein*[4] *and Sonin.*[6])
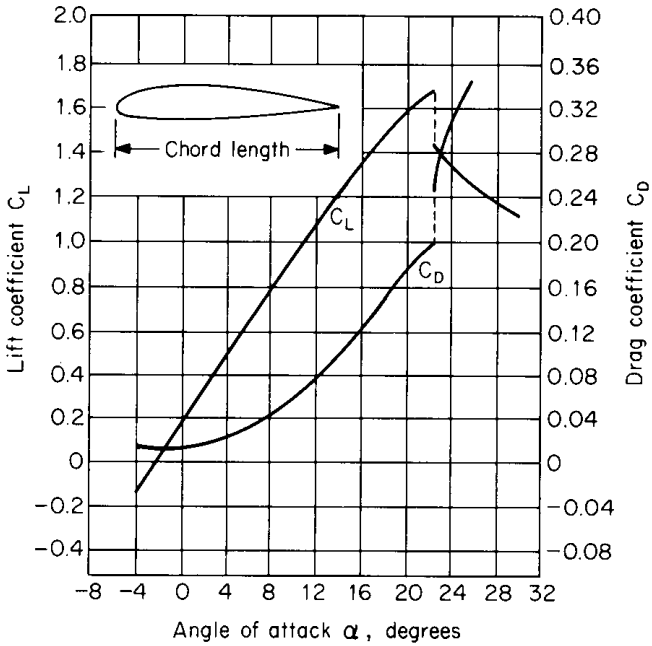
FIGURE 8.38 Typical lift and drag coefficients for an airfoil. (*from Streeter and Wylie,*[1] *p. 224.*)
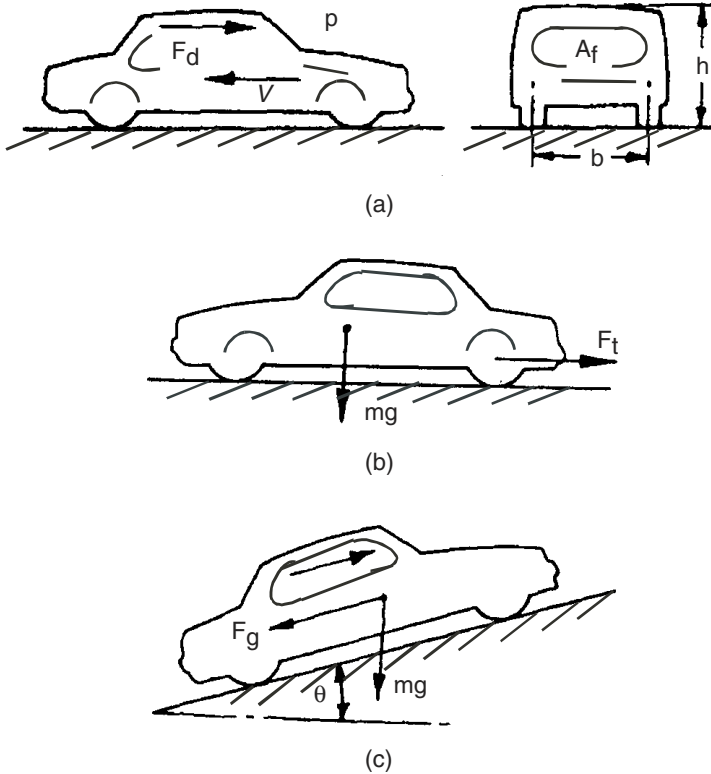
FIGURE 8.38a, b, c   Resistance forces on a vehicle.

TABLE 8.1a   Typical Values of Drag Coefficient of Motor Vehicles

	C_d		C_d
Sports car, sloping rear	0.2–0.3	Motorcycle and rider	1.8
Saloon, stepped rear	0.4–0.5	Long stream-lined body	0.1
Convertible, open top	0.6–0.7		
Bus	0.6–0.8		
Truck	0.8–1.0		

Sources: Refs. 11, 18.

$$F_d = C_D A_f \rho \frac{V^2}{2}$$

where C_D = drag coefficient
   A_f = frontal area (approx. $0.9b \cdot h$ m²) (see Fig. 8.38a)
   ρ = density of air (≈ 1.2 kg/m³)
   V = velocity (m/s)

Typical values of drag coefficient of motor vehicles are given in Table 8.1a. The rolling resistance force

$$F_r = C_r mg$$

where the coefficient of rolling resistance[4,18]

$$C_r \simeq 0.005 + \frac{0.01}{p}\left[1 + \left(\frac{v}{100}\right)^2\right]$$

m = mass of vehicle
p = tire pressure (bars)
v = speed (km/h).

The "gradient force" $F_g = m \cdot g \cdot \sin\theta$ (see Fig. 8.38c).
Therefore, the total force of the vehicle, F_t, is made up of:

$$F_t = F_d + F_r + F_g$$

FLOW IN PIPES

An internal flow in a duct is constrained by the bounding walls, and the viscous effects will grow and meet and permeate the entire flow. Figure 8.17 shows an internal flow in a long duct. There is an *entrance region* where a nearly inviscid upstream flow converges and enters the tube. Viscous boundary layers grow downstream, retarding the axial flow $u(r, x)$ at the wall and thereby accelerating the center-core flow to maintain the incompressible continuity requirement

$$Q = \int_A u\, dA = \text{constant}$$

At a finite distance from the entrance, the boundary layers merge and the inviscid core disappears. The tube flow is then entirely viscous, and the axial velocity adjusts slightly further until at $x = L_e$ it no longer changes with x and is said to be *fully developed,* that is, $u \approx u(r)$ only.

Downstream of $x - L_e$ the velocity profile is constant, the wall shear is constant, and the pressure drops linearly with x for either laminar or turbulent flow.

Reynolds number is the only parameter affecting entrance length. The accepted correlations for L_e are

$$\frac{L_e}{D} \approx 0.06\text{Re} \qquad \text{(laminar)} \tag{8.55}$$

$$\frac{L_e}{D} \approx 4.4\text{Re}^{1/6} \qquad \text{(turbulent)} \tag{8.56}$$

An accepted design value for pipe-flow transition from laminar to turbulent is

$$\text{Re}_\text{crit} \approx 2300 \tag{8.57}$$

However, for the Reynolds-number range between 2000 to 4000, the flow is unstable and this zone is called the *transition zone.*

The pressure loss in a pipe flow may be expressed in terms of pipe-head loss [see Eq. (8.30), $h_{1-2} = h_f$]:

$$h_f = \frac{\Delta p}{\rho g} = f \frac{L}{D} \frac{V^2}{2g} \tag{8.58}$$

where the dimensionless parameter f is called the *Darcy friction factor,* given in terms of wall shear stress

$$\frac{f}{4} = \frac{\tau_w}{\frac{1}{2}\rho V^2} \tag{8.59}$$

where V is mean flow velocity and D is tube diameter.

If the duct is not circular, an equivalent hydraulic diameter D_h is used, defined as

$$D_h = \frac{4 \times \text{cross-sectional area of flow}}{\text{perimeter wetted by fluid}} \tag{8.60}$$

In *laminar flow,* the flow resistance is due to viscous forces only, so that it is independent of the pipe surface roughness. For *laminar flow,*

$$f = \frac{64}{\text{Re}} \tag{8.61}$$

And for *turbulent flow,* the friction factor for Re > 4000 is computed using the Colebrook equation, that is,

$$\frac{1}{\sqrt{f}} = -2 \log \left(\frac{2.51}{\text{Re } \sqrt{f}} + \frac{\epsilon/D}{3.7} \right) \tag{8.62}$$

Figure 8.39 (Moody diagram) is a graphical presentation of this equation. In order to determine a value of the friction factor f from the Moody diagram, a knowledge of relative roughness is necessary. Typical values of roughness for various types of pipe are shown in Table 8.2. For noncircular pipes, D is replaced with D_h.

1. Losses in Pipe Fittings and Valves.   In addition to losses due to wall friction in a piping system, there are also losses of mechanical energy and pressure due to flow through valves or fittings (called *minor losses*). The minor losses can be treated as an equivalent frictional loss by expressing the pressure drop due to fittings as

$$\Delta p = -\frac{1}{2} K \rho V^2 \tag{8.63}$$

or in terms of heat loss,

$$h_m = -K \frac{V^2}{2g} \tag{8.64}$$

where K is the loss (resistance) coefficient. Representative values of K are given in Tables 8.3 and 8.4.

In piping systems, three basic types of calculations are encountered:
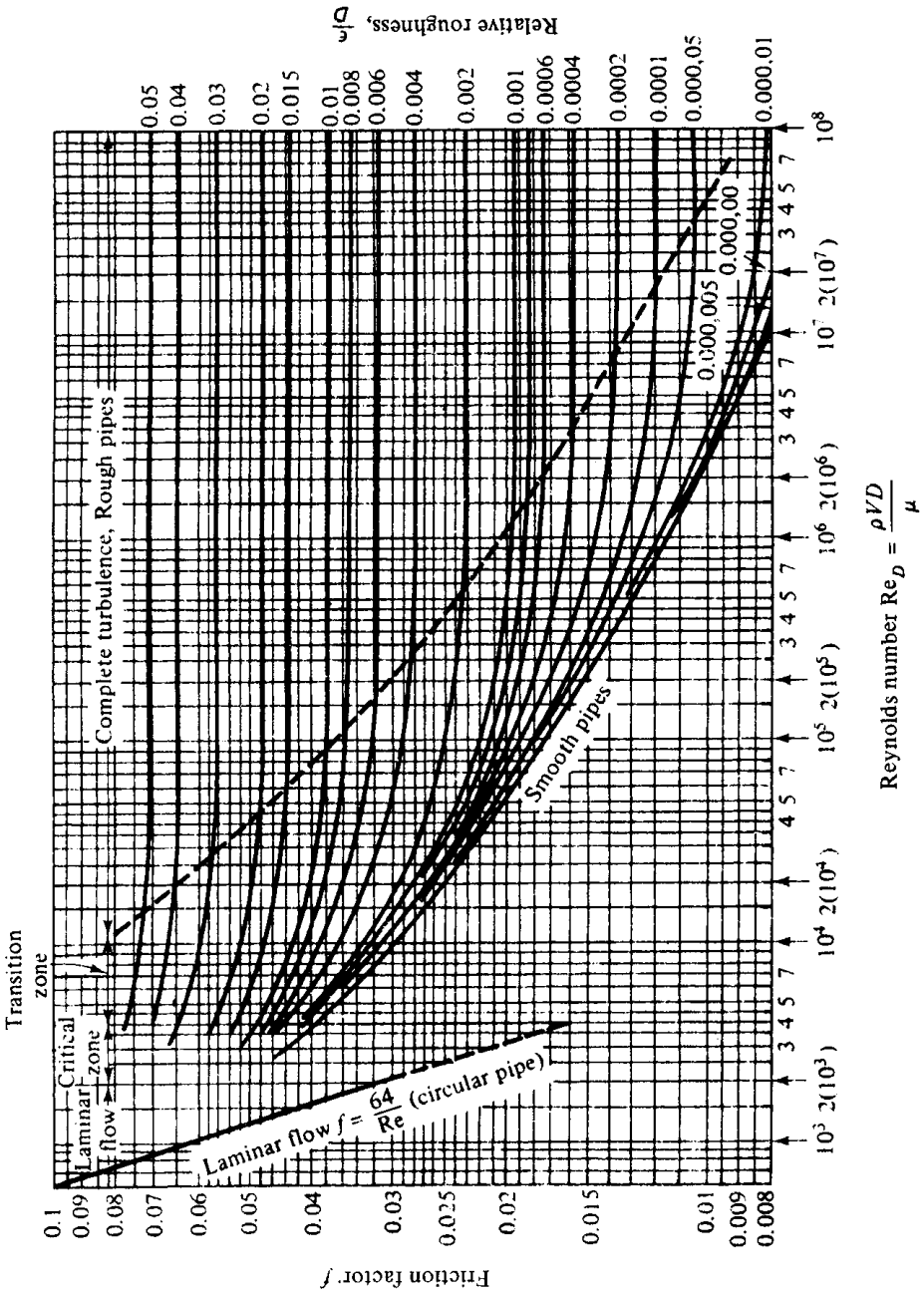
FIGURE 8.39 Friction factors for pipe flow. (*From Moody.*[12])

TABLE 8.2 Roughness of Various Types of Pipe

Type	ϵ (mm)
Glass	Smooth
Asphalted cast iron	0.12
Galvanized iron	0.15
Cast iron	0.26
Wood stave	0.18–0.90
Concrete	0.30–3.0
Riveted steel	1.0–10
Drawn tubing	0.0015

Sources: Adapted from Refs. 1, 2, and 7 to 11.

TABLE 8.3 Values of K for Valves and Fittings
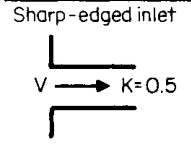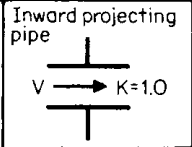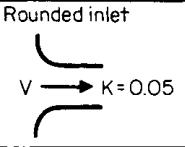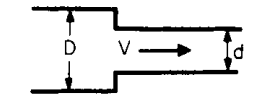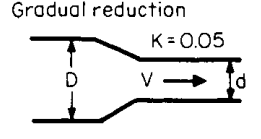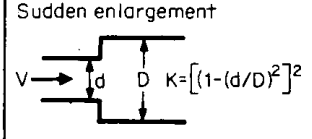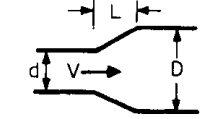
Fitting or valve	K
Standard 45° elbow	0.35
Standard 90° elbow	0.75
Long-radius 90° elbow	0.45
Coupling	0.04
Union	0.04
Gate valve	
Open	0.20
¾ Open	0.90
½ Open	4.5
¼ Open	24.0
Globe valve	
Open	6.4
½ Open	9.5
Tee (along run, line flow)	0.4
Tee (branch flow)	1.5

Sources: Adapted from Refs. 1, 2, and 7 to 11.

1. For a given piping system and flow rate, the pressure drop might be required.

2. For a given piping system and pressure drop, the flow rate might be required.

3. For a given flow and pressure drop, it might be required to design the system, i.e., to determine the necessary pipe diameter.

EXAMPLE 8.7 For the piping shown in Fig. 8.40, determine the pressure p_2. There are 100 m of 15-cm-I.D. cast-iron pipe and 30 m of 7.3-cm-I.D. cast-iron pipe. Water flow rate is 0.01 m³/s, the water temperature is 20°C, and the gage pressure at 1 is 250 kPa.

TABLE 8.4 Values of K for Representative Entrance and Exit Flow Conditions

Sharp-edged inlet	Inward projecting pipe	Rounded inlet
V ⟶ K=0.5	V ⟶ K=1.0	V ⟶ K=0.05

Sudden contraction

D/d	1.5	2.0	2.50	3.0	3.5	4.0
K	0.28	0.36	0.40	0.42	0.44	0.45

Gradual reduction	Sudden enlargement
K=0.05	V→ d D $K=\left[1-(d/D)^2\right]^2$

Gradual enlargement

$$K = K'\left[1-(d/D)^2\right]^2$$

(D-d)/2L	0.05	0.10	0.20	0.30	0.40	0.50	0.80
K'	0.14	0.20	0.47	0.76	0.95	1.05	1.10

Exit loss = (sharp edged, projecting, Rounded), K=1.0

Source: From Avallone and Baumeister,[3] p. 3-58.

FIGURE 8.40 (*Adapted from Probstein*[4] *and Sonin.*[6])

Solution. First write the modified Bernoulli equation [Eq. (8.30)] between 1 and 2, including minor losses:

$$\frac{p_1}{\rho} + \frac{V_1^2}{2} + gz_1 = \left(\frac{fL}{D}\frac{V^2}{2}\right)_{15\text{-cm pipe}} + \left(\Sigma K\frac{V^2}{2}\right)_{15\text{-cm pipe}} + \left(\frac{fL}{D}\frac{V^2}{2}\right)_{7.5\text{-cm pipe}}$$

$$+ \left(\Sigma K\frac{V^2}{2}\right)_{7.5\text{-cm pipe}} + \frac{V_2^2}{2} + gz_2 + \frac{p_2}{\rho}$$

The velocity in the 15-cm pipe is

$$V_{15} = \frac{Q}{A} = \frac{0.01 \text{ m}^3/\text{s}}{(\pi/4)(0.15)^2 \text{ m}^2} = 0.5659 \text{ m/s}$$

and in the 7.5-cm pipe it is

$$V_{7.5} = \frac{Q}{A} = \frac{0.01 \text{ m}^3/\text{s}}{(\pi/4)(0.075)^2 \text{ m}^2} = 2.264 \text{ m/s}$$

At 20°C, for water, $\nu = 1.00 \times 10^{-6} \text{ m}^2/\text{s}$, so

$$\text{Re}_{15} = \frac{(0.5659 \text{ m/s})(0.15 \text{ m})}{1.00 \times 10^{-6} \text{ m}^2/\text{s}} = 8.489 \times 10^4$$

and

$$\text{Re}_{7.5} = \frac{(2.264 \text{ m/s})(0.075 \text{ m})}{1.00 \times 10^{-6} \text{ m}^2/\text{s}} = 1.698 \times 10^5$$

Also,

$$\left(\frac{\epsilon}{D}\right)_{15} = \frac{0.26 \times 10^{-3} \text{ m}}{0.15 \text{ m}} = 0.00173$$

$$\left(\frac{\epsilon}{D}\right)_{7.5} = \frac{0.26 \times 10^{-3} \text{ m}}{0.075 \text{ m}} = 0.00346$$

From the Moody diagram (Fig. 8.39), we obtain

$$f_{15} = 0.025$$

$$f_{7.5} = 0.027$$

The minor losses in the 15-cm pipe are given by

$$(\Sigma K)_{15} = \underset{\substack{\text{standard} \\ \text{elbow}}}{0.75} + \underset{\substack{\text{standard} \\ \text{elbow}}}{0.75} = 1.50$$

For the 7.5-cm pipe,

$$(\Sigma K)_{7.5} = \underset{\text{contraction}}{0.33} + \underset{\text{elbow}}{0.35} + \underset{\text{elbow}}{0.35} + \underset{\text{gate value}}{0.20} = 1.23$$

Substituting into the Bernoulli equation, we have

$$\frac{p_2 - p_1}{\rho} = g(z_1 - z_2) + \frac{V_{15}^2}{2}\left(1 - \frac{fL}{D} - \Sigma K\right)_{15} - \frac{V_{7.5}^2}{2}\left(1 + \frac{fL}{D} + \Sigma K\right)_{7.5}$$

$$= 9.81 \text{ m/s}^2(-10 \text{ m}) + \left(\frac{0.5659^2}{2} \text{ m}^2/\text{s}^2\right)\left(1 - \frac{0.025 \times 100}{0.15} - 1.50\right)$$

$$- \left(\frac{2.264^2}{2} \text{ m}^2/\text{s}^2\right)\left(1 + \frac{0.027 \times 30}{0.075} + 1.23\right)$$

$$= -98.1 \text{ m}^2/\text{s}^2 + (0.1601 \text{ m}^2/\text{s}^2)(-17.17) - (2.563 \text{ m}^2/\text{s}^2)(13.03)$$

$$= -98.1 \text{ m}^2/\text{s}^2 - 2.749 \text{ m}^2/\text{s}^2 - 33.40 \text{ m}^2/\text{s}^2$$

$$= -134.2 \text{ m}^2/\text{s}^2$$

or
$$p_2 - p_1 = -(998.3 \text{ km/m}^3)(134.2 \text{ m}^2/\text{s}^2)$$
$$= -134 \text{ kPa}$$

or
$$p_2 = 250 - 134$$
$$= 116 \text{ kPa (gage)}$$

EXAMPLE 8.8 For the piping system shown in Fig. 8.41, determine the water flow rate. The water temperature is 25°C, with all pipe asphalted cast iron. There are 250 m of 10-cm.I.D. pipe.

Solution. Write the modified Bernoulli equation between the reservoir surface and pipe outlet:

$$\frac{p_0}{\rho} + gz_0 + \frac{V_0^2}{2} = \frac{fL}{D}\frac{V^2}{2} + \Sigma K \frac{V^2}{2} + \frac{p_e}{\rho} + gz_e + \frac{V_e^2}{2}$$

where $p_0 = p_e$ = atmospheric pressure, $z_0 - z_e = 20$ m, and $V_0 = 0$. For this case, with flow velocity the unknown, Reynolds numbers cannot be directly determined; a trial-and-error solution is called for. As a first trial, assume $f = 0.02$. Then,

$$(9.81 \text{ m/s}^2)(20 \text{ m}) = \left[\underbrace{\frac{0.02 \times 250}{0.10}}_{} + \underbrace{(0.5}_{\substack{\text{edged} \\ \text{inlet}}} + \underbrace{3.0}_{\text{4 elbows}} + \underbrace{6.4}_{\substack{\text{glove} \\ \text{valve}}}) + 1 \right] \frac{V_e^2}{2}$$

$$\frac{V_e^2}{2} = \frac{9.81(20)}{50 + 9.9 + 1} \text{ m}^2/\text{s}^2$$

or
$$V_e = 2.53 \text{ m/s}$$

For this velocity,

$$\text{Re} = \frac{VD}{\nu} = \frac{(2.53 \text{ m/s})(0.10 \text{ m})}{1.00 \times 10^{-6} \text{ m}^2/\text{s}} = 2.53 \times 10^5$$

$$\frac{\epsilon}{D} = \frac{0.12 \times 10^{-3} \text{ m}}{0.10 \text{ m}} = 1.2 \times 10^{-3}$$

From the Moody diagram, we obtain $f = 0.022$. Since this value does not agree with our initial trial, we shall now make a second trial, starting with $f = 0.022$.
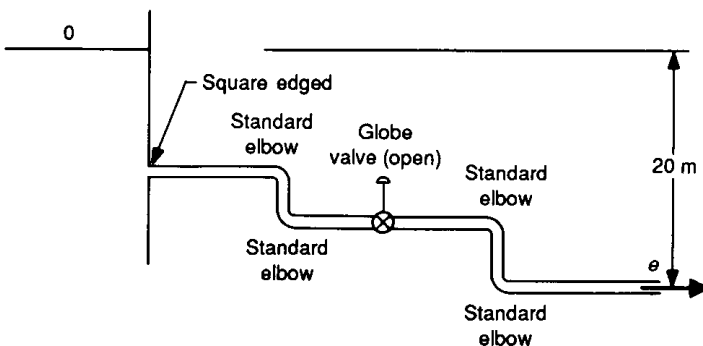
FIGURE 8.41

$$\frac{fL}{D} = \frac{0.022 \times 250}{0.10} = 55$$

$$\frac{V_e^2}{2} = \frac{9.81(20)}{55 + 9.9 + 1}$$

or
$$V = 2.44 \text{ m/s}$$

For this velocity,

$$\text{Re} = \frac{VD}{\nu} = \frac{2.44 \times 0.10}{10^{-6}} = 2.44 \times 10^5$$

From the Moody diagram, we find $f = 0.022$. It can be seen that generally we are able to converge on the correct answer quite rapidly; no more than two trials are required as a rule. The water flow rate is

$$Q = AV$$

$$= \left(\frac{3.14}{4} (0.10)^2 \text{ m}^2\right) 2.44 \text{ m/s}$$

$$= 0.01915 \text{ m}^3/\text{s}$$

$$= 19.15 \text{ L/s}$$

EXAMPLE 8.9 A pump is to be used to supply 5 L/s of water from a reservoir to a point 400 m from the reservoir at the same level as the reservoir surface. Determine the minimum size cast-iron pipe required. The water temperature is 15°C; assume that minor losses can be neglected (see Fig. 8.42). The pump supplies 50 kW of power to the water flow.

Solution. From Eqs. (8.31) and (8.32), we can express the pump work in terms of reservoir surface and outlet conditions:

$$-\frac{1}{\dot{m}} P = \frac{p_e - p_0}{\rho} + g(z_e - z_0) + \frac{fL}{D}\frac{V^2}{2} + \frac{V_e^2}{2}$$

For our case,

Reservoir

Pump
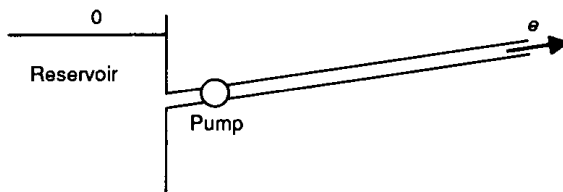
FIGURE 8.42

$$m = \rho Q = (999.2 \text{ kg/m}^3)(5 \times 10^{-3} \text{ m}^3/\text{s}) = 4.996 \text{ kg/s}$$

$$-\frac{1}{\dot{m}P} = -\frac{1}{4.996 \text{ kg/s}}(-50 \times 10^3 \text{ N} \cdot \text{m/s}) = +10{,}010 \text{ m}^2/\text{s}^2$$

Note. P is negative work in a thermodynamic sense, for it represents work done on the fluid.

We now have

$$10{,}010 = \left(\frac{fL}{D} + 1\right)\frac{V^2}{2}$$

where

$$V = \frac{Q}{A} = \frac{5 \times 10^{-3} \text{ m}^3/\text{s}}{(\pi/4)D^2 \text{ m}^2}$$

$$= \frac{6.366 \times 10^{-3}}{D^2} \text{ m/s} \qquad \text{with } D \text{ in meters}$$

$$10{,}010 = \left(\frac{400f}{D} + 1\right)\frac{4.053 \times 10^{-5}}{2D^4}$$

For this case, with D the unknown, the Reynolds number and f cannot be immediately determined. A trial-and-error procedure is required. As a first trial, assume that f = 0.025. Then,

$$10{,}010 = \left[\frac{400(0.025)}{D} + 1\right]\frac{4.053 \times 10^{-5}}{2D^4}$$

$$= \frac{20.26 \times 10^{-5}}{D^5} + \frac{2.026 \times 10^{-5}}{D^4}$$

The second term on the right is small compared to the first, so, to a first approximation,

$$10{,}010 \approx \frac{2.026 \times 10^{-4}}{D^5} \qquad \text{and} \qquad D = 0.0289 \text{ m}$$

For this first trial,

$$V = \frac{6.366 \times 10^{-3}}{(0.0289)^2} = 7.622 \text{ m/s}$$

$$\text{Re} = \frac{7.622 \times 0.0289}{1.15 \times 10^{-6}} = 1.915 \times 10^5$$

$$\frac{\epsilon}{D} = \frac{0.26 \times 10^{-3}}{0.0289} = 0.0090$$

From the Moody diagram, we obtain f = 0.036. To start a second iteration, let f = 0.036. Therefore,

TABLE 8.5   Viscosity of Water

Approximate values at room temperature:
$\mu = 10^{-3}$ Ns/m^2
$\nu = 10^{-6}$ m^2/s

Temperature (°C)	Dynamic viscosity ($\times 10^{-3}$ Ns/m^2)
0.01	1.755
20	1.002
40	0.651
60	0.462
80	0.350
100	0.278

Source: From Refs. 5 and 18.

$$10,010 = \left[\frac{400(0.036)}{D} + 1 \right] \frac{2.026 \times 10^{-5}}{D^4}$$

or

$$10,010 \approx \frac{0.0002917}{D^5} \quad \text{and} \quad D = 0.0311 \text{ m}$$

For this second trial,

$$V = \frac{6.366 \times 10^{-3}}{(0.0311)^2} = 6.582 \text{ m/s}$$

$$\text{Re} = \frac{6.582 \times 0.0311}{1.15 \times 10^{-6}} = 1.78 \times 10^5$$

$$\frac{\epsilon}{D} = \frac{0.26 \times 10^{-3}}{0.0311} = 0.0084$$

From the Moody diagram, we obtain $f = 0.036$. The agreement with the assumed value is good enough that $D = 0.0311$ m can be taken as the required diameter.

OPEN-CHANNEL FLOW

1. One-Dimensional Approximation.   *Open-channel flow* is the flow of a liquid in a conduit with a free surface. Owing to the presence of the free surface, the pressure can be taken constant along the free surface and essentially equal to atmospheric pressure. Unlike flow in closed ducts, pressure gradient is not a direct factor in open-channel flow, where the balance of forces is confined to gravity and friction. Surface tension is rarely important because open channels are normally quite large and have very large Weber numbers. But the free surface complicates the analysis because its shape is a priori unknown: The depth profile changes with

conditions and must be computed as part of the problem, especially in unsteady problems involving wave motion.

An open channel always has two sides and a bottom, where the flow satisfies the nonslip condition. Therefore, even a straight channel has a three-dimensional velocity distribution. Some measurements of straight-channel velocity contours are shown in Fig. 8.43. The profiles are quite complex, with maximum velocity typically occurring in the midplane about 20 percent below the surface. Little theoretical work has been done on velocity distributions like those in Fig. 8.43. Instead, the engineering approach is to assume one-dimensional flow with an average velocity $V(x)$ at each cross-sectional area $A(x)$, where x is distance along the channel. Since density variations are negligible in low-speed liquid flow, for steady flow the volume flux Q is constant along the channel from continuity:

$$Q = V(x)A(x) = \text{constant} \qquad (8.65)$$

A second relation between velocity and depth is the Bernoulli equation including friction losses. If points 1 (upstream) and 2 (downstream) are on the free surface, $p_1 = p_2 = p_a$, and we have, for steady flow,

$$\frac{V_1^2}{2g} + z_1 = \frac{V_2^2}{2g} + z_2 + h_f \qquad (8.66)$$

where h_f is the friction-head loss. The wall friction in channel flow is quite similar to that in steady duct flow and can be correlated adequately with Eq. (8.62) for turbulent flow with a rough surface where $D = D_h = 4A/P$. Most open-channel analyses use the hydraulic radius R_h instead, which is one-fourth the hydraulic diameter:

Natural irregular channel   Narrow
                                             rectangular
                                             section
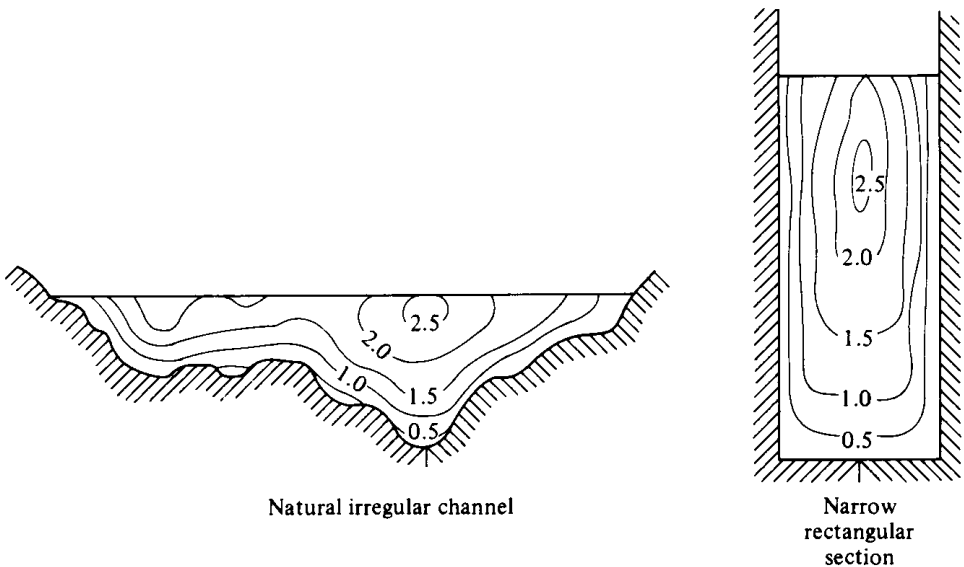
FIGURE 8.43 Measured isovelocity contours in typical straight open-channel flow. (*From White,*[2] *p. 598.*)

$$R_h = \tfrac{1}{4}D_h = \frac{A}{P} \tag{8.67}$$

where P is the wetted perimeter, which includes the sides and bottom of the channel but not the free surface and not the parts of the sides above the water lever.

Open channels are usually large and deep and contain low-viscosity water; hence they are almost always turbulent.

2. Falling Liquid Films.   In the case of liquid film flowing down an inclined or vertical wall (Fig. 8.44a), three different flow regimes have been observed experimentally:

1. At Reynolds numbers $Re_\Gamma = 4\Gamma/\mu$ that do not exceed 20 to 30, there exists the usual viscous flow regime, and the film thickness is constant.
2. At Reynolds numbers $Re_\Gamma > 30$ to 50, a so-called wave regime appears, in which wave motion is superimposed on the forward motion of the film.
3. At $Re_\Gamma \approx 1600$, the laminar regime is replaced by turbulent motion.

In general, waves at the interface create good mixing within the film, especially in the vicinity of the interface. Experimental measurements suggest that the turbulence influences the momentum transfer relatively little as compared to the wave effect.

Note that Γ is the mass flow rate of the film per unit width. For example, in the case of flow on the outside surface of a vertical tube with outer diameter d, Γ is equal to the mass flow rate divided by πd.

By equating the gravity forces with friction forces and satisfying the condition of no slip at the wall, the film thickness of a smooth laminar falling film is given as

$$\delta^+ = 0.866 Re_\Gamma^{0.5} \tag{8.68}$$
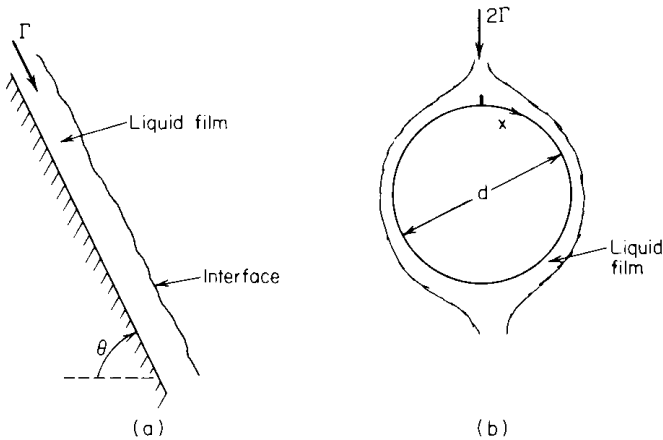
FIGURE 8.44   Falling film on (a) an inclined plate and (b) a horizontal tube. (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 12-74.*)

where

$$\delta^+ = \frac{\delta(g\delta)^{0.5}}{\nu} \tag{8.69}$$

If the film surface is inclined at $\theta < 90°$ to the horizontal, then g in Eq. (8.69) should be replaced with $g \sin \theta$. For calculation of the mean film thicknesses of turbulent falling films, several empirical relations have been proposed in the form

$$\delta^+ = C\mathrm{Re}_\Gamma^n \tag{8.70}$$

Recent comparisons[5] indicated that the specifications of $C = 0.095$, $n = 0.8$ are the best. The thickness of the liquid film on the horizontal tube (Fig. 8.44b) varies around the tube periphery because the gravity force component in the flow direction varies; in this case, g in Eq. (8.69) should be replaced with $g \sin (2x/d)$. Note that Γ for the ease of a horizontal tube is the flow rate per axial unit length over one side of the horizontal tube (i.e., Γ is equal to the mass flow rate falling on a single tube divided by $2L$, where L is the tube length).

TWO-PHASE FLOW

1. General Relations. In general, when two phases flow in a pipe (Figs. 8.45 and 8.46), they do not move at the same velocity. The continuity equation for each phase fixes a relation between the phase velocities V_l and V_g, the void fraction α, and the flowing quality x. This relation is as follows. First, the definitions of the phase velocities:

$$V_l = \frac{Q_l}{A(1 - \alpha)} \tag{8.71}$$

$$V_g = \frac{Q_g}{A(\alpha)} \tag{8.72}$$
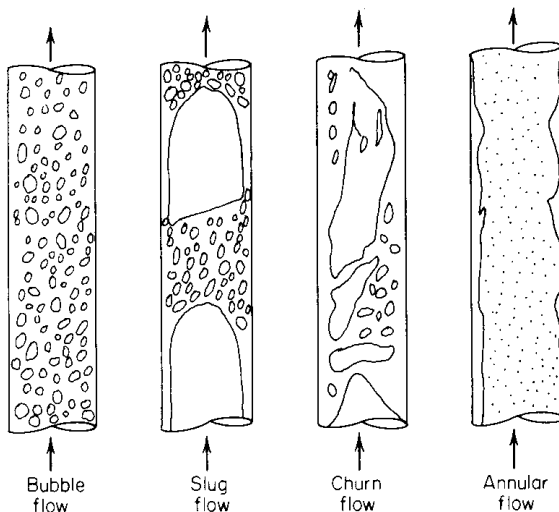
FIGURE 8.45 Flow patterns in vertical flow. (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 13-13.*)
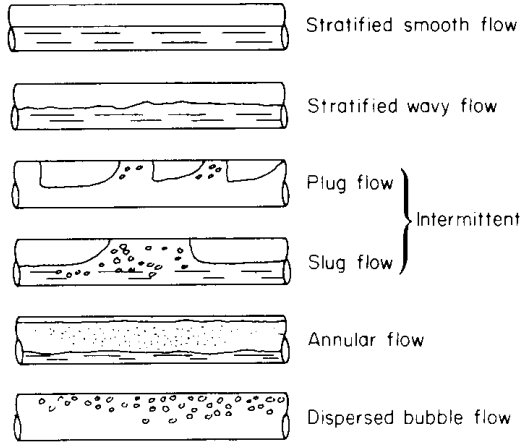
FIGURE 8.46 Flow patterns in horizontal flow.
(*From Rohsenow, Hartnett, and Ganić,*[5] *p. 13-16.*)

Using these definitions, it is possible to develop the following relationship between the velocity ratio, the void, and the quality:

$$\frac{1 - \alpha}{\alpha} = \left(\frac{V_g}{V_l}\right) \left(\frac{\rho_g}{\rho_l}\right) \left(\frac{1 - x}{x}\right) \tag{8.73}$$

2. Pressure-Drop Calculation by Homogeneous Method. The pressure drop in a straight pipe is the result of the action of three factors: the wall friction force, the gravity force, and momentum changes. The general equation for pressure drop in a straight pipe is

$$\Delta p = \Delta p_f + \Delta p_G + \Delta p_m \tag{8.74}$$

The density of a homogeneous two-phase mixture is calculated assuming that the velocity of each phase is the same. That is,

$$V = \frac{1}{\rho} = V_l + xV_{lg} \tag{8.75}$$

The friction pressure drop is determined from

$$\Delta p_f = f\left(\frac{L}{D}\right)\frac{G^2 V}{2} \tag{8.76}$$

The specific volume is calculated from Eq. (8.75). The mass velocity is calculated from the total flow rate and pipe cross-sectional area. Thus

$$G = \frac{\dot{m}}{A} \tag{8.77}$$

The friction factor is determined from the Moody diagram assuming that pure liquid is flowing in the pipe at the mixture mass velocity. The friction factor so

determined will generally be low at qualities below 70 percent and will be too high at qualities above 70 percent.

The gravity pressure drop is calculated from

$$\Delta p_G = \rho g L \sin \theta \tag{8.78}$$

The density ρ is evaluated from Eq. (8.75). The angle θ is measured from the horizontal.

The momentum pressure drop is a result of either a pressure drop or heat transfer giving rise to a change in the volume flow rate. It is

$$\Delta p_m = G^2(V_2 - V_1) \tag{8.79}$$

The first term in the parentheses is the specific volume at the discharge, and the second term is the specific volume at the entrance. Both are evaluated using Eq. (8.75).

Note. Use this method where simplicity and an analytical expression are needed. Precision is low for low velocities or low pressure where the gravity contribution to the pressure drop is large. Precision improves for high quality and high pressure. The homogeneous method is used as the basis of the calculation of pressure drop for fittings.[5]

ACKNOWLEDGMENTS

The numerical examples in the text are based on material from Refs. 1, 2, 4, and 6 to 10.

NOMENCLATURE

Symbol = definition, SI units (U.S. Customary units)

A = area, m^2 (ft^2)

c = velocity of sound, m/s (ft/s)

C_D = drag coefficient

C_f = skin-friction coefficient

C_L = lift coefficient

D = inside tube diameter, m (ft); outside diameter (in "Flow over Immersed Bodies" section), m (ft)

D_h = hydraulic diameter, m (ft)

d = tube outer diameter, m (ft)

F, **F** = force (magnitude and vector), N (lb$_f$)

F_D = drag force, N (lb$_f$)

F_L = lift force, N (lb$_f$)

f = friction factor

G = flow rate per unit area $(= \dot{m}/A)$, kg/(s · m^2) [lb/(s · ft^2)]

g, $\mathbf{g}$ = acceleration of gravity (magnitude and vector) m/s² (ft/s²)

h = vertical distance, m (ft)

h_f = friction head, m · N/N (ft · lb$_f$/lb$_f$)

K = loss coefficient

L = length, m (ft)

L_e = entrance length, m (ft)

$\dot{m}$ = mass flow rate, kg/s (lb/s)

p = pressure, N/m², Pa (lb$_f$/ft²)

P = power, W (Btu/h)

Q = volume flow rate, m³/s (ft³/s)

Δp = pressure drop, N/m² (lb$_f$/ft²)

r = radius, m (ft)

Re = Reynolds number

Re$_\Gamma$ = film Reynolds number (= $4\Gamma/\mu$)

R = gas constant, $J/(\text{kg} \cdot \text{K})$ [Btu/(lb · °R)]

s = distance, m (ft)

T = temperature, K (°R)

T = torque (in ''Moment-of-Momentum Equation'' section), N · m (lb$_f$ · ft)

t = time, s (s)

u = velocity in the axial direction (x direction), m/s (ft/s)

U = free-stream velocity (in ''Boundary-Layer Flows'' section), m/s (ft/s)

V = volume (in ''Fluid Statics'' section), m³ (ft³)

V, $\mathbf{V}$ = velocity (magnitude and vector), m/s (ft/s)

W = weight, N (lb$_f$)

w_{1-2} = work per unit mass, N · m/kg (lb$_f$ · ft/lb)

ν = specific volume, m³/kg (ft³/lb)

x = rectangular coordinate, m (ft)

y = rectangular coordinate, m (ft)

z = rectangular coordinate, m (ft)

Greek

α = void fraction (in ''Two-Phase Flow'' section)

δ = liquid film thickness (in ''Falling Liquid Films'' section)

δ^+ = dimensionless film thickness (in ''Falling Liquid Films'' section)

ϵ = surface roughness, m (ft)

μ = viscosity (dynamic viscosity), Pa · s [lb/(h · ft)]

τ = shear stress, N/m² (lb$_f$/ft²)

ν = kinematic viscosity, m²/s (ft²/s)

ρ = density, kg/m³ (lb/ft³)

σ = surface tension, N/m (lb$_f$/ft)

γ = specific weight (= ρg), N/m^3 (lb$_f$/ft^3)

θ = angle, rad (degree)

η = eddy viscosity, Pa · s [lb/(h · ft)]

Γ = liquid film flow rate per perimeter, kg/(s · m) [lb/(s · ft)]

ω = angular velocity, 1/s

Subscripts

1 = cross section 1

2 = cross section 2

f = friction

l = liquid

g = gas

lg = g-l

m = momentum

G = gravity

x = x direction; function of x

w = wall

∞ = free-stream condition

Mathematical Operation Symbols

∇ = del operator (∇ = **i** $\partial/\partial x$ + **j** $\partial/\partial y$ + **k** $\partial/\partial z$), 1/m (1/ft)

i, j, k = unit vectors

d/dx = derivative with respect to x, 1/m (1/ft)

$\partial/\partial x$ = partial derivative with respect to x, 1/m (1/ft)

*REFERENCES**

1. V. L. Streeter and E. B. Wylie, *Fluid Mechanics,* 7th ed., McGraw-Hill, New York, 1979.

2. F. M. White, *Fluid Mechanics,* McGraw-Hill, New York, 1979.

3. E. A. Avallone and T. Baumeister, III (eds.), *Marks' Standard Handbook for Mechanical Engineering,* 9th ed., McGraw-Hill, New York, 1987.

4. R. F. Probstein, *Advanced Fluid Dynamics* 2.25, MIT class notes, 1974–1975.

5. W. M. Rohsenow, J. P. Hartnett, and E. N. Ganić (eds.), *Handbook of Heat Transfer Fundamentals,* 2d ed., McGraw-Hill, New York, 1985.

6. A. A. Sonin, *Fluid Mechanics* 2.20, MIT class notes, 1974–1975.

7. I. H. Shames, *Mechanics of Fluids,* McGraw-Hill, New York, 1962.

*Those references listed here but not cited in the text were used for comparison of different data sources, clarification, clarity of presentation, and, most important, reader's convenience when further interest in the subject exists.

8. W.-H. Li and S.-H. Lam, *Principles of Fluid Mechanics,* Addison-Wesley, Reading, Mass., 1976.

9. R. H. F. Pao, *Fluid Dynamics,* Charles E. Merrill Books, Inc., Columbus, Ohio, 1967.

10. J. E. A. John and W. L. Haberman, *Introduction to Fluid Mechanics,* 2d ed., Prentice-Hall, Englewood Cliffs, N.J., 1980.

11. H. Schlichting, *Boundary Layer Theory,* 7th ed., McGraw-Hill, New York, 1979.

12. L. E. Moody, ''Friction Factors for Pipe Flow,'' *Trans. ASME,* 68: 672, 1944.

13. R. J. Goldstein, *Fluid Mechanics Measurements,* Hemisphere, Washington, 1983.

14. R. H. Perry, D. W. Green, and J. O. Maloney, *Perry's Chemical Engineers' Handbook,* 6th ed., McGraw-Hill, New York, 1984.

15. L. Prandtl and O. G. Tietjens, *Fundamentals of Hydro- and Aeromechanics,* Dover Books, New York, 1934.

16. C. E. O'Rourke, *General Engineering Handbook,* 2d ed., McGraw-Hill, New York, 1940.

17. A. H. Shapiro, *Illustrated Experiments in Fluid Mechanics* (The NCFMF Book of Film Notes), MIT Press, Cambridge, Mass., 1972.

18. J. Carvill, *Mechanical Engineer's Data Handbook,* Butterworth-Heinemann, 2001.

19. F. S. Merritt, M. Kent Loftin, and J. T. Ricketts, *Standard Handbook for Civil Engineers,* 4th ed., McGraw-Hill, 1996.

20. G. P. Celata, P. Di Marco, A. Goulas, and A. Mariani, Eds., *Experimental Heat Transfer, Fluid Mechanics and Thermodynamics-2001,* Edizioni ETS, Pisa, Italy, 2001.

CHAPTER 9
HEAT AND MASS TRANSFER

Heat is defined as energy transferred by virtue of a temperature difference. It flows from regions of higher temperature to regions of lower temperature. It is customary to refer to different types of heat-transfer mechanisms as *modes.* The basic modes of heat transfer are *conduction* and *radiation.* In some textbooks on heat transfer, *convection* is also considered as a separate mechanism.

CONDUCTION

Conduction is the transfer of heat from one part of a body at a higher temperature to another part of the same body at a lower temperature or from one body at a higher temperature to another body at a lower temperature in physical contact with it. The conduction process takes place at the molecular level and involves the transfer of energy from the more energetic molecules to those with a lower energy level. This can be easily visualized within gases, where we note that the average kinetic energy of molecules in the higher-temperature regions is greater than those in the lower-temperature regions. The more energetic molecules, being in constant and random motion, periodically collide with molecules of a lower energy level and exchange energy and momentum. In this manner there is a continuous transport of energy from the high-temperature regions to those of lower temperature. In liquids, the molecules are more closely spaced than in gases, but the molecular energy exchange process is qualitatively similar to that in gases. In solids that are conductors of electricity (dielectrics), heat is conducted by lattice waves caused by atomic motion. In solids that are good conductors of electricity, this lattice vibration mechanism is only a small contribution to the energy-transfer process, the principal contribution being that due to the motion of free electrons, which move in a similar way to molecules in a gas.

At the macroscopic level the heat flux (i.e., the heat transfer rate per unit area normal to the direction of heat flow) q'' is proportional to the temperature gradient:

$$q'' = -k \frac{dT}{dx} \tag{9.1}$$

where the proportionality constant k is a *transport* property known as the *thermal*

conductivity and is a characteristic of the material. The minus sign is a consequence of the fact that heat is transferred in the direction of decreasing temperature. Equation (9.1) is the one-dimensional form of *Fourier's law* of heat conduction. Recognizing that the heat flux is a vector quantity, we can write a more general statement of Fourier's law (i.e., the conduction rate equation) as

$$\mathbf{q}'' = -k\, \nabla T \qquad (9.2)$$

where ∇ is the three-dimensional del operator, and T is the scalar temperature field. From Eq. (9.2) it is seen that the heat flux vector $\mathbf{q}''$ represents actually a current of heat (thermal energy) which flows in the direction of the steepest temperature gradient. Scalar components of the heat flux vector $\mathbf{q}''$ are given in Table 10.8.

1. Composite Walls

Flat Solid Wall. If we consider a one-dimensional heat flow along the x direction in the plane wall shown in Fig. 9.1a, direct application of Eq. (9.1) can be made, and then integration yields

$$q = \frac{kA}{\Delta x} (T_2 - T_1) \qquad (9.3)$$

where the thermal conductivity is considered constant, Δx is the wall thickness, and T_1 and T_2 are the wall-face temperatures. Note that $q/A = q''$, where q is the heat transfer rate through an area A. Equation (9.3) can be written in the form

$$q = \frac{T_2 - T_1}{\Delta x/kA} = \frac{T_2 - T_1}{R_{th}} = \frac{\text{thermal potential difference}}{\text{thermal resistance}} \qquad (9.4)$$

where $\Delta x/kA$ assumes the role of a *thermal resistance* R_{th}. The relation of Eq. (9.4)
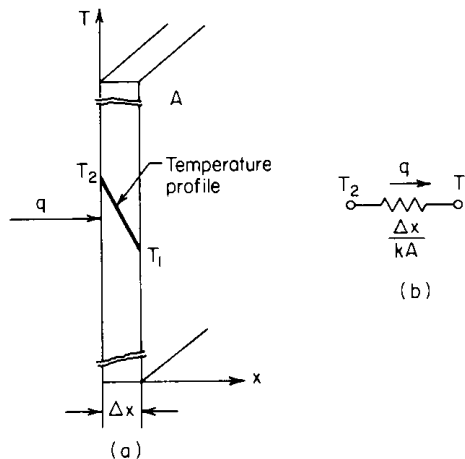
FIGURE 9.1 One-dimensional heat conduction through a plane wall (*a*) and electrical analogue (*b*). (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 1-3.*)

is quite like Ohm's law in electric circuit theory. The equivalent electric circuit for this case is shown in Fig. 9.1*b*.

The electrical analogy may be used to solve more complex problems involving both series and parallel resistances. Typical problems and their analogous electric circuits are given in many heat transfer textbooks.[1–4]

In treating conduction problems it is often convenient to introduce another property which is related to the thermal conductivity, namely, the thermal diffusivity α,

$$\alpha = \frac{k}{\rho c} \tag{9.5}$$

where ρ is the density and c is the specific heat.

Cylindrical Solid Wall. The rate of heat transfer through a cylindrical solid wall of length L is calculated from the equation

$$q = 2\pi L k \frac{T_1 - T_2}{\ln (r_2/r_1)} \tag{9.6}$$

where r_2 and r_1 are outside and inside radii, and ln is logarithm with base e. Similarly for the spherical solid wall

$$q = \frac{4\pi k r_1 r_2 (T_1 - T_2)}{r_2 - r_1} \tag{9.7}$$

An important case of heat transfer is that from a hot fluid on one side of a solid wall to a cooler fluid on the other side. The wall may be a cylindrical or a flat wall of a single material or composite of different materials. The rate of heat transfer is calculated from

$$q = UA(T_h - T_c) \tag{9.8}$$

where T_h and T_c are the temperatures of the hot and cold fluids, respectively.

Flat Composite Wall. Equation (9.8) is used to calculate the rate of heat transfer from a hot fluid successively through the hot-side film, layers of solid material of the wall, and cold-side film to the cold fluid (Fig. 9.2). For this case, U of Eq. (9.8) is
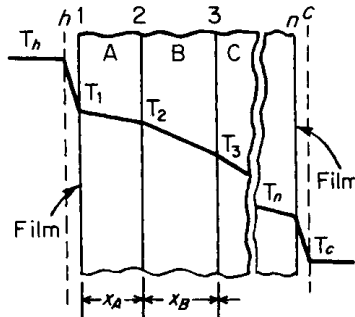
FIGURE 9.2 Composite wall.

$$U = \frac{1}{(1/h_h) + (x_A/k_A) + (x_B/k_B) + \cdots + (1/h_c)} \tag{9.9}$$

where h_h and h_c are the hot- and cold-side film coefficients of heat transfer (see the section "Convection" for a definition of h).

Cylindrical Composite Wall. Equation (9.8) is also used to calculate the rate of heat transfer from a hot to a cold fluid through a composite cylindrical pipe wall (Fig. 9.3). For this case, the product of UA of Eq. (9.8) is

$$UA = \frac{1}{\dfrac{1}{2r_1 L h_h} + \dfrac{\ln (r_2/r_1)}{2L k_A} + \dfrac{\ln (r_3/r_2)}{2L k_B} + \cdots + \dfrac{1}{2r_n L h_C}} \tag{9.10}$$

Spherical Composite Wall

$$UA = \frac{1}{\dfrac{1}{4r_1^2 h_h} + \dfrac{r_2 - r_1}{4k r_1 r_2} + \cdots + \dfrac{1}{4r_n^2 h_c}} \tag{9.11}$$

General Conduction Equation. The differential equation for temperature distributions in solids is given by Eq. (10.23).

2. Mass Transfer by Diffusion. As mentioned above, heat transfer will occur whenever there exists a temperature difference in a medium. Similarly, whenever there exists a difference in the concentration or density of some chemical species in a mixture, mass transfer must occur. Hence, just as a temperature gradient constitutes the driving potential for heat transfer, the existence of a concentration gradient for some species in a mixture provides the driving potential for transport of that species. Therefore, the term *mass transfer* describes the relative motion of species in a mixture due to the presence of concentration gradients.

Since the same physical mechanism is associated with heat transfer by conduction (i.e., heat diffusion) and mass transfer by diffusion, the corresponding rate
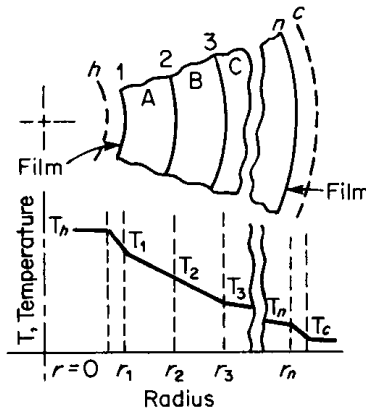
FIGURE 9.3 Composite cylinder or sphere.

equations are of the same form. The rate equation for mass diffusion is known as *Fick's law,* and for a transfer of species 1 in a binary mixture, it may be expressed as

$$j_1 = -\mathbf{D}\,\frac{dC_1}{dx} \tag{9.12}$$

where C_1 is a mass concentration of species 1 in units of mass per unit volume. This expression is analogous to Fourier's law, Eq. (9.1). Moreover, just as Fourier's law serves to define one important transport property, the thermal conductivity, Fick's law defines a second important transport property, namely the *binary diffusion coefficient* or mass diffusivity **D.** The quantity j_1 [mass/(time $\times$ surface area)] is defined as the mass flux of species 1, i.e., the amount of species 1 that is transferred per unit time and per unit area perpendicular to the direction of transfer. In vector form, Fick's law is given as

$$j_1 = -\mathbf{D}\,\nabla C_1 \tag{9.13}$$

In general, the diffusion coefficient **D** for gases at low pressure is almost composition-independent; it increases with temperature and varies inversely with pressure. Liquid and solid diffusion coefficients are markedly concentration-dependent and generally increase with temperature. Differential equations for concentration distribution in solids and fluid medium are given in the section "The Conservation Equation for Species" in Chap. 10.

RADIATION

Radiation, or more correctly *thermal radiation,* is electromagnetic radiation emitted by a body by virtue of its temperature and at the expense of its internal energy. Thus thermal radiation is of the same nature as visible light, x-rays, and radio waves, the difference between them being in their wavelengths and the source of generation. The eye is sensitive to electromagnetic radiation in the region from 0.39 to 0.78 μm; this is identified as the visible region of the spectrum. Radio and hertzian waves have a wavelength of 1×10^3 to 2×10^{10} μm, and x-rays have wavelengths of 1×10^{-5} to 2×10^{-2} μm, while the bulk of thermal radiation occurs in rays from approximately 0.1 to 100 μm. All heated solids and liquids as well as some gases emit thermal radiation. The transfer of energy by conduction requires the presence of a material medium, while radiation does not. In fact, radiation transfer occurs most efficiently in a vacuum. On the macroscopic level, the calculation of thermal radiation is based on the *Stefan-Boltzmann law,* which relates the energy flux emitted by an ideal radiator (or *blackbody*) to the fourth power of the absolute temperature:

$$e_b = \sigma T^4 \tag{9.14}$$

Here σ is the Stefan-Boltzmann constant with a value of 5.669×10^{-8} W/(m^2 · K^4), or 1.714×10^{-9} Btu/(h · ft^2 · °R^4). Engineering surfaces in general do not perform as ideal radiators, and for real surfaces, the preceding law is modified to read

$$e = \epsilon \sigma T^4 \tag{9.15}$$

The term ϵ is called the *emissivity* of the surface and has a value between 0 and 1. When two blackbodies exchange heat by radiation, the net heat exchange is then proportional to the difference in T^4. If the first body "sees" only body 2, then the net heat exchange from body 1 to body 2 is given by

$$q = \sigma A_1(T_1^4 - T_2^4) \tag{9.16}$$

When, because of the geometrical arrangement, only a fraction of the energy leaving body 1 is intercepted by body 2,

$$q = \sigma A_1 F_{1-2}(T_1^4 - T_2^4) \tag{9.17}$$

where F_{1-2} (usually called a *shape factor* or a *view factor*) is the fraction of energy leaving body 1 that is intercepted by body 2.

Since the shape factor is a purely geometrical quantity, consider the coordinate system of Fig. 9.4. The two infinitesimal surfaces dA_1 and dA_2 are separated by a distance r, and the line of connection forms the polar angles β_1 and β_2 respectively, with surface normals n_1 and n_2. For dA_1 and dA_2 the definition of a shape factor gives

$$F_{1-2} = \frac{1}{A_1} \int_{A_1} \int_{A_2} \frac{\cos \beta_1 \cos \beta_2}{\pi r^2} dA_1 \, dA_2 \tag{9.18}$$

and
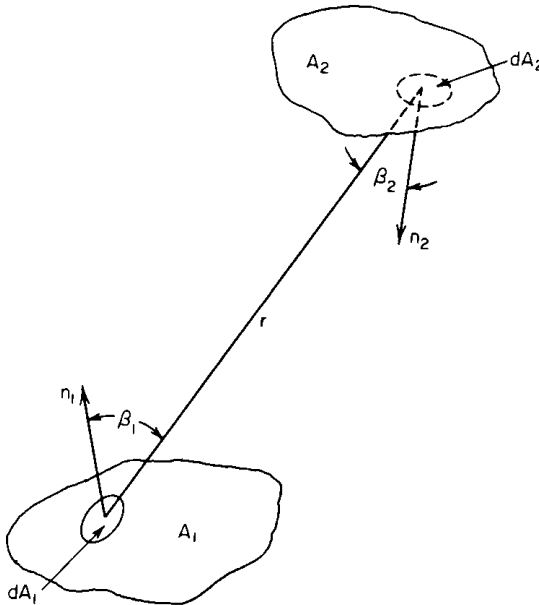
FIGURE 9.4 Coordinate system for shape factors. (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 14-38.*)

$$A_1 F_{1-2} = A_2 F_{2-1} \tag{9.19}$$

The preceding reciprocity relation for the shape factors of two finite surfaces is particularly useful in many engineering calculations. Values of shape factor for most common engineering applications (Fig. 9.5) are given in Table 9.1.

If the bodies are not black, then the view factor F_{1-2} must be replaced by a new factor, $\mathfrak{F}_{1-2}$, which depends on the emissivity ϵ of the surfaces involved (see Table
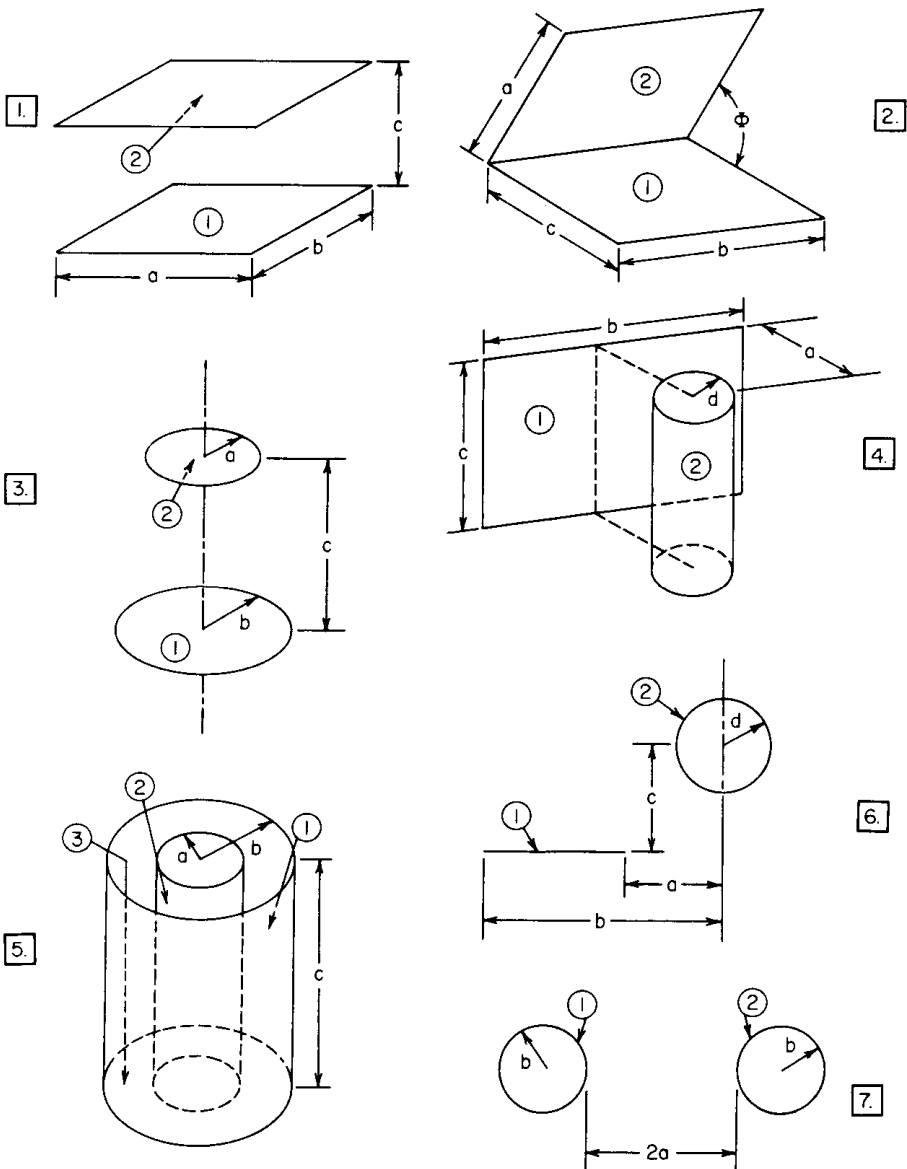
FIGURE 9.5 Schematic representations of configurations 1 through 7 (Table 9.1). (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 14-39.*)

TABLE 9.1   Shape Factor Relations (Fig. 9.5)

Configuration 1: $X = a/c$, $Y = b/c$

$$F_{12}\left(\frac{\pi XY}{2}\right) = \ln\left[\frac{(1 + X^2)(1 + Y^2)}{1 + X^2 + Y^2}\right]^{1/2}$$

$$+ Y\sqrt{1 + X^2}\ \tan^{-1}\frac{Y}{\sqrt{1 + X^2}}$$

$$+ X\sqrt{1 + Y^2}\ \tan^{-1}\frac{X}{\sqrt{1 + Y^2}}$$

$$- Y\tan^{-1} Y - X\tan^{-1} X$$

Configuration 2: $X = a/b$, $Y = c/b$, $Z = X^2 + Y^2 + Y^2 - 2XY\cos\Phi$

$$F_{12}(\pi Y) = -\frac{\sin 2\Phi}{4}\left[XY\sin\Phi + \left(\frac{\pi}{2} - \Phi\right)(X^2 + Y^2)\right.$$

$$+ Y^2\tan^{-1}\frac{X - Y\cos\Phi}{Y\sin\Phi}$$

$$\left. + X^2\tan^{-1}\frac{Y - X\cos\Phi}{X\sin\Phi}\right]$$

$$+ \frac{\sin^2\Phi}{4}\left[\left(\frac{2}{\sin^2\Phi} - 1\right)\ln\frac{(1 + X^2)(1 + Y^2)}{1 + Z}\right.$$

$$\left. + Y^2\ln\frac{Y^2(1 + Z)}{(1 + Y^2)Z} + X^2\ln\frac{X^2(1 + X^2)^{\cos 2\Phi}}{Z(1 + Z)^{\cos 2\Phi}}\right]$$

$$+ Y\tan^{-1}\frac{1}{Y} + X\tan^{-1}\frac{1}{X} - \sqrt{Z}\tan^{-1}\frac{1}{\sqrt{Z}}$$

$$+ \frac{\sin\Phi\sin 2\Phi}{2} X\sqrt{1 + X^2\sin^2\Phi}$$

$$\times\left(\tan^{-1}\frac{X\cos\Phi}{\sqrt{1 + X^2\sin^2\Phi}}\right.$$

$$\left. + \tan^{-1}\frac{Y - X\cos\Phi}{\sqrt{1 + X^2\sin^2\Phi}}\right)$$

$$+ \cos\Phi\int_0^Y \sqrt{1 + \xi^2\sin^2\Phi}\left(\tan^{-1}\frac{X - \xi\cos\Phi}{\sqrt{1 + \xi^2\sin^2\Phi}}\right.$$

$$\left. + \tan^{-1}\frac{\xi\cos\Phi}{\sqrt{1 + \xi^2\sin^2\Phi}}\right)d\xi$$

Configuration 3: $X = a/c$, $Y = c/b$, $Z = 1 + (1 + X^2)Y^2$

$$F_{12} = \tfrac{1}{2}(Z - \sqrt{Z^2 - 4X^2Y^2})$$

TABLE 9.1   Shape Factor Relations (Fig. 9.5) (*Continued*)

Configuration 4: $X = a/d$, $Y = b/d$, $Z = c/d$

$$A = Z^2 + X^2 + \xi^2 - 1, \quad B = Z^2 - X^2 - \xi^2 + 1$$

$$F_{12} = \frac{2}{Y} \int_0^{Y/2} f(\xi)\, d\xi$$

$$f(\xi) = \frac{X}{X^2 + \xi^2} - \frac{X}{\pi(X^2 + \xi^2)} \times \left[\cos^{-1} \frac{B}{A} - \frac{1}{2Z} \left(\sqrt{A^2 + 4Z^2}\, \cos^{-1} \frac{B}{A\sqrt{X^2 + \xi^2}} \right. \right.$$

$$\left. \left. + B \sin^{-1} \frac{1}{\sqrt{X^2 + \xi^2}} - \frac{\pi A}{2} \right) \right]$$

Configuration 5: $X = b/a$, $Y = c/a$

$$A = Y^2 + X^2 - 1, \quad B = Y^2 - X^2 + 1$$

$$F_{12} = \frac{1}{X} - \frac{1}{\pi X} \left\{ \cos^{-1} \frac{B}{A} - \frac{1}{2Y} \left[\sqrt{(A+2)^2 - (2X)^2}\, \cos^{-1} \frac{B}{XA} + B \sin^{-1} \frac{1}{X} - \frac{\pi A}{2} \right] \right\}$$

$$F_{11} = 1 - \frac{1}{X} + \frac{2}{\pi X} \tan^{-1} \frac{2\sqrt{X^2 - 1}}{Y}$$

$$- \frac{Y}{2\pi X} \left[\frac{\sqrt{4X^2 + Y^2}}{Y} \sin^{-1} \frac{4(X^2 - 1) + \frac{Y^2}{X^2}(X^2 - 2)}{Y^2 + 4(X^2 - 1)} \right.$$

$$\left. - \sin^{-1} \frac{X^2 - 2}{X^2} + \frac{\pi}{2} \left(\frac{\sqrt{4X^2 + Y^2}}{Y} - 1 \right) \right]$$

$$F_{13} = \tfrac{1}{2}(1 - F_{12} - F_{11})$$

Configuration 6: $X = c/d$, $Y = a/d$, $Z = b/d$

$$F_{12} = \frac{1}{Z - Y} \left(\tan^{-1} \frac{Z}{X} - \tan^{-1} \frac{Y}{X} \right)$$

Configuration 7: $X = 1 + (a/b)$

$$F_{12} = \frac{2}{\pi} \left(\sqrt{X^2 - 1} - X + \frac{\pi}{2} - \cos^{-1} \frac{1}{X} \right)$$

Source: From Rohsenow, Hartnett, and Ganić,[5] pp. 14-40 to 14-41.

9.2) as well as the geometrical view. Finally, if the bodies are separated by gases or liquids that impede the radiation of heat through them, a formulation of the heat-exchange process becomes more involved (see Chap. 14 of Ref. 5).

CONVECTION

Convection, sometimes identified as a separate mode of heat transfer, relates to the transfer of heat from a bounding surface to a fluid in motion or to the heat transfer

TABLE 9.2 Emissivity of Various Surfaces

Surface	Emissivity, ϵ	T, °C
A. Metals and their oxides		
Aluminum		
Highly polished plate, 98.3% pure	0.039–0.057	226–576
Commercial sheet	0.09	100
Heavily oxidized	0.20–0.31	148–504
Al-surfaced roofing	0.216	38
Chromium (see nickel alloys for Ni–Cr steels), polished	0.08–0.36	38–1093
Copper		
Polished	0.023	117
Polished	0.052	100
Plate, heated long time, covered with thick oxide layer	0.78	25
Gold, pure, highly polished	0.018–0.035	226–627
Iron and steel (not including stainless)		
Steel, polished	0.066	100
Iron, polished	0.14–0.38	427–1027
Cast iron, newly turned	0.44	22
Cast iron, turned and heated	0.60–0.70	882–988
Mild steel	0.20–0.32	232–1065
Oxidized surfaces		
Iron plate, pickled, then rusted red	0.61	20
Iron, dark-gray surface	0.31	100
Rough ingot iron	0.87–0.95	927–1115
Sheet steel with strong, rough oxide layer	0.80	24
Nickel		
Polished	0.072	100
Nickel oxide	0.59–0.86	649–1254
Nickel alloys		
Copper nickel, polished	0.059	100
Nichrome wire, bright	0.65–0.79	49–1000
Nichrome wire, oxidized	0.95–0.98	49–500
Platinum, polished plate, pure	0.054–0.104	226–627
Silver		
Polished, pure	0.020–0.032	226–627
Polished	0.022–0.031	38–371
Stainless steels		
Polished	0.074	100
Type 301	0.54–0.63	232–940
Tin, bright tinned iron	0.043 and 0.064	24
Tungsten, filament	0.39	3315
B. Refractories, building materials, paints, and miscellaneous		
Alumina (85–99.5% Al_2O_3, 0–12% SiO_2, 0–1% Fe_2O_3);		
effect of mean grain size (μm)		
10 μm	0.30–0.18	
50 μm	0.39–0.28	
100 μm	0.50–0.40	
Asbestos, board	0.96	23
Brick		
Red, rough, but no gross irregularities	0.93	21
Fireclay	0.75	1000

TABLE 9.2 Emissivity of Various Surfaces (*Continued*)

Surface	Emissivity, ϵ	T, °C
B. Refractories, building materials, paints, and miscellaneous (*Continued*)		
Carbon		
T carbon (Gebruder Siemens) 0.9% ash, started with emissivity at 260°F of 0.72, but on heating changed to values given	0.81–0.79	127–627
Filament	0.526	1038–1404
Rough plate	0.77	100–320
Lampblack, rough deposit	0.84–0.78	100–500
Concrete tiles	0.63	1000
Enamel, white fused, on iron	0.90	19
Glass		
Smooth	0.94	22
Pyrex, lead, and soda	0.95–0.85	260–538
Paints, lacquers, varnishes		
Snow-white enamel varnish on rough iron plate	0.906	23
Black shiny lacquer, sprayed on iron	0.875	24
Black shiny shellac on tinned iron sheet	0.821	21
Black matte shellac	0.91	77–146
Black or white lacquer	0.80–0.95	38–93
Flat black lacquer	0.96–0.98	38–93
Aluminum paints and lacquers		
10% Al, 22% lacquer body, on rough or smooth surface	0.52	100
Other Al paints, varying age and Al content	0.27–0.67	100
Porcelain, glazed	0.92	22
Quartz, rough, fused	0.93	21
Roofing paper	0.91	21
Rubber, hard, glossy plate	0.94	23
Water	0.95–0.963	0–100

Source: Ref. 11.

across a flow plane within the interior of the flowing fluid. If the fluid motion is induced by a pump, a blower, a fan, or some similar device, the process is called *forced convection.* If the fluid motion occurs as a result of the density difference produced by the heat transfer itself, the process is called *free* or *natural convection.*

Detailed inspection of the heat-transfer process in these cases reveals that although the bulk motion of the fluid gives rise to heat transfer, the basic heat-transfer mechanism is *conduction;* i.e., the energy transfer is in the form of heat transfer by conduction within the moving fluid. More specifically, it is not *heat* that is being convected but *internal energy.*

However, there are convection processes for which there is, in addition, *latent heat exchange.* This latent heat exchange is generally associated with a phase change between the liquid and vapor states of the fluid. Two special cases are *boiling* and *condensation.*

1. Heat-Transfer Coefficient. In convective processes involving heat transfer from a boundary surface exposed to a relatively low velocity fluid stream, it is convenient to introduce a heat-transfer coefficient *h,* defined by Eq. (9.20), which is known as *Newton's law of cooling*:

$$q'' = h(T_w - T_f) \tag{9.20}$$

Here T_w is the surface temperature and T_f is a characteristic fluid temperature.

For surfaces in unbounded convection, such as plates, tubes, bodies of revolution, etc., immersed in a large body of fluid, it is customary to define h in Eq. (9.20) with T_f as the temperature of the fluid far away from the surface, often identified as T_∞ (Fig. 9.6). For bounded convection, including such cases as fluids flowing in tubes or channels, across tubes in bundles, etc., T_f is usually taken as the enthalpy-mixed-mean temperature, customarily identified as T_m.

The heat-transfer coefficient defined by Eq. (9.20) is sensitive to the geometry, to the physical properties of the fluid, and to the fluid velocity. However, there are some special situations when h can depend on the temperature difference $T_w - T_f = \Delta T$. For example, if the surface is hot enough to boil a liquid surrounding it, h will typically vary as ΔT^2, or in the case of natural convection, h varies as some weak power of ΔT—typically as $\Delta T^{1/4}$ or $\Delta T^{1/3}$. It is important to note that Eq. (9.20) as a definition of h is valid in these cases too, although its usefulness may well be reduced.

Since $q'' = q/A$, from Eq. (9.20) the thermal resistance in convection heat transfer is given by

$$R_{th} = \frac{1}{hA}$$

which is actually the resistance at a surface-to-fluid interface.

At the wall, the fluid velocity is zero, and the heat transfer takes place by conduction. Therefore, we may apply Fourier's law to the fluid at $y = 0$ (where y is the axis normal to the flow direction; Fig. 9.6):

$$q'' = -k \left. \frac{\partial T}{\partial y} \right|_{y=0} \tag{9.21}$$

where k is the thermal conductivity of fluid. By combining this equation with Newton's law of cooling [Eq. (9.20)], we then obtain

$$h = \frac{q''}{T_w - T_f} = -\frac{k(\partial T/\partial y)|_{y=0}}{T_w - T_f} \tag{9.22}$$

so that we need to find the temperature gradient at the wall in order to evaluate the heat-transfer coefficient.
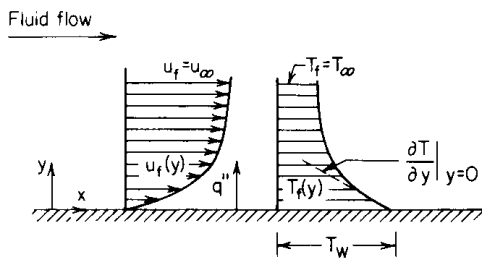
FIGURE 9.6 Velocity and temperature distributions in flow over a flat plate. (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 1-5.*)

The temperature gradient at the wall can be obtained by solving Eq. (10.22). The heat-transfer coefficient h can be expressed conveniently in terms of a nondimensional Nusselt number

$$\text{Nu} = \frac{hL}{k} \tag{9.23}$$

where k is the fluid thermal conductivity and L is the characteristic length (i.e., for the case of flow in a circular tube, $L = D,$ where D is the tube diameter).

2. Convective Mass Transfer. Similar results may be obtained for *convective mass transfer.* If a fluid of species concentration $C_{1,\infty}$ flows over a surface at which the species concentration is maintained at some value $C_{1,w} \neq C_{1,\infty}$, transfer of the species by convection will occur. Species 1 is typically a vapor that is transferred into a gas stream by evaporation or sublimation at a liquid or solid surface, and we are interested in determining the rate at which this transfer occurs. As for the case of heat transfer, such a calculation may be based on the use of a convection coefficient.[3,5] In particular, we may relate the mass flux of species 1 to the product of a transfer coefficient and a concentration difference:

$$j_1 = h_{\mathbf{D}}(C_{1,w} - C_{1,\infty}) \tag{9.24}$$

Here $h_{\mathbf{D}}$ is the convection mass-transfer coefficient, and it has a dimension of L/t.

At the wall, $y = 0$, the fluid velocity is zero, and species transfer is due only to diffusion; hence

$$j_1 = -\mathbf{D} \left.\frac{\partial C_1}{\partial y}\right|_{y=0} \tag{9.25}$$

Combining Eqs. (9.24) and (9.25), it follows that

$$h_{\mathbf{D}} = -\frac{\mathbf{D} \left.(\partial C_1/\partial y)\right|_{y=0}}{C_{1,w} - C_{1,\infty}} \tag{9.26}$$

Therefore, conditions that influence the surface concentration gradient, $\left.(\partial C_1/\partial y)\right|_{y=0}$, will also influence the convection mass-transfer coefficient and the rate of species transfer across the fluid layer near the wall.

Conservation equations related to convective mass transfer are given in the section "The Conservation Equation for Species" in Chap. 10. Relations for estimating convective heat-transfer and mass-transfer coefficients for different flow geometries are given in Table 9.3. All symbols from Table 9.3 are defined in the list of nomenclature, see also Table 10.10 for physical interpretations of dimensionless groups.

Note. For convective processes involving high-velocity gas flows (high subsonic or supersonic flows), a more meaningful and useful definition of the heat-transfer coefficient is given by

$$q'' = h(T_w - T_{aw}) \tag{9.27}$$

Here T_{aw}, commonly called the *adiabatic wall temperature* or the *recovery temperature,* is the equilibrium temperature the surface would attain in the absence of any heat transfer to or from the surface and in the absence of radiation exchange be-

TABLE 9.3 Illustrative Relations for Estimating Convective Transfer Coefficients

(Isothermal wall conditions, * properties of fluids at reference temperature†; for mass transfer, replaced Pr with Sc and Nu with Sh in flow geometries 1 to 9.)

Flow geometry	Laminar flow	Transition	Turbulent flow
1. Flow parallel to a flat plate	$\mathrm{Nu}_x = 0.332\,\mathrm{Re}_x^{1/2}\mathrm{Pr}^{0.33}\,(\mathrm{Pr} \geq 0.5)$ $\mathrm{Nu}_x = 0.565\,\mathrm{Re}_x^{1/2}\mathrm{Pr}^{1/2}\,(\mathrm{Pr} \leq 0.025)$	$\mathrm{Re}_x \doteq 5 \times 10^5$	$\mathrm{Nu}_x = 0.296\,\mathrm{Re}_x^{0.8}\mathrm{Pr}^{0.6}(10.0 \geq \mathrm{Pr} \geq 0.5)$ $\mathrm{Nu}_L = 0.664\,\mathrm{Re}_{\mathrm{trans}}^{1/2}\mathrm{Pr}^{1/3}$ $+\dfrac{5}{4}\left[1 - \left(\dfrac{\mathrm{Re}_{\mathrm{trans}}}{\mathrm{Re}_L}\right)^{0.8}\right](0.0296)\mathrm{Re}_L^{0.8}\mathrm{Pr}^{0.6}$
2. Flow in a duct (a) A straight pipe	$\mathrm{Nu}_D = 3.65 + \dfrac{0.065(D/L)\mathrm{Re}_D\mathrm{Pr}}{1 + 0.04[(D/L)\mathrm{Re}_D\mathrm{Pr}]^{2/3}}$	$\mathrm{Re}_D \doteq 2300$	$\mathrm{Nu}_D = 0.23\,\mathrm{Re}_D^{0.8}\mathrm{Pr}^{0.33},\ \mathrm{Pr} > 0.5$ $\mathrm{Nu}_D = 4.8 + 0.003\,\mathrm{Re}_D\mathrm{Pr};\ \mathrm{Pr} < 0.1$
(b) Parallel plates	$\mathrm{Nu}_{D_h} = 7.54 + \dfrac{0.03(D_h/L)\mathrm{Re}_{D_h}\mathrm{Pr}}{1 + 0.016[D_h/L)\mathrm{Re}_{D_h}\mathrm{Pr}]^{2/3}}$	$\mathrm{Re}_{D_h} \doteq 2800$	Nu_{D_h} same as 2a
3. Flow across a circular cylinder	$\mathrm{Nu}_D = 0.3 + \dfrac{0.62\mathrm{Re}_D^{1/2}\mathrm{Pr}^{1/3}}{[1 + (0.4/\mathrm{Pr})^{2/3}]^{1/4}}$ $\mathrm{Re}_D\mathrm{Pr} > 0.2$ $\mathrm{Re}_D < 10{,}000$	$\mathrm{Re}_D \doteq 10{,}000$	$\mathrm{Nu}_D = \mathrm{Nu}_D(\text{laminar})$ $+\left[1 + \left(\dfrac{\mathrm{Re}_D}{282{,}000}\right)^{5/8}\right]^{4/5}$

4. Flow across a sphere

$Nu = 2 + 0.3\ Re^{0.6}Pr^{0.33}$, $Pr \geq 0.6$
$Nu = 2 + 0.4(RePr)^{0.5}$; $Pr < 0.6$

$Re_D \doteq 150{,}000$

5. Free convection on a horizontal cylinder

$$Nu_D = \frac{2}{\ln\left[1 + 4/(Gr_D Pr)^{1/4}\right]}$$

$Gr_D Pr \doteq 10^9$

$$Nu_D^{1/2} = 0.60 + 0.387 \times \left(\frac{Gr_D Pr}{[1 + (0.559/Pr)^{9/16}]^{16/9}}\right)^{1/6}$$

6. Free convection on a sphere
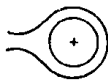
$Nu_D = 2 + K(Gr_D Pr)^{1/4}$
$0 < Gr_D Pr < 50$; $K = 0.3$
$50 < Gr_D Pr < 200$; $K = 0.4$
$200 < Gr_D Pr < 10^6$; $K = 0.5$
$10^6 < Gr_D Pr < 10^8$; $K = 0.6$

$Gr_D \doteq 10^8$

7. Free convection on a vertical wall
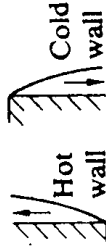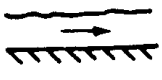
$$Nu_x = 0.508\ \frac{Gr_x^{1/4} Pr^{1/2}}{[0.952 + Pr]^{1/4}}$$

$$Nu_L = \frac{4}{3}\ Nu_x(x = L)$$

$Gr_x \doteq 10^8$

$$Nu_x = 0.149(Pr^{0.175} - 0.55)Gr_x^{0.36}$$

$$Nu_L = \left(\frac{4}{3}\right)(0.508)\ \frac{Gr_{\text{trans}}^{1/4} Pr^{1/2}}{[0.952 + Pr]^{1/4}}$$
$$+ \frac{1}{1.08}\left[1 - \left(\frac{Gr_{\text{trans}}}{Gr_L}\right)^{0.92}\right](0.149)$$
$$\times (Pr^{0.175} - 0.55)Gr_L^{0.36}$$

TABLE 9.3 Illustrative Relations for Estimating Convective Transfer Coefficients (*Continued*)

(*Isothermal wall conditions, * properties of fluids at reference temperature†; for mass transfer, replaced Pr with Sc and Nu with Sh in flow geometries 1 to 9.*)

Flow geometry	Laminar flow	Transition	Turbulent flow
8. Free convection on a horizontal square (a) Facing up	$Nu_L = 0.54(Gr_L Pr)^{1/4}$ $10^3 < Gr_L < 7 \times 10^7$ (For the rectangle use the shorter side for L)	$Gr_L \doteq 7 \times 10^7$	$Nu_L = 0.12(Gr_L Pr)^{1/3}$ (Note that L is of no importance)
(b) Facing down	$Nu_L = 0.27 \, (Gr_L Pr)^{1/4}$ $3 \times 10^5 < Gr_L < 3 \times 10^{10}$		
9. Flow through a packed bed of spheres	$St Pr^{2/3} = 1.625 Re_D^{-1/2}$ $15 < Re_D < 120$ $Re_D = \dfrac{\dot{m}D}{A\mu}$ D: sphere diameter A: bed cross-sectional area	$Re_D = 120$ $\dfrac{\mathscr{P}}{A} = \dfrac{6(1 - \epsilon_v)}{D}$ ϵ_v: volume void fraction $\mathscr{P}$: perimeter ($=$ mass-transfer surface area per unit bed length)	$St Pr^{2/3} = 0.687 Re^{-0.327}$ $120 < Re < 2000$
10. Falling film evaporation	Nusselt: $Nu = 1.10 Re_F^{-1/3}; \ Re_F < 30$ Wavy laminar: $Nu = 0.82 Re_F^{-0.22}; \ Re_F > 30$ $Nu = \dfrac{h(v^2/g)^{1/3}}{k}; \ Re_F = \dfrac{4\Gamma}{\mu}$	$Re_F = 1000 - 1800$ (Use the larger of the wavy laminar and turbulent film correlations.)	$Nu = 3.8 \times 10^{-3} Re_F^{0.4} Pr^{0.65}$

9.16

11. Film condensation on a vertical surface

Nusselt:

$\mathrm{Nu}_x = [\mathrm{Gr}_x\mathrm{Pr}_l/4\mathrm{Ja}_l]^{1.4}$; $\mathrm{Re}_F < 30$

Wavy laminar

$$\frac{(h_L v_L^2/g)^{1/3}}{k_L} = \frac{\mathrm{Re}_F}{1.08\mathrm{Re}_F^{1.22} - 5.2};$$

$\mathrm{Re}_F > 30$

$\mathrm{Re}_F = 1000 - 1800$
(See item 10.)

12. Film condensation on tubes

$$\mathrm{Nu}_D = [\mathrm{Gr}_D\mathrm{Pr}_l/3.6N\mathrm{Ja}_l]^{1/4}$$

N: number of tubes

13. Nucleate, saturated, pool boiling
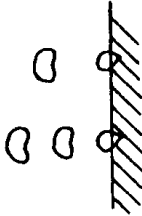
$$\mathrm{Nu}_{Lcl} = \frac{\mathrm{Ja}_L^2}{C^3\mathrm{Pr}_l^m}$$

$$L_c = \sqrt{\sigma_l/g(\rho_l - \rho_v)}$$

$m = 2$ for water

$m = 4.1$ for other fluids

$C = 0.013$ water-copper or stainless steel

$C = 0.006$ water-nickel or brass

$C = 0.027$ ethyl alcohol-chromium

$q \leq q_{\max}$

$$\mathrm{St}_{\max} = \frac{q_{\max}}{\rho_v V_{\max} i_{lg}} = 0.145$$

$$V_{\max} = [\sigma_l(\rho_l - \rho_v)g/\rho_v^2]^{1/4}$$

TABLE 9.3 Illustrative Relations for Estimating Convective Transfer Coefficients (*Continued*)

(Isothermal wall conditions, * properties of fluids at reference termperature†; for mass transfer, replaced Pr with Sc and Nu with Sh in flow geometries 1 to 9.)

Flow geometry	Laminar flow	Transition	Turbulent flow
14. Film boiling on a horizontal plate	$\overline{Nu}_{L_{c,v}} = 0.425 \left[Gr_{L_{c,v}} Pr_v \left(\frac{1 + 0.4Ja_v}{Ja_v} \right) \right]^{1/4}$ $Gr_{L_{c,v}} = \dfrac{g[(\rho_l - \rho_v)/\rho_v]L_c^3}{\omega_v^2}$ $L_c = \sqrt{\sigma_l/g(\rho_l - \rho_v)}$		$q \geq q_{min}$ $St_{min} = \dfrac{q_{min}}{\rho_v V_{min} i_{fg}} = 0.09$ $V_{min} = [\sigma_l(\rho_l - \rho_v)g_c/(\rho_l + \rho_v)^2]^{1/4}$
15. Film boiling on a horizontal cylinder	$\overline{Nu}_{D,v} = 0.62 \left[Gr_{D,v} Pr_v \left(\frac{1 + 0.5Ja_v}{Ja_v} \right) \right]^{1/4}$ $Gr_{D,v} = \dfrac{g[(\rho_l - \rho_v)/\rho_v]D^2}{v_v^2}$		
16. Film boiling on a sphere	$\overline{Nu}_{D,v} = 0.4 \left[Gr_{D,v} Pr_v \left(\frac{1 + 0.4Ja_v}{Ja_v} \right) \right]^{1/3}$ $Gr_{D,v} = \dfrac{g[(\rho_l - \rho_v)/\rho_v]D^3}{v_v^2}$		

*For external flows $\overline{Nu}$ same for uniform wall heat flux.
†$T_{ref} = T_a + \alpha(T_{a\,(or\,b)} - T_a)$ for low-speed flows. For general engineering use $\alpha \doteq \frac{1}{2}$, but often more precise specifications are available; e.g., for laminar natural convection of gases $\alpha = 0.38$ except for $\beta = 1/T_a$. Dimension less parameters are also defined in Table 10.10.

Source: Adapted from Edwards, Denny, and Mills,[6] pp. 166–170.

tween the surroundings and the surface. In general, the adiabatic wall temperature is dependent on the fluid properties and the properties of the bounding wall. Generally, the adiabatic wall temperature is reported in terms of a dimensionless recovery factor **r** defined as

$$T_{aw} = T_f + \mathbf{r}\,\frac{V^2}{2c_p} \tag{9.28}$$

The value of **r** for gases normally lies between 0.8 and 1.0. It can be seen that for low-velocity flows the recovery temperature is equal to the free-stream temperature T_f. In this case, Eq. (9.27) reduces to Eq. (9.20). From this point of view, Eq. (9.27) can be taken as the generalized definition of the heat-transfer coefficient.

COMBINED HEAT-TRANSFER MECHANISMS

In practice, heat transfer frequently occurs by two mechanisms in parallel. A typical example is shown in Fig. 9.7. In this case, the heat conducted through the plate is removed from the plate surface by a combination of convection and radiation. An energy balance in this case gives

$$-k_s A\,\frac{dT}{dy}\bigg|_w = hA(T_w - T_\infty) + \sigma A\epsilon(T_w^4 - T_a^4) \tag{9.29}$$

where T_a is the temperature of the surroundings, k_s is the thermal conductivity of the solid plate, and ϵ is the emissivity of the plate (i.e., in this special case, $\mathfrak{F}_{1-2} \equiv \epsilon$, since the area of the plate is much smaller than the area of the surroundings[3]). The plate and the surroundings are separated by a gas that has no effect on radiation.

There are many applications where radiation is combined with other modes of heat transfer, and the solution of such problems can often be simplified by using a thermal resistance R_{th} for radiation. The definition of R_{th} is similar to that of the
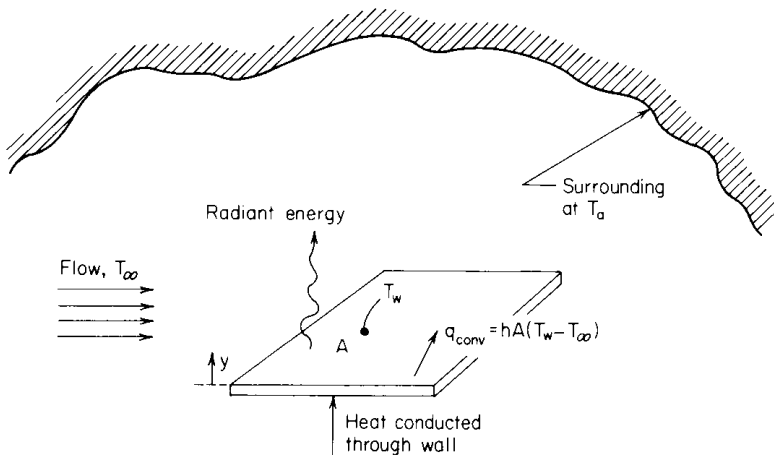
FIGURE 9.7 Combination of conduction, convection, and radiation heat transfer. (*From Rohsenow, Hartnett, and Ganić,*[5] *p. 1-12.*)

thermal resistance for convection and conduction. If the heat transfer by radiation, for the example in Fig. 9.7, is written

$$q = \frac{T_w - T_a}{R_{th}} \tag{9.30}$$

the resistance is given by

$$R_{th} = \frac{T_w - T_a}{\sigma A \epsilon (T_w^4 - T_a^4)} \tag{9.31}$$

Also, a heat-transfer coefficient h_r can be defined for radiation:

$$h_r = \frac{1}{R_{th}A} = \frac{\sigma \epsilon (T_w^4 - T_a^4)}{T_w - T_a} = \sigma \epsilon (T_w + T_a)(T_w^2 + T_a^2) \tag{9.32}$$

Here we have linearized the radiation rate equation, making the heat rate proportional to a temperature difference rather than to the difference between two temperatures to the fourth power. Note that h_r depends strongly on temperature, while the temperature dependence of the convection heat-transfer coefficient h is generally weak.

Another example where all three heat transfer mechanisms are present is the vacuum flask. A cross-section of a typical vacuum flask is shown in Fig. 9.8 and is seen to be a double-walled bottle with a vacuum space between them, the whole supported in a protective outer case.

Very little heat can be transferred by conduction because of the vacuum space and the cork stopper (cork is a bad conductor of heat). Also, because of the vacuum space, no convection is possible. Radiation is minimized by silvering the two glass surfaces (radiation is reflected off shining surfaces).

Thus a vacuum flask is an example of prevention of all three types of heat transfer and is, therefore, able to keep hot liquid hot and cold liquids cold.

Note on Insulation. By the careful use of insulation, heat can be retained in a building for longer periods, thereby minimizing the cost of heating. Since convection causes hot air to rise, it is important to insulate the roof space, which is probably the greatest source of heat loss in the home. This can be achieved by laying fiber-glass between the wooden joists in the roof space.
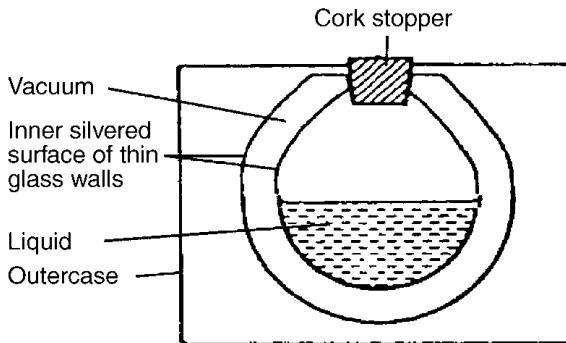
FIGURE 9.8 Vacuum flask.

Glass is a poor conductor of heat. However, large losses can occur through thin panes of glass and such losses can be reduced by using double-glazing. Two sheets of glass, separated by air, are used. Air is a very good insulator, but the air space must not be too large, otherwise convection currents can occur that would carry heat across the space.

Hot water tanks should be lagged to prevent conduction and convection or heat to the surrounding air.

Brick, concrete, plaster and wood are all poor conductors of heat. A house is made from two walls with an air gap between them. Air is a poor conductor and trapped air minimizes losses through the wall. Heat losses through the walls can be prevented almost completely by using cavity wall insulation, i.e. plastic-foam.

HEAT EXCHANGERS

The application of principles of heat transfer to the design of equipment to accomplish a certain engineering objective is of extreme importance. In a heat exchanger, heat transfer from one fluid to another occurs either by direct contact or through an intervening wall. The following section deals with the recuperative-type heat exchanger in which the fluids exchange heat through a wall.

One fluid flows through a series of tubes and the other through a shell surrounding them (shell-and-tube heat exchanger). Flow may be either "parallel" (Fig. 9.9) (both fluids moving in the same direction) or "counter flow" (Fig. 9.10) (fluids moving in opposite directions). Another possibility is the "cross flow" arrangement in which the flows are at right angles.

$$\Delta T_1 = T_{h1} - T_{c1}, \ \Delta T_2 = T_{h2} - T_{c2}$$

(a)   (b)
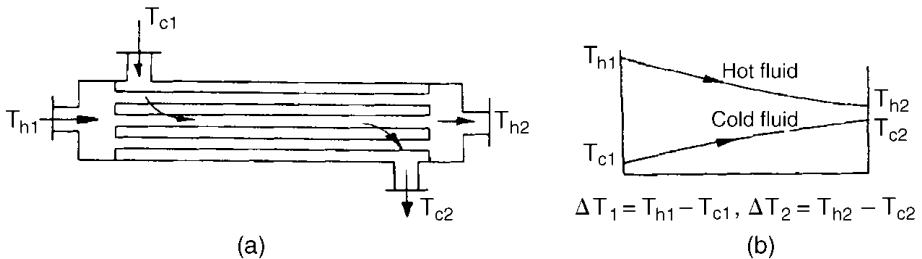
FIGURE 9.9   Parallel flow (*a*) heat exchanger and (*b*) temperature distribution.

$$\Delta T_1 = T_{h1} - T_{c2}, \ \Delta T_2 = T_{h2} - T_{c1}$$

(a)   (b)
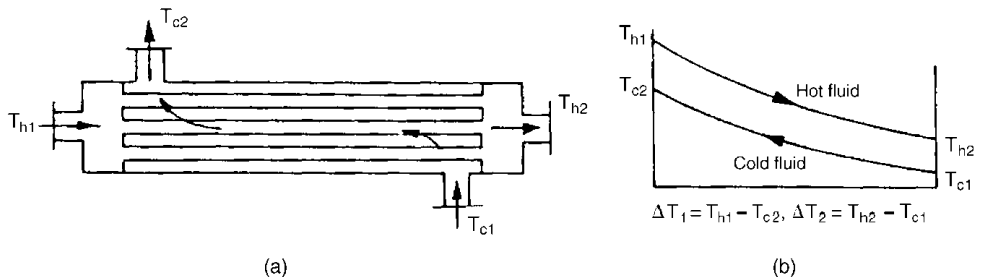
FIGURE 9.10   Counter flow (*a*) heat exchanger and (*b*) temperature distribution.

Other types have more complex flows: e.g., the "multi-pass" and "mixed-flow" types.

The heat transfer rate is calculated from Eq. 9.8 where U is the overall heat transfer coefficient given by Eq. 9.11; A is heat transfer area (surface area of tubes).

The following relations give the heat transferred, the logarithmic mean temperature difference ΔT_m and the "effectiveness" ϵ.

$$\Delta T_m = \frac{\Delta T_2 - \Delta T_1}{\ln \dfrac{\Delta T_2}{\Delta T_1}} \qquad (9.33)$$

$$\epsilon = \frac{\text{actual heat transfer}}{\text{maximum possible heat transfer}} = \frac{T_{c2} - T_{c1}}{T_{h1} - T_{c1}} \qquad (9.34)$$

where heat transfer is

$$q = U \cdot A \cdot \Delta T_m$$

and

$$U \approx \frac{1}{1/h_a + 1/h_c}$$

assuming that the main resistance to heat transfer is on the fluid side, i.e., the heat transfer conduction resistance through the tube wall is much smaller.

$$A = \text{mean surface area of tubes}$$

$$h_a = \text{heat transfer coefficient for hot fluid side}$$

$$h_c = \text{heat transfer coefficient for cold fluid side}$$

It is important to note that if one of the fluids is a condensing vapor or a boiling liquid, the temperature is constant on that side, $T_1 = T_2$.

Counter Flow (Fig. 9.10). For this case the temperature range possible is greater than for the parallel-flow type (Fig. 9.9). The same formulas apply.

Cross Flow. Instead of using ΔT_m as above, $\Delta T_m \cdot F$ is used, where F is a factor obtained from tables (see Ref. 9).

$$q = U \cdot A \cdot F \cdot \Delta T_m$$

If one fluid is a condensing vapor (constant temperature), ΔT_m is the same as for parallel-flow and counter-flow types.

RELATION OF HEAT TRANSFER TO THERMODYNAMICS

The subjects of thermodynamics and heat transfer are highly complementary, although some fundamental differences exist between them. Although thermodynamics is concerned with heat interaction and the vital role that it plays in the first and second laws, it considers neither the basic mechanisms that provide for heat

exchange nor the methods that exist for computing the rate of heat exchange.[3,7] Thermodynamics is concerned with equilibrium states. More specifically, thermodynamics may be used to determine the amount of energy required in the form of heat for a system to pass from one equilibrium state to another. The discipline of heat transfer does what thermodynamics is inherently unable to do. It considers the rate at which heat transfer occurs in terms of thermal nonequilibrium. This is done through the introduction of a new set of physical principles. These principles are transport laws, which are not a part of the subject of thermodynamics. They include Fourier's law, the Stefan-Boltzmann law, and Newton's law of cooling, which were defined earlier. In summary, it is important to remember that a description of heat transfer requires that the additional principles be combined with the first law of thermodynamics.

NOMENCLATURE

Symbol = Definition, SI Units (U.S. Customary Units)

C = mass concentration of species, kg/m^3 (lb_m/ft^3)

c = specific heat, $J/(kg \cdot K)$ [$Btu/(lb_m \cdot °F)$]

c_p = specific heat at constant pressure, $J/(kg \cdot K)$[$Btu/(lb_m \cdot °F)$]

c_v = specific heat at constant volume, $J/(kg \cdot K)$ [$Btu/(lb_m \cdot °F)$]

D = tube inside diameter; tube outside diameter (Table 9.3); diameter, m (ft)

D_h = hydraulic diameter (4 × flow area/wetted perimeter), m (ft)

D = diffusion coefficient, m^2/s (ft^2/s)

e = emissive power, W/m^2 [$Btu/(h \cdot ft^2)$]

e_b = blackbody emissive power, W/m^2 [$Btu/(h \cdot ft^2)$]

F_{1-2} = view factor (geometric shape factor for radiation from one blackbody to another)

$\mathscr{F}_{1-2}$ = real body view factor (geometric shape and emissivity factor for radiation from one gray body to another)

Gr = Grashof number (Table 9.3 and Table 10.10)

g = gravitational acceleration, m/s^2 (ft/s^2)

h = heat-transfer coefficient, $W/(m^2 \cdot K)$ [$Btu/(h \cdot ft^2 \cdot °F)$]

h_D = mass-transfer coefficient, m/s (ft/s)

i = enthalpy per unit mass, J/kg (Btu/lb_m)

i_{lg} = latent heat of evaporation, J/kg (Btu/lb_m)

j = mass diffusion flux of species, $kg/(s \cdot m^2)$ [$lb_m/(h \cdot ft^2)$]

j = mass diffusion flux of species (vector), $kg/(s \cdot m^2)$ [$lb_m/(h \cdot ft^2)$]

k = thermal conductivity, $W/(m \cdot K)$ [$Btu/(h \cdot ft \cdot °F)$]

L = length, m (ft)

m = mass flow rate, kg/s (lb/s)

Nu = Nusselt number (Table 9.3 and Table 10.10)

P = pressure, Pa, N/m^2 (lb_f/ft^2)

Pr = prandtl number (Table 9.3 and Table 10.10)

ΔP = pressure drop, Pa, N/m² (lb$_f$/ft²)

q = heat-transfer rate, W (Btu/h)

$\mathbf{q}''$ = heat flux (vector), W/m² [Btu/(h · ft²)]

q'' = heat flux, W/m² [Btu/(h · ft²)]

q''' = volumetric heat generation, W/m³ [Btu/(h · ft³)]

R_{th} = thermal resistance, K/W (h · °F/Btu)

Re = Reynolds number

Re_x = local Reynolds number (= $\rho V x / \mu$)

r = radial distance in cylindrical or spherical coordinate, m (ft)

$\mathbf{r}$ = recovery factor, Eq. (9.28)

Sc = Schmidt number (Table 9.3 and Table 10.10)

Sh = Sherwood number (Table 9.3 and Table 10.10)

St = Stanton number (Table 9.3 and Table 10.10)

T = temperatures °C, K (°F, °R)

ΔT = temperature difference, °C(°F)

t = time, s

u = velocity component in the axial direction (x direction) in rectangular coordinates, m/s (ft/s)

V = velocity, m/s (ft/s)

$\mathbf{V}$ = velocity (vector), m/s (ft/s)

v = velocity component in the y direction in rectangular coordinates, m/s (ft/s)

x = rectangular coordinate, m (ft)

y = rectangular coordinate, m (ft)

z = rectangular or cylindrical coordinate, m (ft)

Greek

α = thermal diffusivity, m²/s (ft²/s)

β = coefficient of thermal expansion, K^{-1} (°R^{-1})

δ = hydrodynamic boundary-layer thickness, in (ft)

δ_D = concentration boundary-layer thickness, m (ft)

δ_T = thermal boundary-layer thickness, m (ft)

ϵ = emissivity

Γ = film flow rate, kg/s · m (lb/s · ft)

μ = dynamic viscosity, Pa · s [lb/(h · ft)]

ν = kinematic viscosity, m²/s (ft²/s)

ρ = density, kg/m³ (lb/ft³)

σ = surface tension (Table 9.3), N/m (lb$_f$/ft); Stefan-Boltzmann radiation constant, W/(m$^2 \cdot$ K^4) [Btu/(h $\cdot$ ft$^2 \cdot$ °R^4)]

σ_t = surface tension (Table 9.3), N/m (lb$_f$/ft)

Subscripts

a = surroundings

aw = adiabatic wall

cr = critical

f = fluid

g = gas (vapor)

i = species i

l = liquid

m = mean

r = radiation, Eq. (9.36)

s = solid

sat = saturation

t = total

w = wall

x = x component

y = y component

z = z component

θ = θ component

ϕ = ϕ component

v = vapor

max = maximum

1 = species 1 in binary mixture of 1 and 2

∞ = free-stream condition

Superscripts

$^-$ = average (e.g., $\overline{X}$ is the average of X)

Mathematical Operation Symbols

d/dx = derivative with respect to x, m^{-1} (ft^{-1})

$\partial/\partial t$ = partial time derivative operator, s^{-1}

d/dt = total time derivative operator, s^{-1}

D/Dt = substantial time derivative operator, s^{-1}

∇ = "del" operator (vector), m^{-1} (ft^{-1})

∇^2 = Laplacian operator, m^{-2} (ft^{-2})

*REFERENCES**

1. F. Kreith and V. Z. Black, *Basic Heat Transfer,* Harper & Row, New York, 1980.

2. J. P. Holman, *Heat Transfer,* 8th ed., McGraw-Hill, New York, 1997.

3. F. P. Incropera and D. P. DeWitt, *Fundamentals of Heat Transfer,* Wiley, New York 1981.

4. M. N. Ozisik, *Basic Heat Transfer,* McGraw-Hill, New York, 1977.

5. W. M. Rohsenow, J. P. Hartnett, and E. N. Ganić, eds., *Handbook of Heat Transfer Fundamentals,* McGraw-Hill, New York, 1985.

6. D. K. Edwards, V. E. Denny, and A. F. Mills, *Transfer Processes: An Introduction to Diffusion, Convection, and Radiation,* 2d ed., Hemisphere, Washington, and McGraw-Hill, New York, 1979.

7. J. H. Lienhard, *A Heat Transfer Textbook,* Prentice-Hall, Englewood Cliffs, N.J., 1981.

8. R. E. Treybal, *Mass-Transfer Operations,* 3d ed., McGraw-Hill, New York, 1980.

9. W. M. Rohsenow, J. P. Hartnett, and E. N. Ganić, eds., *Handbook of Heat Transfer Applications,* McGraw-Hill, New York, 1985.

10. V. P. Isachenkon, V. A. Osipova and A. S. Sukomel, *Heat Transfer,* Mir Publishers, 1980 (in Russian).

11. W. C. Reynolds and H. C. Perkins, *Engineering Thermodynamics,* McGraw-Hill, 1977.

12. E. N. Ganic, *Heat and Mass Transfer* (in Bosnian), University of Sarajevo, 2002.

13. A. Bejan, *Convection Heat Transfer,* Wiley, 1984.

14. J. F. Keffer, R. K. Shah, and E. N. Ganić, Eds., *Experimental Heat Transfer, Fluid Mechanics and Thermodynamics,* Elsevier, 1992.

---

* Those references listed here but not cited in the text were used for comparison of different data sources, clarification, clarity of presentation, and, most important, reader's convenience when further interest in the subject exists.

CHAPTER 10

CONSERVATION EQUATIONS AND DIMENSIONLESS GROUPS

CONSERVATION EOUATIONS IN FLUID MECHANICS, HEAT TRANSFER, AND MASS TRANSFER

Each time we try to solve a new problem related to momentum and heat and mass transfer in a fluid, it is convenient to start with a set of equations based on basic laws of conservation for physical systems. These equations include

1. The continuity equation (conservation of mass)
2. The equation of motion (conservation of momentum)
3. The energy equation (conservation of energy, or the first law of thermodynamics)
4. The conservation equation for species (conservation of species)

These equations are sometimes called the *equations of change,* inasmuch as they describe the change of velocity, temperature, and concentration with respect to time and position in the system.

The first three equations are sufficient for problems involving a pure fluid (a *pure* substance is a single substance characterized by an unvarying chemical structure). The fourth equation is added for a mixture of chemical species, i.e., when mass diffusion with or without chemical reactions is present.

- *The control volume:* When deriving the conservation equations, it is necessary to select a control volume. The derivation can be performed for a volume element of any shape in a given coordinate system, although the most convenient shape is usually assumed for simplicity (e.g., a rectangular shape in a rectangular coordinate system). For illustration purposes, different coordinate systems are shown in Fig. 10.1. In selecting a control volume, we have the option of using a volume fixed in space, in which case the fluid flows through the boundaries, or a volume containing a fixed mass of fluid and moving with the fluid. The former is known
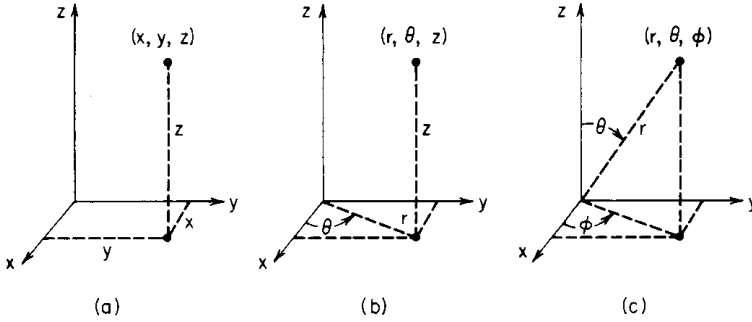
FIGURE 10.1   Coordinate systems: (*a*) rectangular, (*b*) cylindrical, (*c*) spherical.

as the *eulerian viewpoint,* and the latter is the *lagrangian viewpoint.* Both approaches yield equivalent results.

- *The partial time derivative $\partial B / \partial t$:*   The partial time derivative of $B(x,y,z,t)$, where B is any continuum property (e.g., density, velocity, temperature, concentration, etc.), represents the change of B with time at a fixed position in space. In other words, $\partial B / \partial t$ is the change of B with t as seen by a *stationary observer.*

- *Total time derivative dB/dt:* The total time derivative is related to the partial time derivative as follows:

$$\frac{dB}{dt} = \frac{\partial B}{\partial t} + \frac{dx}{dt}\frac{\partial B}{\partial x} + \frac{dy}{dt}\frac{\partial B}{\partial y} + \frac{dz}{dt}\frac{\partial B}{\partial z} \tag{10.1}$$

where dx/dt, dy/dt, and dz/dt are the components of the velocity of the moving observer. Therefore, dB/dt is the change of B with time as seen by the moving observer.

- *Substantial time derivative DB/Dt:*   This derivative is a special kind of total time derivative where now the velocity of the observer is just the same as the velocity of the stream; i.e., the observer drifts along with the current:

$$\frac{DB}{Dt} = \frac{\partial B}{\partial t} + u\frac{\partial B}{\partial x} + v\frac{\partial B}{\partial y} + w\frac{\partial B}{\partial z} \tag{10.2}$$

where u, v, and w are the components of the local fluid velocity $\mathbf{V}$. The substantial time derivative is also called the *derivative following the motion.* The sum of the last three terms on the right-hand side of Eq. (10.2) is called the *convective contribution* because it represents the change in B due to translation.

The use of the operator D/Dt is always made when rearranging various conservation equations related to the volume element fixed in space to an element following the fluid motion. The operator D/Dt also may be expressed in vector form:

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + (\mathbf{V} \cdot \nabla) \tag{10.3}$$

Mathematical operations involving ∇ are given in many textbooks. Applications of ∇ in various operations involving the conservation equations are given in Refs. 1 and 2. Table 10.1 gives the expressions for D/Dt in different coordinate systems.

TABLE 10.1 Substantial Derivative in Different Coordinate Systems

---

Rectangular coordinates (x, y, z):

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}$$

Cylindrical coordinates (r, θ, x):

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + v_r\frac{\partial}{\partial r} + \frac{v_\theta}{r}\frac{\partial}{\partial \theta} + v_z\frac{\partial}{\partial z}$$

Spherical coordinates (r, θ, ϕ):

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + v_r\frac{\partial}{\partial r} + \frac{v_\theta}{r}\frac{\partial}{\partial \theta} + \frac{v_\phi}{r \sin \theta}\frac{\partial}{\partial \phi}$$

---

1. The Equation of Continuity.   For a volume element fixed in space,

$$\frac{\partial \rho}{\partial t} = -\underbrace{(\nabla \cdot \rho\mathbf{V})}_{\substack{\text{net rate of mass} \\ \text{efflux per unit} \\ \text{volume}}} \tag{10.4}$$

The continuity equation in this form describes the rate of change of density at a fixed point in the fluid. By performing the indicated differentiation on the right side of Eq. (10.4) and collecting all derivatives of ρ on the left side, we obtain an equivalent form of the equation of continuity:

$$\frac{D\rho}{Dt} = -\rho(\nabla \cdot \mathbf{V}) \tag{10.5}$$

The continuity equation in this form describes the rate of change of density as seen by an observer "floating along" with the fluid.

For a fluid of constant density (incompressible fluid), the equation of continuity becomes

$$\nabla \cdot \mathbf{V} = 0 \tag{10.6}$$

Table 10.2 gives the equation of continuity in different coordinate systems.

2. The Equation of Motion (Momentum Equation).   The momentum equation for a stationary volume element (i.e., a balance over a volume element fixed in space) with gravity as the only body force is given by

$$\underbrace{\frac{\partial \rho\mathbf{V}}{\partial t}}_{\substack{\text{rate of increase} \\ \text{of momentum} \\ \text{per unit volume}}} = \underbrace{-(\nabla \cdot \rho\mathbf{V})\mathbf{V}}_{\substack{\text{rate of} \\ \text{momentum gain} \\ \text{by convection} \\ \text{per unit volume}}} - \underbrace{\nabla P}_{\substack{\text{pressure force} \\ \text{on element per} \\ \text{unit volume}}} + \underbrace{\nabla \cdot \tau}_{\substack{\text{rate of} \\ \text{momentum gain} \\ \text{by viscous} \\ \text{transfer per unit} \\ \text{volume}}} + \underbrace{\rho\mathbf{g}}_{\substack{\text{gravitational} \\ \text{force on} \\ \text{element per} \\ \text{unit volume}}} \tag{10.7}$$

Equation (10.7) may be rearranged, with the help of the equation of continuity, to give

TABLE 10.2 Equation of Continuity in Different Coordinate Systems

---

Rectangular coordinates (x, y, z):

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} (\rho u) + \frac{\partial}{\partial y} (\rho v) + \frac{\partial}{\partial z} (\rho w) = 0 \tag{A}$$

Cylindrical coordinates (r, θ, z):

$$\frac{\partial \rho}{\partial t} + \frac{1}{r} \frac{\partial}{\partial r} (\rho r v_r) + \frac{1}{r} \frac{\partial}{\partial \theta} (\rho v_\theta) + \frac{\partial}{\partial z} (\rho v_z) = 0 \tag{B}$$

Spherical coordinates (r, θ, ϕ):

$$\frac{\partial \rho}{\partial t} + \frac{1}{r^2} \frac{\partial}{\partial r} (\rho r^2 v_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\rho v_\theta \sin \theta) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} (\rho v_\phi) = 0 \tag{C}$$

Incompressible flow

Rectangular coordinates (x, y, z):

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \tag{D}$$

Cylindrical coordinates (r, θ, z):

$$\frac{1}{r} \frac{\partial}{\partial r} (r v_r) + \frac{1}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{\partial v_z}{\partial z} = 0 \tag{E}$$

Spherical coordinates (r, θ, ϕ):

$$\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 v_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (v_\theta \sin \theta) + \frac{1}{r \sin \theta} \frac{\partial v_\phi}{\partial \phi} = 0 \tag{F}$$

$$\rho \frac{D\mathbf{V}}{Dt} = -\nabla P + \nabla \cdot \tau + \rho \mathbf{g} \tag{10.8}$$

The last equation is a statement of Newton's second law of motion in the form *mass* × *acceleration* = *sum of forces.*

These two forms of the equation of motion, Eqs. (10.7) and (10.8), correspond to the two forms of the equation of continuity, Eqs. (10.4) and (10.5).

As indicated above, the only body force included in Eqs. (10.7) and (10.8) is gravity. In general, electromagnetic forces also may act on a fluid.[4]

The scalar components of Eq. (10.8) are listed in Table 10.3, and the components of the stress tensor τ are given in Table 10.4.

For the flow of a newtonian fluid with varying density but constant viscosity μ, Eq. (10.8) becomes

$$\rho \frac{D\mathbf{V}}{Dt} = -\nabla P + \frac{1}{3} \mu \nabla (\nabla \cdot V) + \mu \nabla^2 \mathbf{V} + \rho \mathbf{g} \tag{10.9}$$

If ρ and μ are constant, Eq. (10.9) may be simplified by means of the equation of continuity $(\nabla \cdot \mathbf{V} = 0)$ to give for a newtonian fluid

$$\rho \frac{D\mathbf{V}}{Dt} = -\nabla P + \mu \nabla^2 \mathbf{V} + \rho \mathbf{g} \qquad (10.10)$$

This is the famous Navier-Stokes equation in vector form. The scalar components of Eq. (10.10) are given in Table 10.5.

For $\nabla \cdot \tau = 0$, Eq. (10.8) reduces to Euler's equation:

$$\rho \frac{D\mathbf{V}}{Dt} = -\nabla P + \rho \mathbf{g} \qquad (10.11)$$

which is applicable for describing flow systems in which viscous effects are relatively unimportant.

TABLE 10.3   Equation of Motion in Terms of Viscous Stresses [Eq. (10.8)]*

Rectangular coordinates (x, y, z)

x direction

$$\rho \left(\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} \right) = -\frac{\partial P}{\partial x} + \left(\frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z} \right) + \rho g_x \qquad (A)$$

y direction

$$\rho \left(\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} \right) = -\frac{\partial P}{\partial y} + \left(\frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} + \frac{\partial \tau_{zy}}{\partial z} \right) + \rho g_y \qquad (B)$$

z direction

$$\rho \left(\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} \right) = -\frac{\partial P}{\partial z} + \left(\frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \tau_{zz}}{\partial z} \right) + \rho g_z \qquad (C)$$

Cylindrical coordinates (r, θ, z)

r direction

$$\rho \left(\frac{\partial v_r}{\partial t} + v_r \frac{\partial v_r}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_r}{\partial \theta} - \frac{v_\theta^2}{r} + v_z \frac{\partial v_r}{\partial z} \right)$$
$$= -\frac{\partial P}{\partial r} + \left[\frac{1}{r} \frac{\partial}{\partial r} (r\tau_{rr}) + \frac{1}{r} \frac{\partial \tau_{\theta r}}{\partial \theta} - \frac{\tau_{\theta\theta}}{r} + \frac{\partial \tau_{zy}}{\partial z} \right] + \rho g_r \qquad (A)$$

θ direction

$$\rho \left(\frac{\partial v_\theta}{\partial t} + v_r \frac{\partial v_\theta}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{v_r v_\theta}{r} + v_z \frac{\partial v_\theta}{\partial z} \right)$$
$$= -\frac{1}{r} \frac{\partial P}{\partial \theta} + \left[\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \tau_{r\theta}) + \frac{1}{r} \frac{\partial \tau_{\theta\theta}}{\partial \theta} + \frac{\partial \tau_{z\theta}}{\partial z} \right] + \rho g_\theta \qquad (B)$$

z direction

$$\rho \left(\frac{\partial v_z}{\partial t} + v_r \frac{\partial v_z}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_z}{\partial \theta} + v_z \frac{\partial v_z}{\partial z} \right) = -\frac{\partial P}{\partial z} + \left[\frac{1}{r} \frac{\partial}{\partial r} (r\tau_{rz}) + \frac{1}{r} \frac{\partial \tau_{\theta z}}{\partial \theta} + \frac{\partial \tau_{zz}}{\partial z} \right] + \rho g_z \qquad (C)$$

TABLE 10.3 Equation of Motion in Terms of Viscous Stresses (*Continued*) [Eq. (10.8)]*

Spherical coordinates (r, θ, ϕ)

r direction

$$\rho \left(\frac{\partial v_r}{\partial t} + v_r \frac{\partial v_r}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_r}{\partial \theta} + \frac{v_\phi}{r \sin \theta} \frac{\partial v_r}{\partial \phi} - \frac{v_\theta^2 + v_\phi^2}{r} \right) = - \frac{\partial P}{\partial r} + \left[\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \tau_{rr}) \right.$$

$$\left. + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\tau_{\theta r} \sin \theta) + \frac{1}{r \sin \theta} \frac{\partial \tau_{\phi r}}{\partial \phi} - \frac{\tau_{\theta\theta} + \tau_{\phi\phi}}{r} \right] + \rho g_r \quad (A)$$

θ direction

$$\rho \left(\frac{\partial v_\theta}{\partial t} + v_r \frac{\partial v_\theta}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{v_\phi}{r \sin \theta} \frac{\partial v_\theta}{\partial \phi} + \frac{v_r v_\theta}{r} - \frac{v_\phi^2 \cot \theta}{r} \right) = - \frac{1}{r} \frac{\partial P}{\partial \theta}$$

$$+ \left[\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \tau_{r\theta}) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\tau_{\theta\theta} \sin \theta) + \frac{1}{r \sin \theta} \frac{\partial \tau_{\phi\theta}}{\partial \phi} + \frac{\tau_{r\theta}}{r} - \frac{\tau_{\phi\phi} \cot \theta}{r} \right] + \rho g_\theta \quad (B)$$

ϕ direction

$$\rho \left(\frac{\partial v_\phi}{\partial t} + v_r \frac{\partial v_\phi}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_\phi}{\partial \theta} + \frac{v_\phi}{r \sin \theta} \frac{\partial v_\phi}{\partial \phi} + \frac{v_\phi v_r}{r} + \frac{v_\theta v_\phi \cot \theta}{r} \right) = - \frac{1}{r \sin \theta} \frac{\partial P}{\partial \phi}$$

$$+ \left[\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \tau_{r\phi}) + \frac{1}{r} \frac{\partial \tau_{\theta\phi}}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial \tau_{\phi\phi}}{\partial \phi} + \frac{\tau_{r\phi}}{r} + \frac{2 \tau_{\theta\phi} \cot \theta}{r} \right] + \rho g_\phi \quad (C)$$

*Components of the stress tensor τ for newtonian fluids are given in Table 10.4. This equation also may be used for describing nonnewtonian flow. However, we need relations between the components of τ and the various velocity gradients; in other words, we have to replace the expressions given in Table 10.4 by other relations appropriate for the nonnewtonian fluid of interest. The expressions for τ for some non-newtonian fluid models are given in Ref. 2. See also Ref. 4.

TABLE 10.4 Components of the Stress Tensor τ for Newtonian Fluids*

Rectangular coordinates (x, y, z)

$$\tau_{xx} = \mu \left[2 \frac{\partial u}{\partial x} - \frac{2}{3} (\nabla \cdot \mathbf{V}) \right] \qquad (A)$$

$$\tau_{yy} = \mu \left[2 \frac{\partial v}{\partial y} - \frac{2}{3} (\nabla \cdot \mathbf{V}) \right] \qquad (B)$$

$$\tau_{zz} = \mu \left[2 \frac{\partial w}{\partial z} - \frac{2}{3} (\nabla \cdot \mathbf{V}) \right] \qquad (C)$$

$$\tau_{xy} = \tau_{yx} = \mu \left[\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right] \qquad (D)$$

$$\tau_{yz} = \tau_{zy} = \mu \left[\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right] \qquad (E)$$

$$\tau_{zx} = \tau_{xz} = \mu \left[\frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \right] \qquad (F)$$

$$(\nabla \cdot \mathbf{V}) = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \qquad (G)$$

TABLE 10.4 Components of the Stress Tensor τ for Newtonian Fluids* (*Continued*)

<center>Cylindrical coordinates (r, θ, z)</center>

$$\tau_{rr} = \mu \left[2 \frac{\partial v_r}{\partial r} - \frac{2}{3} (\nabla \cdot \mathbf{V}) \right] \tag{A}$$

$$\tau_{\theta\theta} = \mu \left[2 \left(\frac{1}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{v_r}{r} \right) - \frac{2}{3} (\nabla \cdot \mathbf{V}) \right] \tag{B}$$

$$\tau_{zz} = \mu \left[2 \frac{\partial v_z}{\partial z} - \frac{2}{3} (\nabla \cdot \mathbf{V}) \right] \tag{C}$$

$$\tau_{r\theta} = \tau_{\theta r} = \mu \left[r \frac{\partial}{\partial r} \left(\frac{v_\theta}{r} \right) + \frac{1}{r} \frac{\partial v_r}{\partial \theta} \right] \tag{D}$$

$$\tau_{\theta z} = \tau_{z\theta} = \mu \left[\frac{\partial v_\theta}{\partial z} + \frac{1}{r} \frac{\partial v_z}{\partial \theta} \right] \tag{E}$$

$$\tau_{zr} = \tau_{rz} = \mu \left[\frac{\partial v_z}{\partial r} + \frac{\partial v_r}{\partial z} \right] \tag{F}$$

$$(\nabla \cdot \mathbf{V}) = \frac{1}{r} \frac{\partial}{\partial r} (r v_r) + \frac{1}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{\partial v_z}{\partial z} \tag{G}$$

<center>Spherical coordinates (r, θ, ϕ)</center>

$$\tau_{rr} = \mu \left[2 \frac{\partial v_r}{\partial r} - \frac{2}{3} (\nabla \cdot \mathbf{V}) \right] \tag{A}$$

$$\tau_{\theta\theta} = \mu \left[2 \left(\frac{1}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{v_r}{r} \right) - \frac{2}{3} (\nabla \cdot \mathbf{V}) \right] \tag{B}$$

$$\tau_{\phi\phi} = \mu \left[2 \left(\frac{1}{r \sin \theta} \frac{\partial v_\phi}{\partial \phi} + \frac{v_r}{r} + \frac{v_\theta \cot \theta}{r} \right) - \frac{2}{3} (\nabla \cdot \mathbf{V}) \right] \tag{C}$$

$$\tau_{r\theta} = \tau_{\theta r} = \mu \left[r \frac{\partial}{\partial r} \left(\frac{v_\theta}{r} \right) + \frac{1}{r} \frac{\partial v_r}{\partial \theta} \right] \tag{D}$$

$$\tau_{\theta\phi} = \tau_{\phi\theta} = \mu \left[\frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \left(\frac{v_\phi}{\sin \theta} \right) + \frac{1}{r \sin \theta} \frac{\partial v_\theta}{\partial \phi} \right] \tag{E}$$

$$\tau_{\phi r} = \tau_{r\phi} = \mu \left[\frac{1}{r \sin \theta} \frac{\partial v_r}{\partial \phi} + r \frac{\partial}{\partial r} \left(\frac{v_\phi}{r} \right) \right] \tag{F}$$

$$(\nabla \cdot \mathbf{V}) = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 v_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (v_\theta \sin \theta) + \frac{1}{r \sin \theta} \frac{\partial v_\phi}{\partial \phi} \tag{G}$$

*It should be noted that the sign convention adopted here for components of the stress tensor is consistent with that found in many fluid mechanics and heat-transfer books; however, it is opposite to that found in some books on transport phenomena, e.g., Refs. 2, 3, and 5.

As mentioned before, there is a subset of flow problems, called *natural convection,* where the flow pattern is due to buoyant forces caused by temperature differences. Such buoyant forces are proportional to the coefficient of thermal expansion β, defined as

$$\beta = -\frac{1}{\rho}\left(\frac{\partial\rho}{\partial T}\right)_P \tag{10.12}$$

TABLE 10.5 Equation of Motion in Terms of Velocity Gradients for a Newtonian Fluid with Constant ρ and μ, Eq. (10.10)

Rectangular coordinates (x, y, z)

x direction

$$\rho\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z}\right) = -\frac{\partial P}{\partial x} + \mu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) + \rho g_x \tag{A}$$

y direction

$$\rho\left(\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z}\right) = -\frac{\partial P}{\partial y} + \mu\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right) + \rho g_y \tag{B}$$

z direction

$$\rho\left(\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z}\right) = -\frac{\partial P}{\partial z} + \mu\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 ww}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right) + \rho g_z \tag{C}$$

Cylindrical coordinates (r, θ, z)

r direction

$$\rho\left(\frac{\partial v_r}{\partial t} + v_r\frac{\partial v_r}{\partial r} + \frac{v_\theta}{r}\frac{\partial v_r}{\partial\theta} - \frac{v_\theta^2}{r} + v_z\frac{\partial v_r}{\partial z}\right)$$

$$= -\frac{\partial P}{\partial r} + \mu\left[\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial}{\partial r}(rv_r)\right) + \frac{1}{r^2}\frac{\partial^2 v_r}{\partial\theta^2} - \frac{2}{r^2}\frac{\partial v_\theta}{\partial\theta} + \frac{\partial^2 v_r}{\partial z^2}\right] + \rho g_r \tag{A}$$

θ direction

$$\rho\left(\frac{\partial v_\theta}{\partial t} + v_r\frac{\partial v_\theta}{\partial r} + \frac{v_\theta}{r}\frac{\partial v_\theta}{\partial\theta} + \frac{v_r v_\theta}{r} + v_z\frac{\partial v_\theta}{\partial z}\right)$$

$$= -\frac{1}{r}\frac{\partial P}{\partial\theta} + \mu\left[\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial}{\partial r}(rv_\theta)\right) + \frac{1}{r^2}\frac{\partial^2 v_\theta}{\partial\theta^2} + \frac{2}{r^2}\frac{\partial v_r}{\partial\theta} + \frac{\partial^2 v_\theta}{\partial z^2}\right] + \rho g_\theta \tag{B}$$

z direction

$$\rho\left(\frac{\partial v_z}{\partial t} + v_r\frac{\partial v_z}{\partial r} + \frac{v_\theta}{r}\frac{\partial v_z}{\partial\theta} + v_z\frac{\partial v_z}{\partial z}\right) = -\frac{\partial P}{\partial z} + \mu\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_z}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 v_z}{\partial\theta^2} + \frac{\partial^2 v_z}{\partial z^2}\right] + \rho g_z$$

$$\tag{C}$$

TABLE 10.5 Equation of Motion in Terms of Velocity Gradients for a Newtonian Fluid with Constant ρ and μ, Eq. (10.10) (*Continued*)

Spherical coordinates (r, θ, ϕ)†

r direction

$$\rho \left(\frac{\partial v_r}{\partial t} + v_r \frac{\partial v_r}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_r}{\partial \theta} + \frac{v_\phi}{r \sin \theta} \frac{\partial v_r}{\partial \phi} - \frac{v_\theta^2 + v_\phi^2}{r} \right)$$

$$= -\frac{\partial P}{\partial r} + \mu \left(\nabla^2 v_r - \frac{2v_r}{r^2} - \frac{2}{r^2} \frac{\partial v_\theta}{\partial \theta} - \frac{2v_\theta \cot \theta}{r^2} - \frac{2}{r^2 \sin \theta} \frac{\partial v_\phi}{\partial \phi} \right) + \rho g_r \quad (A)$$

θ direction

$$\rho \left(\frac{\partial v_\theta}{\partial t} + v_r \frac{\partial v_\theta}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{v_\phi}{r \sin \theta} \frac{\partial v_\theta}{\partial \phi} + \frac{v_r v_\theta}{r} - \frac{v_\phi^2 \cot \theta}{r} \right)$$

$$= -\frac{1}{r} \frac{\partial P}{\partial \theta} + \mu \left(\nabla^2 v_\theta + \frac{2}{r^2} \frac{\partial v_r}{\partial \theta} - \frac{v_\theta}{r^2 \sin^2 \theta} - \frac{2 \cos \theta}{r^2 \sin^2 \theta} \frac{\partial v_\phi}{\partial \phi} \right) + \rho g_\theta \quad (B)$$

ϕ direction

$$\rho \left(\frac{\partial v_\phi}{\partial t} + v_r \frac{\partial v_\phi}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_\phi}{\partial \theta} + \frac{v_\phi}{r \sin \theta} \frac{\partial v_\phi}{\partial \phi} + \frac{v_\phi v_r}{r} + \frac{v_\theta v_\phi}{r} \cot \theta \right)$$

$$= -\frac{1}{r \sin \theta} \frac{\partial P}{\partial \phi} + \mu \left(\nabla^2 v_\phi - \frac{v_\phi}{r^2 \sin^2 \theta} + \frac{2}{r^2 \sin \theta} \frac{\partial v_r}{\partial \phi} + \frac{2 \cos \theta}{r^2 \sin^2 \theta} \frac{\partial v_\theta}{\partial \phi} \right) + \rho g_\phi \quad (C)$$

† For spherical coordinates the Laplacian is

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \left(\frac{\partial^2}{\partial \phi^2} \right)$$

where T is absolute temperature. Using an approximation that applies to low fluid velocities and small temperature variations, it can be shown[2,3] that

$$\nabla P - \rho \mathbf{g} = \rho \beta \mathbf{g} (T - T_\infty) \qquad (10.13)$$

Then Eq. (10.8) becomes

$$\rho \frac{D\mathbf{V}}{Dt} = \nabla \cdot \tau - \underbrace{\rho \beta \mathbf{g} (T - T_\infty)}_{\substack{\text{buoyant force on} \\ \text{element per unit} \\ \text{volume}}} \qquad (10.14)$$

The preceding equation of motion is used for setting up problems in natural convection when the ambient temperature T_∞ may be defined.

3. The Energy Equation.   For a stationary volume element through which a pure fluid is flowing, the energy equation reads

$$\frac{\partial}{\partial t} \rho(\mathbf{u} + \tfrac{1}{2}V^2) = -\nabla \cdot \rho \mathbf{V}(\mathbf{u} + \tfrac{1}{2}V^2) - \nabla \cdot \mathbf{q}'' + \rho(\mathbf{V} \cdot \mathbf{g})$$

rate of gain of energy per unit volume / rate of energy input per unit volume by convection / rate of energy input per unit volume by conduction / rate of work done on fluid per unit volume by gravitational forces

$$-\nabla \cdot P\mathbf{V} + \nabla \cdot (\tau \cdot \mathbf{V}) + q''' \qquad (10.15)$$

rate of work done on fluid per unit volume by pressure forces / rate of work done on fluid per unit volume by viscous forces / rate of heat generation per unit volume ("source term")

The left side of the preceding equation, which represents the rate of accumulation of internal and kinetic energy, does not include the potential energy of the fluid, since this form of energy is included in the work term on the right side. Equation (10.15) may be rearranged, with the aid of the equations of continuity and motion, to give[2,6]

$$\rho \frac{D\mathbf{u}}{Dt} = -\nabla \cdot \mathbf{q}'' - P(\nabla \cdot \mathbf{V}) + \nabla \mathbf{V}:\tau + q''' \qquad (10.16)$$

A summary of $\nabla \mathbf{V}:\tau$ in different coordinate systems is given in Table 10.6. For a newtonian fluid,

TABLE 10.6 Summary of Dissipation Term $\nabla \mathbf{V}:\tau$ in Different Coordinate Systems

Rectangular coordinates (x, y, z):

$$\nabla \mathbf{V}:\tau = \tau_{xx}\left(\frac{\partial u}{\partial x}\right) + \tau_{yy}\left(\frac{\partial v}{\partial y}\right) + \tau_{zz}\left(\frac{\partial w}{\partial z}\right) + \tau_{xy}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)$$

$$+ \tau_{yz}\left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right) + \tau_{zx}\left(\frac{\partial w}{\partial z} + \frac{\partial u}{\partial z}\right) \qquad (A)$$

Cylindrical coordinates (r, θ, z):

$$\nabla \mathbf{V}:\tau = \tau_{rr}\left(\frac{\partial v_r}{\partial r}\right) + \tau_{\theta\theta}\left(\frac{1}{r}\frac{\partial v_\theta}{\partial \theta} + \frac{v_r}{r}\right) + \tau_{zz}\left(\frac{\partial v_z}{\partial z}\right) + \tau_{r\theta}\left[r\frac{\partial}{\partial r}\left(\frac{v_\theta}{r}\right) + \frac{1}{r}\frac{\partial v_r}{\partial \theta}\right]$$

$$+ \tau_{\theta z}\left(\frac{1}{r}\frac{\partial v_z}{\partial \theta} + \frac{\partial v_\theta}{\partial z}\right) + \tau_{rz}\left(\frac{\partial v_z}{\partial r} + \frac{\partial v_r}{\partial z}\right) \qquad (B)$$

Spherical coordinates (r, θ, ϕ):

$$\nabla \mathbf{V}:\tau = \tau_{rr}\left(\frac{\partial v_r}{\partial r}\right) + \tau_{\theta\theta}\left(\frac{1}{r}\frac{\partial v_\theta}{\partial \theta} + \frac{v_r}{r}\right) + \tau_{\phi\phi}\left(\frac{1}{r\sin\theta}\frac{\partial v_\phi}{\partial \phi} + \frac{v_r}{r} + \frac{v_\theta\cot\theta}{r}\right)$$

$$+ \tau_{r\theta}\left(\frac{\partial v_\theta}{\partial r} + \frac{1}{r}\frac{\partial v_r}{\partial \theta} - \frac{v_\theta}{r}\right) + \tau_{r\theta}\left(\frac{\partial v_\phi}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial v_r}{\partial \phi} - \frac{v_\phi}{r}\right)$$

$$+ \tau_{\theta\phi}\left(\frac{1}{r}\frac{\partial v_\phi}{\partial \theta} + \frac{1}{r\sin\theta}\frac{\partial v_\theta}{\partial \phi} - \frac{v_\phi\cot\theta}{r}\right) \qquad (C)$$

$$\nabla \mathbf{V} : \tau = \mu \Phi \tag{10.17}$$

and values of the dissipation function Φ in different coordinate systems are given in Table 10.7. Components of the heat flux vector $\mathbf{q}'' = -k\nabla T$ are given in Table 10.8 for different coordinate systems.

Often it is more convenient to work with enthalpy rather than internal energy. Using the definition of enthalpy, $i = \mathbf{u} + P/\rho$, and the mass conservation equation [Eq. (10.5)], then Eq. (10.16) can be rearranged to give

TABLE 10.7 The Viscous Dissipation Function Φ

Rectangular coordinates (x, y, z):

$$\Phi = 2\left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial w}{\partial z}\right)^2\right] + \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z}\right)^2$$

$$+ \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)^2 - \frac{2}{3}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right)^2 \tag{A}$$

Cylindrical coordinates (r, θ, z):

$$\Phi = 2\left[\left(\frac{\partial v_r}{\partial r}\right)^2 + \left(\frac{1}{r}\frac{\partial v_\theta}{\partial \theta} + \frac{v_r}{r}\right)^2 + \left(\frac{\partial v_z}{\partial z}\right)^2\right] + \left[r\frac{\partial}{\partial r}\left(\frac{v_\theta}{r}\right) + \frac{1}{r}\frac{\partial v_r}{\partial \theta}\right]^2$$

$$+ \left[\frac{1}{r}\frac{\partial v_z}{\partial \theta} + \frac{\partial v_\theta}{\partial z}\right]^2 + \left(\frac{\partial v_r}{\partial z} + \frac{\partial v_z}{\partial r}\right)^2 - \frac{2}{3}\left[\frac{1}{r}\frac{\partial}{\partial r}(rv_r) + \frac{1}{r}\frac{\partial v_\theta}{\partial \theta} + \frac{\partial v_z}{\partial z}\right]^2 \tag{B}$$

Spherical coordinates (r, θ, ϕ):

$$\Phi = 2\left[\left(\frac{\partial v_r}{\partial r}\right)^2 + \left(\frac{1}{r}\frac{\partial v_\theta}{\partial \theta} + \frac{v_r}{r}\right)^2 + \left(\frac{1}{r\sin\theta}\frac{\partial v_\phi}{\partial \phi} + \frac{v_r}{r} + \frac{v_\theta\cot\theta}{r}\right)^2\right]$$

$$+ \left[r\frac{\partial}{\partial r}\left(\frac{v_\theta}{r}\right) + \frac{1}{r}\frac{\partial v_r}{\partial \theta}\right]^2 + \left[\frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\left(\frac{v_\phi}{\sin\theta}\right) + \frac{1}{r\sin\theta}\frac{\partial v_\theta}{\partial \phi}\right]^2$$

$$+ \left[\frac{1}{r\sin\theta}\frac{\partial v_r}{\partial \phi} + r\frac{\partial}{\partial r}\left(\frac{v_\phi}{r}\right)\right]^2 - \frac{2}{3}\left[\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 v_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}(v_\theta\sin\theta) + \frac{1}{r\sin\theta}\frac{\partial v_\phi}{\partial \phi}\right]^2 \tag{C}$$

TABLE 10.8 Scalar Components of the Heat Flux Vector $\mathbf{q}''$

Rectangular (x, y, z)		Cylindrical (r, θ, z)		Spherical (r, θ, ϕ)	
$q_x'' = -k\dfrac{\partial T}{\partial x}$	(A)	$q_r'' = -k\dfrac{\partial T}{\partial r}$	(D)	$q_r'' = -k\dfrac{\partial T}{\partial r}$	(G)
$q_y'' = -k\dfrac{\partial T}{\partial y}$	(B)	$q_\theta'' = -k\dfrac{1}{r}\dfrac{\partial T}{\partial \theta}$	(E)	$q_\theta'' = -k\dfrac{1}{r}\dfrac{\partial T}{\partial \theta}$	(H)
$q_z'' = -k\dfrac{\partial T}{\partial z}$	(C)	$q_z'' = -k\dfrac{\partial T}{\partial z}$	(F)	$q_\phi'' = -k\dfrac{1}{r\sin\theta}\dfrac{\partial T}{\partial \phi}$	(I)

$$\rho \frac{Di}{Dt} = \nabla \cdot k\nabla T + \frac{DP}{Dt} + \mu\Phi + q''' \tag{10.18}$$

For most engineering applications, it is convenient to have the equation of thermal energy in terms of the fluid temperature and heat capacity rather than the internal energy or enthalpy. In general, for pure substances,[3]

$$\frac{Di}{Dt} = \left(\frac{\partial i}{\partial P}\right)_T \frac{DP}{Dt} + \left(\frac{\partial i}{\partial T}\right)_P \frac{DT}{Dt} = \frac{1}{\rho}(1 - \beta T)\frac{DP}{Dt} + c_p \frac{DT}{Dt} \tag{10.19}$$

where β is defined by Eq. (10.12). Substituting this into Eq. (10.18), we have the following general relation:

$$\rho c_p \frac{DT}{Dt} = \nabla \cdot k\nabla T + T\beta \frac{DP}{Dt} + \mu\Phi + q''' \tag{10.20}$$

For an ideal gas, $\beta = 1/T$, and then

$$\rho c_p \frac{DT}{Dt} = \nabla \cdot k\nabla T + \frac{DP}{Dt} + \mu\Phi + q''' \tag{10.21}$$

Note that c_p need not be constant.

We could have obtained Eq. (10.21) directly from Eq. (10.18) by noting that for an ideal gas, $di = c_p \, dT$, where c_p is constant, and thus

$$\frac{Di}{Dt} = c_p \frac{DT}{Dt}$$

For an incompressible fluid with specific heat $c = c_p = c_v$, we go back to Eq. (10.16) ($d\mathbf{u} = c \, dT$) to obtain

$$\rho c \frac{DT}{Dt} = \nabla \cdot k\nabla T + \mu\Phi + q''' \tag{10.22}$$

Equations (10.16), (10.18), and (10.20) can be easily written in terms of energy (heat) and momentum fluxes using relations for fluxes given in Tables 10.4, 10.6, and 10.8. The energy equation given by Eq. (10.22) (with $q''' = 0$ for simplicity) is given in Table 10.9 in different coordinate systems.

For *solids,* the density may usually be considered constant, and we may set $\mathbf{V} = 0$, and Eq. (10.22) reduces to

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot k\nabla T + q''' \tag{10.23}$$

which is the starting point for most problems in heat conduction.

The Energy Equation for a Mixture.   The energy equations in the preceding section are applicable for a pure fluid. A thermal energy equation valid for a mixture of chemical species is required for situations involving simultaneous heat and mass transfer. For a pure fluid, conduction is the only diffusive mechanism of heat flow; hence Fourier's law was used, which resulted in the term $\nabla \cdot k\nabla T$. More generally, this term may be written $-\nabla\mathbf{q}''$, where $\mathbf{q}''$ is the diffusive heat flux, i.e., the heat flux relative to the mass average velocity. More specifically, for a mixture, $\mathbf{q}''$ is now made from three contributions: (1) ordinary conduction, described by Fourier's

TABLE 10.9 The Energy Equation (for Newtonian Fluids of Constant ρ and k)*

Rectangular coordinates (x, y, z):

$$\rho c_p \left(\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + w \frac{\partial T}{\partial z} \right) = k \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right) + 2\mu \left\{ \left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial y} \right)^2 \right.$$

$$\left. + \left(\frac{\partial w}{\partial z} \right)^2 \right\} + \mu \left\{ \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right)^2 \right\} \quad (A)$$

Cylindrical coordinates (r, θ, z):

$$\rho c_p \left(\frac{\partial T}{\partial t} + v_r \frac{\partial T}{\partial r} + \frac{v_\theta}{r} \frac{\partial T}{\partial \theta} + v_z \frac{\partial T}{\partial z} \right) = k \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial T}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 T}{\partial \theta^2} + \frac{\partial^2 T}{\partial z^2} \right]$$

$$+ 2\mu \left\{ \left(\frac{\partial v_r}{\partial r} \right)^2 + \left[\frac{1}{r} \left(\frac{\partial v_\theta}{\partial \theta} + v_r \right) \right]^2 + \left(\frac{\partial v_z}{\partial z} \right)^2 \right\} + \mu \left\{ \left(\frac{\partial v_\theta}{\partial z} + \frac{1}{r} \frac{\partial v_z}{\partial \theta} \right)^2 \right.$$

$$\left. + \left(\frac{\partial v_z}{\partial r} + \frac{\partial v_r}{\partial z} \right)^2 + \left[\frac{1}{r} \frac{\partial v_r}{\partial \theta} + r \frac{\partial}{\partial r} \left(\frac{v_\theta}{r} \right) \right]^2 \right\} \quad (B)$$

Spherical coordinates (r, θ, ϕ):

$$\rho c_p \left(\frac{\partial T}{\partial t} + v_r \frac{\partial T}{\partial r} + \frac{v_\theta}{r} \frac{\partial T}{\partial \theta} + \frac{v_\phi}{r \sin \theta} \frac{\partial T}{\partial \phi} \right) = k \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial T}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial T}{\partial \theta} \right) \right.$$

$$\left. + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 T}{\partial \phi^2} \right] + 2\mu \left\{ \left(\frac{\partial v_r}{\partial r} \right)^2 + \left(\frac{1}{r} \frac{\partial v_\theta}{\partial \theta} + \frac{v_r}{r} \right)^2 + \left(\frac{1}{r \sin \theta} \frac{\partial v_\phi}{\partial \phi} + \frac{v_r}{r} + \frac{v_\theta \cot \theta}{r} \right)^2 \right\}$$

$$+ \mu \left\{ \left[r \frac{\partial}{\partial r} \left(\frac{v_\theta}{r} \right) + \frac{1}{r} \frac{\partial v_r}{\partial \theta} \right]^2 + \left[\frac{1}{r \sin \theta} \frac{\partial v_r}{\partial \phi} + r \frac{\partial}{\partial r} \left(\frac{v_\phi}{r} \right) \right]^2 \right.$$

$$\left. + \left[\frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \left(\frac{v_\phi}{\sin \theta} \right) + \frac{1}{r \sin \theta} \frac{\partial v_\theta}{\partial \phi} \right]^2 \right\} \quad (C)$$

*The terms contained in braces { } are associated with viscous dissipation and may usually be neglected except in systems with large velocity gradients.

law, $-k\nabla T$, where k is the *mixture* thermal conductivity; (2) the contribution due to interdiffusion of species, given by $\Sigma_i \mathbf{j}_i i_i$; and (3) diffusional conduction (also called the *diffusion-thermo effect* or *Dufour effect*[1,7]). The third contribution is of the second order and is usually negligible:

$$\mathbf{q}'' = -k \nabla T + \sum_i \mathbf{j}_i i_i \quad (10.24)$$

Here $\mathbf{j}_i$ is a diffusive mass flux of species i, with units of mass/(area $\times$ time), as mentioned before. Substituting Eq. (10.24) in, for example, Eq. (10.18), we obtain the energy equation for a mixture:

$$\rho \frac{Di}{Dt} = \frac{DP}{Dt} + \nabla \cdot k\nabla T - \nabla \cdot \left(\sum_i \mathbf{j}_i i_i \right) + \mu \Phi + q''' \quad (10.25)$$

For a nonreacting mixture the term $\nabla \cdot (\Sigma i \mathbf{j}_i i_i)$ is often of minor importance. But

when endothermic or exothermic reactions occur, the term can play a dominant role. For reacting mixtures, the species enthalpies

$$i_i = i_i^0 + \int_{T^0}^{T} c_{pi} \, dT$$

must be written with a consistent set of heats of formation i_i^0 at T^0.[8]

4. The Conservation Equation for Species. For a stationary control volume, the conservation equation for species is

$$\frac{\partial C_i}{\partial t} = -\nabla \cdot (C_i \mathbf{V}) - \nabla \cdot \mathbf{j}_i + r_i''' \tag{10.26}$$

| rate of storage of species i per unit volume | net rate of convection of species i per unit volume | net rate of diffusion of species i per unit volume | production rate of species i per unit volume |

Using the mass conservation equation, the preceding equation can be rearranged to obtain

$$\rho \frac{Dm_i}{Dt} = -\nabla \cdot \mathbf{j}_i + r_i''' \tag{10.27}$$

where m_i is mass fraction of species i, i.e., where $m_i = C_i/\rho$, where ρ is the density of the mixture, $\Sigma_i C_i = \rho$, and C_i is a partial density of species i (i.e., a mass concentration of species i).

The conservation equation of species also can be written in terms of mole concentration and mole fractions, as shown in Refs. 2, 3, and 7. The mole concentration of species i is $c_i = C_i/M_i$, where M_i is the molecular weight of the species. The mole fraction of species i is defined as $x_i = c_i/c$, where $c = \Sigma_i c_i$. As is obvious, $\Sigma_i m_i = 1$ and $\Sigma_i x_i = 1$.

Equations (10.26) and (10.27) written in different coordinate systems are given in Ref. 2.

5. Use of Conservation Equations to Set Up Problems. For a problem involving fluid flow and simultaneous heat and mass transfer, equations of continuity, momentum, energy, and chemical species, Eqs. (10.5), (10.8), (10.18), and (10.27), are a formidable set of partial differential equations. There are four *independent variables:* three space coordinates (say, x, y, z) and a time coordinate t.

If we consider a pure fluid, then there are five equations: the continuity equation, three momentum equations, and the energy equation. The accompanying five *dependent variables* are pressure, three components of velocity, and temperature. Also, a thermodynamic equation of state serves to relate density to the pressure, temperature, and composition. (Notice that for natural-convection flows, the momentum and energy equations are coupled.)

For a mixture of n chemical species, there are n species conservation equations, but one is redundant, since the sum of mass fractions is equal to unity.

A complete mathematical statement of a problem requires specification of boundary and initial conditions. Boundary conditions are based on a physical statement or principle (for example, for viscous flow, the component of velocity parallel to a stationary surface is zero at the wall; for an insulated wall, the derivative of temperature normal to the wall is zero; etc.).

A general solution, even by numerical methods, of the full equations in the four independent variables is difficult to obtain. Fortunately, however, many problems of engineering interest are adequately described by simplified forms of the full conservation equations, and these forms can often be solved easily. The governing equations for simplified problems are obtained by deleting superfluous terms in the full conservation equations. This applies directly to laminar flows only. In the case of turbulent flows, some caution must be exercised. For example, on an average basis a flow may be two-dimensional and steady, but if it is unstable and as a result turbulent, fluctuations in the three components of velocity may be occurring with respect to time and the three spatial coordinates. Then the remarks about dropping terms apply only to the time-averaged equations.[7,8]

When simplifying the conservation equation given in a full form, we have to rely on physical intuition or on experimental evidence to judge which terms are negligibly small. Typical resulting classes of simplified problems are

Constant transport properties

Constant density

Timewise steady flow (or quasi-steady flow)

Two-dimensional flow

One-dimensional flow

Fully developed flow (no dependence on the streamwise coordinate)

Stagnant fluid or rigid body

Terms also may be shown to be negligibly small by order-of-magnitude estimates.[7,8] Some classes of flow that result are

Creeping flows: Inertia terms are negligible.

Forced flows: Gravity forces are negligible.

Natural convection: Gravity forces predominate.

Low-speed gas flows: Viscous dissipation and compressibility terms are negligible.

Boundary-layer flows: Streamwise diffusion terms are negligible.

DIMENSIONLESS GROUPS AND SIMILARITY IN FLUID MECHANICS AND HEAT TRANSFER

Modern engineering practice in the fields of fluid mechanics and heat transfer is based on a combination of theoretical analysis and experimental data. Often the engineer is faced with the necessity of obtaining practical results in situations where, for various reasons, physical phenomena cannot be described mathematically or the differential equations describing the problem are too difficult to solve. An experimental program must be considered in such cases. However, in carrying the experimental program, the engineer should know how to relate the experimental data (i.e., data obtained on the model under consideration) to the actual, usually larger, system (prototype). A determination of the relevant dimensionless parameters (groups) provides a powerful tool for that purpose.

The generation of such dimensionless groups in heat transfer and fluid mechanics (known generally as *dimensional analysis)* is basically done (1) by using differential equations and their boundary conditions (this method is sometimes called a *differential similarity)* and (2) by applying the dimensional analysis in the form of the Buckingham pi theorem.

The first method (differential similarity) is used when the governing equations and their boundary conditions describing the problem are known. The equations are first made dimensionless. For demonstration purposes, let us consider the relatively simple problem of a binary mixture with constant properties and density flowing at low speed, where body forces, heat source term, and chemical reactions are neglected. The conservation equations are, from Eqs. (10.6), (10.10), (10.16), and (10.27),

Mass
$$\nabla \cdot \mathbf{V} = 0 \tag{10.28}$$

Momentum
$$\rho \frac{D\mathbf{V}}{Dt} = -\nabla P + \mu\nabla^2\mathbf{V} \tag{10.29}$$

Thermal energy
$$\rho c \frac{DT}{Dt} = k\nabla^2 T + \mu\Phi \tag{10.30}$$

Species
$$\frac{Dm_1}{Dt} = \mathbf{D}\nabla^2 m_1 \tag{10.31}$$

Using L and V as characteristic length and velocity, respectively, we define the dimensionless variables

$$x* = \frac{x}{L} \qquad y* = \frac{y}{L} \qquad z* = \frac{z}{L} \tag{10.32}$$

$$\mathbf{V}* = \frac{\mathbf{V}}{V} \tag{10.33}$$

$$t* = \frac{t}{L/V} \tag{10.34}$$

$$P* = \frac{P}{\rho V^2} \tag{10.35}$$

and also

$$T* = \frac{T - T_w}{T_\infty - T_w} \tag{10.36}$$

$$m* = \frac{m_1 - m_{1,w}}{m_{1,\infty} - m_{1,w}} \tag{10.37}$$

where the subscript ∞ refers to the external free-stream condition or some average condition and the subscript w refers to conditions adjacent to a bounding surface across which transfer of heat and mass occurs. If we introduce the dimensionless quantities, Eqs. (10.32) to (10.37), into Eqs. (10.28) to (10.31), we obtain, respectively,

$$\nabla^* \cdot \mathbf{V}^* = 0 \tag{10.38}$$

$$\frac{D\mathbf{V}^*}{Dt^*} = -\nabla^*P^* + \frac{1}{Re}\nabla^{*2}\mathbf{V}^* \tag{10.39}$$

$$\frac{DT^*}{Dt^*} = \frac{1}{Re\ Pr}\nabla^{*2}T^* + \frac{2\ Ec}{Re}\Phi^* \tag{10.40}$$

$$\frac{Dm^*}{Dt^*} = \frac{1}{Re\ Sc}\nabla^{*2}m^* \tag{10.41}$$

Obviously, the solutions of Eqs. (10.38) to (10.41) depend on the coefficients that appear in these equations. Solutions of Eqs. (10.38) to (10.41) are equally applicable to the model and prototype (where the model and prototype are geometrically similar systems of different linear dimensions in streams of different velocities, temperatures, and concentration), if the coefficients in these equations are the same for both model and prototype. These coefficients, Pr, Re, Sc, and Ec (called *dimensionless parameters* or *similarity parameters*), are defined in Table 10.10.

Focusing attention now on heat transfer, from Eq. (9.22), using the dimensionless quantities, the heat-transfer coefficient is given as

$$h = \frac{k\partial T^*}{L\ \partial y^*}\bigg|_{y^*=0} \tag{10.42}$$

or in dimensionless form,

$$\frac{hL}{k} = \frac{\partial T^*}{\partial y^*}\bigg|_{y^*=0} = Nu \tag{10.43}$$

where the dimensionless group Nu is known as the *Nusselt number*. Since Nu is the dimensionless temperature gradient at the surface, according to Eq. (10.40) it must therefore depend on the dimensionless groups that appear in this equation; hence

$$Nu = f_1(Re, Pr, Ec) \tag{10.44}$$

For processes where viscous dissipation and compressibility are negligible, which is the case in many industrial applications, we have

$$Nu = f_2(Re, Pr) \quad \text{(forced convection)} \tag{10.45}$$

In the case of buoyancy-induced flow, Eq. (10.29) should be replaced with the simplified version[9] of Eq. (10.14), and following a similar procedure, we obtain

$$Nu = f_3(Gr, Pr) \quad \text{(natural convection)} \tag{10.46}$$

where Gr is the Grashof number, defined in Table 10.10. Also, using the relation of Eq. (9.26), and dimensionless quantities,

$$h_{\mathbf{D}} = \frac{\mathbf{D}}{\mathbf{L}}\frac{\partial m^*}{\partial y^*}\bigg|_{y^*=0} \tag{10.47}$$

TABLE 10.10 Summary of the Chief Dimensionless Groups*

Group	Symbol	Definition	Physical significance (interpretation)	Main area of use
Biot number	Bi	$\dfrac{hL}{k_s}$	Ratio of internal thermal resistance of solid to fluid thermal resistance	Heat transfer between fluid and solid
Biot number† (mass transfer)	Bi_D	$\dfrac{h_D L}{D}$	Ratio of the internal species transfer resistance to the boundary-layer species transfer resistance	Mass transfer between fluid and solid
Coefficient of friction (skin friction coefficient)	c_f	$\dfrac{\tau_w}{\rho V^2/2}$	Dimensionless surface shear stress	Flow resistance
Eckert number	Ec	$\dfrac{V_\infty^2}{c_p(T_w - T_\infty)}$	Kinetic energy of the flow relative to the boundary-layer enthalpy difference	Forced convection (compressible flow)
Euler number	Eu	$\dfrac{\Delta P}{\rho V^2}$	Ratio of friction to velocity head	Fluid friction
Fourier number	Fo	$\dfrac{\alpha t}{L^2}$	Ratio of the heat conduction rate to the rate of thermal energy storage in a solid	Unsteady-state heat transfer
Fourier number (mass transfer)	Fo_D	$\dfrac{Dt}{L^2}$	Ratio of the species diffusion rate to the rate of species storage	Unsteady-state mass transfer
Froude number	Fr	$\dfrac{V^2}{gL}$	Ratio of inertial to gravitational force	Wave and surface behavior (mixed natural and forced convection)
Graetz number	Gz	$Re\ Pr\,\dfrac{D}{L} = \dfrac{\rho c_p VD^2}{kL}$	Ratio of the fluid stream thermal capacity to convective heat transfer	Forced convection
Grashof number	Gr	$\dfrac{g\beta\,\Delta TL^3}{\nu^2}$	Ratio of buoyancy to viscous forces	Natural convection
Colburn j factor (heat transfer)	j_H	$St\,Pr^{2/3}$	Dimensionless heat transfer coefficient	Forced convection (heat, mass, and momentum transfer analogy)
Colburn j factor (mass transfer)	j_D	$St_D\,Sc^{2/3}$	Dimensionless mass transfer coefficient	Forced convection (heat, mass, and momentum transfer analogy)

Name	Symbol	Formula	Interpretation	Use
Jakob number	Ja	$\dfrac{\rho c_{pl}(T_w - T_{sat})}{\rho_g i_{lg}}$	Ratio of sensible heat absorbed by the liquid to the latent heat absorbed	Boiling
Knudsen number	Kn	$\dfrac{\lambda}{L}$	Ratio of molecular mean free path to characteristic dimension	Low-pressure (low-density) gas flow
Lewis number	Le	$\dfrac{\alpha}{\mathbf{D}} = \dfrac{Sc}{Pr}$	Ratio of molecular thermal and mass diffusivities	Combined heat and mass transfer
Mach number	Ma	$\dfrac{V}{a}$	Ratio of the velocity of flow to the velocity of sound	Compressible flow
Nusselt number	Nu	$\dfrac{hL}{k}$	Basic dimensionless convective heat transfer coefficient (ratio of convection heat transfer to conduction in a fluid slab of thickness L)	Convective heat transfer
Péclet number	Pe	$Re\,Pr = \dfrac{\rho c_p VL}{k}$	Dimensionless independent heat transfer parameter (ratio of heat transfer by convection to conduction)	Forced convection
Péclet number (mass transfer)	Pe$_{\mathbf{D}}$	$Re\,Sc = \dfrac{VL}{\mathbf{D}}$	Dimensionless independent mass transfer coefficient (ratio of bulk mass transfer to diffusive mass transfer)	Mass transfer
Prandtl number	Pr	$\dfrac{\mu c_p}{k} = \dfrac{\nu}{\alpha}$	Ratio of molecular momentum and thermal diffusivities	Forced and natural convection
Rayleigh number	Ra	$Gr\,Pr = \dfrac{\rho g \beta\, \Delta T L^3}{\mu \alpha}$	Modified Grashof number (see interpretation for Gr and Pr)	Natural convection
Reynolds number	Re	$\dfrac{\rho VL}{\mu}$	Ratio of inertia to viscous forces	Forced convection; dynamic similarity
Schmidt number	Sc	$\dfrac{\nu}{\mathbf{D}}$	Ratio of molecular momentum and mass diffusivities	Mass transfer

10.19

TABLE 10.10 Summary of the Chief Dimensionless Groups* (*Continued*)

Group	Symbol	Definition	Physical significance (interpretation)	Main area of use
Sherwood number	Sh	$\dfrac{h_\mathbf{D}L}{\mathbf{D}}$	Ratio of convection mass transfer to diffusion in a slab of thickness L	Convective mass transfer
Strouhal number‡	Sr	$\dfrac{Lf}{V}$	Ratio of the velocity of vibration Lf to the velocity of the fluid	Flow past tube (shedding of eddies)
Stanton number	St	$\dfrac{\text{Nu}}{\text{Re Pr}} = \dfrac{h}{\rho c_p V}$	Dimensionless heat transfer coefficient (ratio of heat transfer at the surface to that transported by fluid by its thermal capacity)	Forced convection
Stanton number (mass transfer)	St$_\mathbf{D}$	$\dfrac{\text{Sh}}{\text{Re Sc}} = \dfrac{h_\mathbf{D}}{V}$	Dimensionless mass transfer coefficient	Convective mass transfer
Weber number	We	$\dfrac{\rho V^2 L}{\sigma}$	Ratio of inertia force to surface tension force	Droplet breakup; thin-film flow

* In these dimensionless groups, L designates characteristic dimension (e.g., tube diameter, hydraulic diameter, length of the tube or plate, slab thickness, radius of a cylinder or sphere, droplet diameter, thin-film thickness, etc.). Physical properties are usually evaluated at mean temperature unless otherwise specified.

† *Note:* $\mathbf{D} = \mathbf{D}_{12}$ ($\mathbf{D}_{12}$ is also a commonly used symbol for binary diffusion coefficient; $\mathbf{D}_{ij}$ is the h multicomponent diffusion coefficient). When species 1 is in very small concentration, the symbol $\mathbf{D}_{1m}$ is occasionally used,[7] representing an effective binary diffusion coefficient for species 1 diffusing through the mixture.

‡ In some engineering texts, the symbol St is also used for this group.

or

$$h_{\mathbf{D}}\frac{L}{\mathbf{D}} = \left.\frac{\partial m^*}{\partial y^*}\right|_{y^*=0} = \text{Sh} \qquad (10.48)$$

This parameter, termed the *Sherwood number,* is equal to the dimensionless mass fraction (i.e., concentration) gradient at the surface, and it provides a measure of the convection mass transfer occurring at the surface. Following the same argument as before [but now for Eq. (10.41)), we have

$$\text{Sh} = f_4(\text{Re, Sc}) \quad \text{(forced convection, mass transfer)} \qquad (10.49)$$

The significance of expressions such as Eqs. (10.44) to (10.46) and (10.49) should be appreciated. For example, Eq. (10.45) states that convection heat-transfer results, whether obtained theoretically or experimentally, can be represented in terms of three dimensionless groups, instead of seven parameters (h, L, V, k, c_p, μ, and ρ). The convenience is evident. Once the form of the functional dependence of Eq. (10.45) is obtained for a particular surface geometry (e.g., from laboratory experiments on a small model), it is known to be universally applicable; i.e., it may be applied to different fluids, velocities, temperatures, and length scales, as long as the assumptions associated with the original equations are satisfied (e.g., negligible viscous dissipation and body forces). Note that the relations of Eqs. (10.44) and (10.49) are derived without actually solving the system of Eqs. (10.28) and (10.31). References 7 to 12 cover the preceding procedure with more details and also include many different cases.

It is important to mention here that once the conservation equations are put in dimensionless form, it is also convenient to make an order-of-magnitude assessment of all terms in the equations. Often a problem can be simplified by discovering that a term that would be very difficult to handle if large is in fact negligibly small.[7,8] Even if the primary thrust of the investigation is experimental, making the equations dimensionless and estimating the orders of magnitude of the terms are good practice. It is usually not possible for an experimental test to include (simulate) all conditions exactly; a good engineer will focus on the most important conditions. The same applies to performing an order-of-magnitude analysis. For example, for boundary-layer flows, allowance is made for the fact that lengths transverse to the main flow scale with a much shorter length than those measured in the direction of main flow. References 7, 11, and 13 cover many examples of the order-of-magnitude analysis.

When the governing equations of a problem are unknown, an alternative approach of deriving dimensionless groups is based on use of dimensional analysis in the form of the Buckingham pi theorem.[3,5,9,12,14] The Buckingham pi theorem proves that in a physical problem including n quantities in which there are m dimensions, the quantities can be arranged into $n - m$ independent dimensionless parameters. The success of this method depends on our ability to select, largely from intuition, the parameters that influence the problem. The procedure is best illustrated by an example.

Example 10.1. The discharge through a horizontal capillary tube is thought to depend on the pressure drop per unit length, the diameter, and the viscosity. Find the form of the equation. The quantities with their dimensions are as follows:

Quantity	Symbol	Dimensions
Discharge	Q	$L^3 t^{-1}$
Pressure drop/length	$\Delta p/l$	$ML^{-2}t^{-2}$
Diameter	D	L
Viscosity	μ	$ML^{-1}t^{-1}$

Then

$$F\left(Q, \frac{\Delta p}{l}, D, \mu\right) = 0$$

Three dimensions are used, and with four quantities there will be one Π parameter:

$$\Pi = Q^{x_1}\left(\frac{\Delta p}{l}\right)^{y_1} D^{z_1}\mu$$

Substituting in the dimensions gives

$$\Pi = (L^3 T^{-1})^{x_1}(ML^{-2}T^{-2})^{y_1}L^{z_1}ML^{-1}t^{-1} = M^0 M^0 L^0$$

The exponents of each dimension must be the same on both sides of the equation. With L first,

$$3x_1 - 2y_1 + z_1 - 1 = 0$$

and similarly for M and t,

$$y_1 + 1 = 0$$

$$-x_1 - 2y_1 - 1 = 0$$

from which $x_1 = 1$, $y_1 = -1$, $z_1 = -4$, and

$$\Pi = \frac{Q\mu}{D^4\, \Delta p/l}$$

After solving for Q,

$$Q = C\frac{\Delta p D^4}{l\mu}$$

from which dimensional analysis yields no information about the numerical value of the dimensionless constant C; experiment (or analysis) shows that it is $\pi/128$.

In the preceding example, if kinematic viscosity had been used in place of dynamic viscosity, an incorrect formula would have resulted.

Example 10.2.   A fluid-flow situation depends on the velocity V; the density ρ; several linear dimensions l, l_1, and l_2; pressure drop Δp; gravity g, viscosity μ; surface tension σ; and bulk modulus of elasticity E. Apply dimensional analysis to these variables to find a set of parameters.

$$F(V, \rho, l, l_1, l_2, \Delta p, g, \mu\; \sigma, E) = 0$$

Since three dimensions are involved, three repeating variables are selected. For complex situations, V, p, and l are generally helpful. There are seven Π parameters:

$$\Pi_1 = V^{x_1}\, \rho^{y_1}\, l^{z_1}\, \Delta p$$

$$\Pi_2 = V^{x_2}\, \rho^{y_2}\, l^{z_2}\, g$$

$$\Pi_3 = V^{x_3}\, \rho^{y_3}\, l^{z_3}$$

$$\Pi_4 = V^{x_4}\, \rho^{y_4}\, l^{z_4}\, \sigma$$

$$\Pi_5 = V^{x_5}\, \rho^{y_5}\, l^{z_5}\, E$$

$$\Pi_6 = \frac{l}{l_1}$$

$$\Pi_7 = \frac{l}{l_2}$$

By expanding the quantities into dimensions as in the first example, we have

$$\Pi_1 = \frac{\Delta p}{\rho V^2} \quad \Pi_2 = \frac{gl}{V^2} \quad \Pi_3 = \frac{\mu}{Vl_\rho} \quad \Pi_4 = \frac{\sigma}{V^2\rho l} \quad \Pi_5 = \frac{E}{\rho V^2} \quad \Pi_6 = \frac{l}{l_1} \quad \Pi_7 = \frac{l}{l_2}$$

and

$$f\left(\frac{\Delta p}{\rho V^2}, \frac{gl}{V^2}, \frac{\mu}{Vl_\rho}, \frac{\sigma}{V^2\rho l}, \frac{E}{\rho V^2}, \frac{l}{l_1}, \frac{l}{l_2}\right) = 0$$

It is convenient to invert some of the parameters and to take some square roots:

$$f_1\left(\frac{\Delta p}{\rho V^2}, \frac{V}{\sqrt{gl}}, \frac{Vl_\rho}{\mu}, \frac{V}{\sqrt{E/\rho}}, \frac{l}{l_1}, \frac{l}{l_2}\right) = 0$$

The first parameter, usually written $\Delta p/(\rho V^2/2)$, is the pressure coefficient; the second parameter is the Froude number Fr; the third is the Reynolds number Re; the fourth is the Weber number We; and the fifth is the Mach number Ma. Hence

$$f_1\left(\frac{\Delta p}{\rho V^2}, \text{Fr, Re, We, Ma}, \frac{l}{l_1}, \frac{l}{l_2}\right) = 0$$

After solving for pressure drop,

$$\Delta p = \rho V^2 f_2\left(\text{Fr, Re, We, Ma}, \frac{l}{l_1}, \frac{l}{l_2}\right)$$

in which f_1, f_2 must be determined from analysis or experiment. By selecting other repeating variables, a different set of pi parameters could be obtained.

For example, knowing in advance that the heat-transfer coefficient in fully developed forced convection in a tube is a function of certain variables, that is, $h = f(V, \rho, \mu, c_p, k, D)$, we can use the Buckingham pi theorem to obtain Eq. (10.45), as shown in Ref. 3. However, this method is carried out without any consideration of the physical nature of the process in question; i.e., there is no way to ensure that all essential variables have been included. However, as shown above, starting with the differential form of the conservation equations, we have derived the similarity parameters (dimensionless groups) in rigorous fashion.

In Table 10.10 those dimensionless groups which appear frequently in fluid-flow, heat-, and mass-transfer literature have been listed. The list includes groups already mentioned above as well as those found in special fields of heat transfer. Note that although similar in form, the Nusselt and Biot numbers differ in both definition and interpretation. The Nusselt number is defined in terms of thermal conductivity of the fluid; the Biot number is based on the solid thermal conductivity.

NOMENCLATURE

$$
\begin{aligned}
\textit{Symbol} &= \textit{Definition, SI Units (U.S. Customary Units)}
\end{aligned}
$$

A = heat-transfer area, m² (ft²)

$\mathbf{a}$ = acceleration, m/s² (ft/s²)

a = speed of sound, m/s (ft/s)

C = mass concentration of species, kg/m³ (lb_m/ft³)

c = specific heat, J/(kg · K) [Btu/(lb_m · °F)]

c_p = specific heat at constant pressure, J/(kg · K) [Btu/(lb_m · °F)]

c_v = specific heat at constant volume, J/(kg · K) [Btu/(lb_m · °F)]

D = tube inside diameter, diameter, m (ft)

$\mathbf{D}$ = diffusion coefficient, m²/s (ft²/s)

Ec = Eckert number (Table 10.10)

e = emissive power, W/m² [Btu/(h · ft²)]

e_b = blackbody emissive power, W/m² [Btu/(h · ft²)]

F = force, N (lb_f)

f = frequency of vibration (Table 10.10), s⁻¹

f_1, f_2, f_3, f_4 = denotes function of Eqs. (10.44) to (10.46) and (10.49)

Gr = Grashof number (Table 10.10)

$g, \mathbf{g}$ = gravitational acceleration (magnitude and vector), m/s² (ft/s²)

h = heat-transfer coefficient, W/(m² · K) [Btu/(h · ft² · °F)]

h_D = mass-transfer coefficient, m/s (ft/s)

i = enthalpy per unit mass, J/kg (Btu/lb_m)

i_{lg} = latent heat of evaporation, J/kg (Btu/lb_m)

i^0 = heat of formation, J/kg (Btu/lb_m)

$j, \mathbf{j}$ = mass diffusion flux of species (magnitude and vector), kg/(s · m²) [lb_m/(h · ft²)]

k = thermal conductivity, W/(m · K) [Btu/(h · ft · °F)]

L = length, m (ft)

M = mass, kg (lb_m)

m = mass fraction of species [Eq. (10.27)]

Nu = Nusselt number (Table 10.10)

P = pressure: Pa, N/m² (lb_f/ft²)

Pr = Prandtl number (Table 10.10)

ΔP = pressure drop, Pa, N/m^2 (lb_f/ft^2)

q = heat-transfer rate, W (Btu/h)

$\mathbf{q''}$ = heat flux (vector), W/m^2 [$Btu/(h \cdot ft^2)$]

q'' = heat flux, W/m^2 [$Btu/(h \cdot ft^2)$]

q''' = volumetric heat generation, W/m^3 [$Btu/(h \cdot ft^3)$]

Re = Reynolds number (Table 10.10)

r = radial distance in cylindrical or spherical coordinate, m (ft)

r''' = volumetric generation rate of species, $kg/(s \cdot m^3)$ [$lb_m/(h \cdot ft^3)$]

Sc = Schmidt number (Table 10.10)

Sh = Sherwood number (Table 10.10)

St = Stanton number (Table 10.10)

T = temperature: °C, K (°F, °R)

ΔT = temperature difference, °C (°F)

t = time, s

u = velocity component in the axial direction (x direction) in rectangular coordinates, m/s (ft/s)

$\mathbf{u}$ = internal energy per unit mass, J/kg (Btu/lb_m)

$V, \mathbf{V}$ = velocity (magnitude and vector), m/s (ft/s)

v = velocity component in the y direction in rectangular coordinates, m/s (ft/s)

v_r = velocity component in the r direction, m/s (ft/s)

v_z = velocity component in the z direction, m/s (ft/s)

v_θ = velocity component in the θ direction, m/s (ft/s)

v_ϕ = velocity component in the ϕ direction, m/s (ft/s)

w = velocity component in the z direction in rectangular coordinates, m/s (ft/s)

x = rectangular coordinate, m (ft)

y = rectangular coordinate, m (ft)

z = rectangular or cylindrical coordinate, m (ft)

Greek

α = thermal diffusivity, m^2/s (ft^2/s)

β = coefficient of thermal expansion, K^{-1} ($°R^{-1}$)

ϵ = emissivity

θ = angle in cylindrical and spherical coordinates, radians (degrees)

λ = molecular mean free path, m (ft)

μ = dynamic viscosity, $Pa \cdot s$ [$lb_m/(h \cdot ft)$]

ν = kinematic viscosity, m^2/s (ft^2/s)

ρ = density, kg/m^3 (lb_m/ft^3)

σ = surface tension (Table 10.10), N/m (lb$_f$/ft)

τ = shear stress, N/m^2 (lb$_f$/ft^2)

τ = shear stress tensor, N/m^2 (lb$_f$/ft^2)

Φ = dissipation function (Table 10.7), s^{-2}

ϕ = angle in spherical coordinate system, rad (degrees)

Subscripts

a = surroundings

aw = adiabatic wall

cr = critical

f = fluid

g = gas (vapor)

i = species i

l = liquid

m = mean

s = solid

sat = saturation

t = total

w = wall

x = x component

y = y component

z = z component

θ = θ component

ϕ = ϕ component

1 = species 1 in binary mixture of 1 and 2

∞ = free-stream condition

Mathematical Operation Symbols

d/dx = derivative with respect to x, m^{-1} (ft^{-1})

$\partial/\partial t$ = partial time derivative operator, s^{-1}

d/dt = total time derivative operator, s^{-1}

D/Dt = substantial time derivative operator, s^{-1}

∇ = "del" operator (vector), m^{-1} (ft^{-1})

∇^2 = laplacian operator, m^{-2} (ft^{-2})

*REFERENCES**

1. W. M. Kays and M. E. Crawford, *Convective Heat and Mass Transfer,* 2d ed., McGraw-Hill, New York, 1980.

2. R. B. Bird, W. E. Stewart, and E. N. Lightfoot, *Transport Phenomena,* Wiley, New York, 1960.

3. W. M. Rohsenow and H. Y. Choi, *Heat, Mass, and Momentum Transfer,* Prentice-Hall, Englewood Cliffs, N.J., 1961.

4. W. M. Rohsenow, J. P. Hartnett, and E. N. Ganić, eds., *Handbook of Heat Transfer Applications,* chap. 2, McGraw-Hill, New York, 1985.

5. A. S. Foust, L. A. Wenzel, C. W. Clump, L. Mans, and L. B. Andersen, *Principles of Unit Operations,* 2d ed., Wiley, New York, 1980.

6. S. Whitaker, *Elementary Heat Transfer Analysis,* Pergamon, New York, 1976.

7. D. K. Edwards, V. E. Denny, and A. F. Mills, *Transfer Processes: An Introduction to Diffusion, Convection, and Radiation,* 2d ed., Hemisphere, Washington, and McGraw-Hill, New York, 1979.

8. H. Schlichting, *Boundary-Layer Theory,* 7th ed., McGraw-Hill, New York, 1979.

9. B. Gebhart, *Heat Transfer,* 2d ed., McGraw-Hill, New York, 1971.

10. F. P. Incropera and D. P. DeWitt, *Fundamentals of Heat Transfer,* Wiley, New York, 1981.

11. F. M. White, *Viscous Fluid Flow,* McGraw-Hill, New York, 1974.

12. V. P. Isachenko, V. A. Osipova, and A. S. Sukomel, *Heat Transfer,* Mir Publishers, Moscow, 1977.

13. E. R. G. Eckert and R. M. Drake, Jr., *Analysis of Heat and Mass Transfer,* McGraw-Hill, New York, 1972.

14. J. H. Lienhard, *A Heat Transfer Textbook,* Prentice-Hall, Englewood Cliffs, N.J., 1981.

15. W. C. Reynolds and H. C. Perkins, *Engineering Thermodynamics,* 2d ed., McGraw-Hill, New York, 1977.

16. W. M. Rohsenow, J. P. Hartnett, and E. N. Ganić, eds, *Handbook of Heat Transfer Fundamentals,* 2d ed., chap. 1, McGraw-Hill, New York, 1985.

---

* Those references listed here but not cited in the text were used for comparison of different data sources, clarification, clarity of presentation, and, most important, reader's convenience when further interest in subject exists.

CHAPTER 11
TOPICS IN APPLIED PHYSICS

ELECTRIC FIELDS

1. Electric Charge. Any observed charge q, positive or negative, is in magnitude an integral multiple of

$$e = 1.6022 \times 10^{-19} \text{ C} \tag{11.1}$$

where $-e$ is the charge on a single electron. Note that C (coulomb) is the derived unit for electric charge, since $1 \text{ C} = 1 \text{ A} \cdot \text{s}$. One coulomb is a large amount of charge; hence the microcoulomb (μC) is frequently used.

An electric current in a wire corresponds to the transport of a certain amount of electric charge through a fixed cross section in a certain time interval. Electric current is measured in amperes (A).

2. Coulomb's Law. If q_1 and q_2 are magnitudes of two charges, separated by a distance r, then the force between them is

$$F = b \frac{q_1 \cdot q_2}{r^2} \tag{11.2}$$

This force is repulsive when q_1 and q_2 are of the same sign and attractive when they are of opposite signs. By experiment, it is found that

$$b \simeq 9 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2$$

For mathematical reasons, it is convenient to replace b by another constant, ϵ_0, called the *permissivity of empty space* and defined by

$$\epsilon_0 = \frac{1}{4\pi b} = 8.85432 \times 10^{-12} \text{ C}^2/(\text{N} \cdot \text{m}^2)$$

In terms of ϵ_0, Coulomb's law is written

$$F = \frac{q_1 \cdot q_2}{4\pi\epsilon_0 r^2} \tag{11.3}$$

Forces between point charges act independently. Suppose that charges q_1, q_2, q_3, . . . are fixed in an inertial frame. Then the force F_1 on a charge due to q_1 is

computed by Coulomb's law as if q_2, q_3, . . . did not exist; etc. Finally, the total force on q is just the vector sum of the individual forces:

$$F_{\text{total}} = F_1 + F_2 + F_3 + \cdots \tag{11.4}$$

3. Electric Fields. If a charge q experiences a force F in the field, then the electric field strength is defined to be the force per unit charge, that is,

$$E = \frac{F}{q} \tag{11.5}$$

Equivalently, the force on a point charge q located at a point at which the electric field is E (regardless of what arrangement of charges may have established E) is given by

$$F = qE \tag{11.6}$$

Note that E is a vector in the direction of F and represents the force (in newtons) on a unit charge (1 C) at the point considered.

From Eqs. (11.4) and (11.5),

$$E_{\text{total}} = E_1 + E_2 + E_3 + \cdots \tag{11.7}$$

That is, the electric field at any point due to a distribution of charges (or charge elements) q_1, q_2, q_3, . . . is found by adding (or integrating) the fields independently established at that point by individual charges.

4. Electric Flux. The *electric flux* through the elementary area dS is defined as

$$d\psi = E \cdot dS \tag{11.8}$$

Note that $d\psi$ can be interpreted as the number of lines (the lines to which E is tangent at each point) cutting dS, provided that the lines are supposedly drawn with normal density E.

The total flux through S in the outward direction is given by integration of Eq. (11.8) as

$$\psi = \int_S E \cdot dS \tag{11.9}$$

5. Gauss's Law. Based on definition in previous sections, Gauss's law reads

$$\psi = \frac{Q}{\epsilon_0} \tag{11.10}$$

That is, the total electric flux ψ out of an arbitrary closed surface in free space is proportional to the net electric charge within the surface Q; the constant of proportionality is $1/\epsilon_0 = 4\pi b$. Gauss's law is valid even for enclosed charges in motion.

6. Electric Potential. The electric force F exerted on a test charge q by some stationary distribution of charges is a conservative force (i.e., it can be derived from a function of position). Therefore, the test charge possesses electric potential energy

U. When U is known, then the components of the electric force F along the x, y, and z axes are given by

$$F_x = -\frac{\partial U}{\partial x} \qquad F_y = -\frac{\partial U}{\partial y} \qquad F_z = -\frac{\partial U}{\partial z}$$

or in vector form

$$\mathbf{F} = -\left(\frac{\partial U}{\partial x}\mathbf{i} + \frac{\partial U}{\partial y}\mathbf{j} + \frac{\partial U}{\partial z}\mathbf{k}\right) \qquad (11.11)$$

If the only force acting on a point charge is the electric force $\mathbf{F}$, the conservation of energy relation is

$$\Delta K_e + \Delta U = 0 \qquad (11.12)$$

where K_e, is the kinetic energy of the point charge.

The electric potential energy per unit test charge is called the *electric potential* ϕ or the *voltage v*. Note that ϕ and $\mathbf{E}$ enjoy the same relation as U and $\mathbf{F}$, that is,

$$\mathbf{E} = -\left(\frac{\partial \phi}{\partial x}\mathbf{i} + \frac{\partial \phi}{\partial y}\mathbf{j} + \frac{\partial \phi}{\partial z}\mathbf{k}\right) \qquad (11.13)$$

Notice that the unit of electric potential is the volt (V), where 1 V = 1 J/C; also, 1 V/m = 1 N/C. It follows also that

$$\phi_{\text{total}} = \phi_1 + \phi_2 + \phi_3 + \cdots \qquad (11.14)$$

The Electronvolt. When an electron (charge $-e$) travels through a potential difference of 1 V (i.e., when it moves from a given location to a location where the electric potential is 1 V higher), the electron loses

$$e\,(1\text{ V}) = 1.602 \times 10^{-19}\text{ J}$$

of electric energy and gains [Eq. (11.12)] the same amount of kinetic energy. This quantity of energy is referred to as *one electronvolt* (1 eV). In general, any energy expressed in joules may be converted to electronvolts by dividing by 1.602×10^{-19} J/eV.

7. Electric Current, Resistance, and Power. The rate of flow of electric charge across a given conductor area is defined as the *electric current I* through that area, or

$$I = \frac{dq}{dt} \qquad (11.15)$$

Also, in terms of J, the electric current through a surface S (e.g., a cross section of the conductor) is

$$I = \int_S J \cdot dS = \int_S J \cdot dA \qquad (11.16)$$

where $dA = dS \cos \theta$ is the projection of dS perpendicular to the flow direction, and J is the electric current density at a point within a conductor.

Ohm's Law. Empirically, it is found that the current I in a resistor (long wire, conducting bar, etc.) is very nearly proportional to the difference v in electric potential between the ends of the resistor. This proportionality is known as *Ohm's law:*

$$v = I \cdot R \tag{11.17}$$

where the proportionality factor R is called the resistance. The unit of resistance, the ohm (Ω), is defined as $1\ \Omega = 1\ \text{V/A}$. The resistance of a conductor of length L and uniform cross-sectional area A is

$$R = \rho \frac{L}{A} \tag{11.18}$$

where ρ, the *resistivity*, is a property of the material and is temperature-dependent. The units of ρ are ($\Omega \cdot \text{m}$).

The difference in potential between the terminals of any source of electric energy such as a battery or a mechanically driven generator, when delivering no current, is a measure of the electromotive force v_e of the source. Note that v_e is often called the *open-circuit voltage* of a battery or generator. The actual source of electric energy is

$$v_t = v_e \pm IR_0 \tag{11.19}$$

where v_t is terminal voltage, I is the current passing through the source, and R_0 is the internal resistance of the source. The negative sign is used when the source is delivering current, and the positive sign is used when I has the opposite direction (e.g., charging a storage cell).

If an amount of charge dq moves through a difference in potential v, the change in electric energy is

$$dE = (dq)v \tag{11.20}$$

The electric power P is the time rate of change of electric energy, that is,

$$P = \frac{dE}{dt} = \left(\frac{dq}{dt}\right) v = I \cdot v \tag{11.21}$$

Note that $(1\ \text{A})(1\ \text{V}) = 1\ \text{W}$ or $1\ \text{V} = 1\ \text{W/A}$. The power output of an energy source having terminal voltage v_t and supplying current I is given as

$$P = I \cdot v_t = I(v_e - I \cdot R_0) \tag{11.22}$$

The power absorbed by a pure resistance R when the potential difference across it is v_{1-2} is given by

$$P = I + v_{1-2} = I^2 R = \frac{v_{1-2}^2}{R} \tag{11.23}$$

This latter power is dissipated as heat.

8. Kirchhoff's Laws. For any network composed of resistors and sources (batteries), given the values of a certain number of the electromotive forces v_e, currents I, and resistances R, the values of the remaining quantities can be found by an application of Kirchhoff's laws.

Kirchhoff's Junction Law. The algebraic sum of all currents into any junction is zero. That is, at any junction,

$$\Sigma I = 0 \tag{11.24}$$

In the last equation, a current out of the junction is counted as negative.

Kirchhoff's Loop Law. The algebraic sum of all electromotive forces around any closed circuit (loop) is equal to the algebraic sum of the voltage drops (the $R \cdot I_R$ and the $R_0 \cdot I$) around the loop, that is,

$$\Sigma v_e = \Sigma R \cdot I_R + \Sigma R_0 I_{R_0} \tag{11.25}$$

where the summations are taken along any continuously directed path from p_1 to p_2 in the network. If v_{1-2} comes out positive, this means that p_1 is at a positive potential relative to p_2.

EXAMPLE 11.1 Referring to Fig. 11.1, find I_1, I_2, and I_3. Directions shown for the currents are arbitrary.

Junction Equations. At junction a,

$$I_1 - I_2 - I_3 = 0 \tag{1}$$

Another junction equation can be written at d, but it is the negative of Eq. (1).

Loop Equations. Traversing the loop *adcba* of the planar network, we obtain
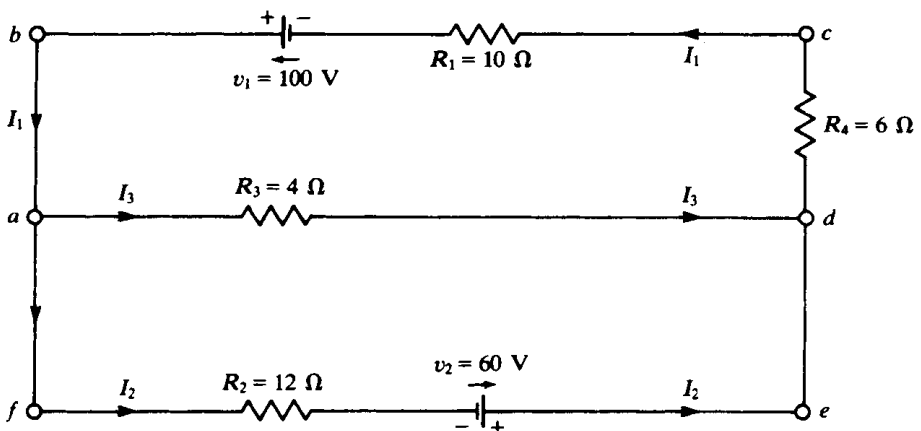
FIGURE 11.1 (*From Wells and Slusher,*[2] *p. 237.*)

$$100 = 4I_3 + 6I_1 + 10I_1 \tag{2}$$

and traversing *afeda*, we obtain

$$60 = 12I_2 + 4(-I_3) \tag{3}$$

(It is assumed that the batteries have no internal resistance.) Note that the loop *afedcba* would not give an independent equation, but rather the sum of Eq. (2) and Eq. (3).

The independent equations (1), (2), and (3) can now be solved simultaneously (applying usual methods) for I_1, I_2, and I_3, yielding

$$I = 6.053 \text{ A} \qquad I_2 = 5.263 \text{ A} \qquad I_3 = 0.7895 \text{ A}$$

Note that each value of I is positive; thus the arbitrarily chosen directions indicated in Fig. 11.1 happen to be correct.

EXAMPLE 11.2   Figure 11.2*a* and *b* shows two possibilities, the so-called series and parallel combinations of resistors. Using also Ohm's law, it can be shown that resistances in series add directly and resistances in parallel add as their reciprocals.

EXAMPLE 11.3   Figure 11.3*a* and *b* shows two possibilities, the capacitors connected in parallel (*a*) and the capacitors connected in series (*b*). Where capacitors C_1, C_2, C_3, ... are added in parallel (Fig. 11.3*a*), the total capacitance C increases, that is,

$$C = C_1 + C_2 + C_3 \tag{1}$$

Where capacitors C_1, C_2, C_3 . . . . are added in series (Fig. 11.3*b*), the total capacitance C decreases, that is,

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} \tag{2}$$

The *capacitance C* of a capacitor is the ratio of quantity of electricity q stored in it and voltage v across it, that is,

FIGURE 11.2

(a)

(b)
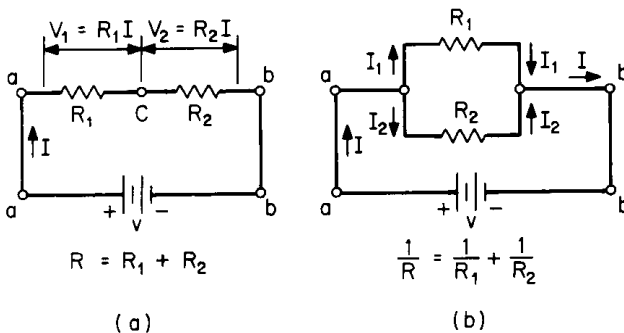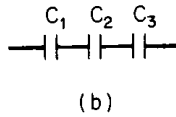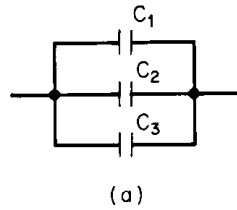
FIGURE 11.3

$$C = \frac{q}{v} \tag{3}$$

Where a capacitor requires a charge of 1 C to be charged to a voltage of 1 V, its capacitance is 1 F (farad), that is,

$$1\ F = 1\ \frac{C}{V} = 1\ \frac{A \cdot s}{V}$$

9. Direct and Alternating Currents. *Direct current* (dc) flows in one direction only through a circuit. The associated direct voltages, in contrast to alternating voltages, are of unchanging polarity. Direct current corresponds to a drift or displacement of electric charge in one unvarying direction around the closed loop or loops of an electric circuit. Direct currents and voltages may be of constant magnitude or may vary with time.

Alternating current (ac) reverses direction periodically, usually many times per second. One complete period, with current flow first in one direction and then in the other, is called a *cycle,* and 60 cycles per second (60 hertz, or Hz) is the customary frequency of alternation in the United States and 50 Hz in Europe.

Alternating current is shown diagrammatically in Fig. 11.4. Time is measured horizontally (beginning at any arbitrary moment), and the current at each instant is
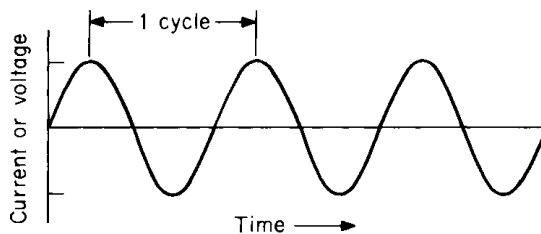
FIGURE 11.4   Diagram of sinusoidal alternating current.

measured vertically. In this diagram it is assumed that the current is alternating sinusoidally; that is, the current i is described by the following relation:

$$i = I_m \sin 2\pi f t \tag{11.26}$$

where I_m is the maximum instantaneous current, f is frequency per second (Hz), and t is time in seconds. Electric energy is ordinarily generated as alternating current.

MAGNETIC FIELDS

1. Magnetism.   *Magnetism* comprises those physical phenomena involving magnetic fields and their effects on materials. Magnetic fields may be set up on a macroscopic scale by electric currents or by magnets. On an atomic scale, individual atoms cause magnetic fields when their electrons have a net magnetic moment as a result of their angular momentum. A magnetic moment arises whenever a charged particle has an angular momentum. It is the cooperative effect of the atomic magnetic moments that causes the macroscopic magnetic field of a permanent magnet.

2. Magnetic Field Forces.   A region in space is the site of a magnetic field if a test charge q moving in the region experiences a force **F** by virtue of its motion with respect to an inertial frame. This force may be described in terms of a field vector **B,** called the *magnetic induction* or *magnetic flux density,* or simply, the *magnetic field.* If u is the velocity of the charge, then the force is given by

$$\mathbf{F} = q(\mathbf{u} \times \mathbf{B}) \tag{11.27}$$

or in magnitude,

$$F = quB \sin \theta \tag{11.28}$$

where θ is the smaller angle between the vectors **u** and **B**. The unit of B is the tesla (T), where

$$1 \text{ T} = 1 \frac{\text{N} \cdot \text{s}}{\text{C} \cdot \text{m}} = 1 \frac{\text{N}}{\text{A} \cdot \text{m}} = 1 \frac{\text{V} \cdot \text{s}}{\text{m}^2} \tag{11.29}$$

When both a magnetic field **B** and an electric field **E** are present in the region, the force on a charge q is the vector sum of the electric and magnetic forces, that is

$$\mathbf{F}_{\text{tot}} = q[\mathbf{E} + (\mathbf{u} \times \mathbf{B})] \tag{11.30}$$

This relation is known as the *Lorentz equation.*
An element of length ds of a wire of arbitrary shape carrying current I in a magnetic field **B** experiences a force $d\mathbf{F}$ given by

$$d\mathbf{F} = I(d\mathbf{s} \times \mathbf{B}) \tag{11.31}$$

Here the d**s** vector is of magnitude ds and in the direction of motion of positive

charge. Net force on the wire (assuming a straight wire of length L) in a uniformed field **B** is, by integration,

$$\mathbf{F} = I(\mathbf{L} \times \mathbf{B}) \tag{11.32}$$

or

$$F = ILB \sin \theta \tag{11.33}$$

Magnetic Flux. Just as electric flux is associated with the electric field **E**, so magnetic flux Φ is associated with the magnetic field **B**. Thus the magnetic flux through an elementary area dS is defined as

$$d\Phi = \mathbf{B} \cdot dS \tag{11.34}$$

or

$$\Phi = \int_S \mathbf{B} \cdot dS \tag{11.35}$$

The unit of magnetic flux is the weber (Wb), where

$$1 \text{ Wb} = 1 \text{ T} \cdot \text{m}^2 = 1 \text{ V} \cdot \text{s} = 1 \text{ J/A}$$

Also, uncommonly, **B** (the magnetic flux density) is given in webers per square meter (Wb/m^2).

3. Sources of the Magnetic Field. The *magnetic field of a moving charge* is given by

$$\mathbf{B} = \frac{\mu_0 q}{4\pi r^3} (\mathbf{u} \times \mathbf{r}) \tag{11.36}$$

where **u** is the velocity of a point charge q, **r** is the displacement vector, and μ_0 is the permeability of empty space ($= 4\pi \times 10^{-7}$ H/m). Here the henry (H) is the unit of inductance, where

$$1 \text{ H} = 1 \frac{\text{Wb}}{\text{A}}$$

In magnitude (in μT), we have

$$B = \frac{qu}{10r^2} \sin \theta \tag{11.37}$$

where θ is the angle between **u** and **r**.
   The *magnetic field of a current filament* is given by

$$d B = \frac{\mu_0 I}{4\pi r^3} (d\mathbf{s} \times \mathbf{r}) \tag{11.38}$$

Equation (11.38) is known as the *Biot-Savart law.*

Ampere's circuital law reads

$$\oint \mathbf{B} \cdot d\mathbf{s} = \mu_0 I \tag{11.39}$$

where the line integral on the left side is around an arbitrary closed path l, and I is the current through any open surface bounded by the path l.

4. Faraday's Law. *Faraday's law of induction* is related to an induced electromotive force (emf) to the change in magnetic flux that produces it. For any flux change that takes place in a circuit, Faraday's law states that the magnitude of the emf (v_i) induced in the circuit is proportional to the rate of change of flux as

$$v_i \approx -\frac{d\Phi}{dt} \tag{11.40}$$

The time rate of change of flux in this expression may refer to any kind of flux change that takes place. If the change is motion of a conductor through a field, $d\Phi/dt$ refers to the rate of cutting flux. If the change is an increase or decrease in flux linking a coil, $d\Phi/dt$ refers to the rate of such change. For a coil of N turns,

$$v_i = -N\frac{d\Phi}{dt} \tag{11.41}$$

5. Inductance. For an explanation of self-inductance of a coil refer to Fig. 11.5. In Fig. 11.5, current I establishes a magnetic field in and around coil C_0 as indicated. Each line of flux threads (or "links") some or all of the turns. The total linkage is $N\Phi$, where N is the total number of turns and Φ is the average flux linking a turn.
    If Φ changes (as by sliding the contact on R), an electromotive force

$$v_i = -N\frac{d\Phi}{dt}$$

is induced in the coil. For an increase in Φ (decrease in R), $d\Phi/dt$ is positive and v_i is negative, that is, in the direction opposite to that of I. For Φ decreasing (increasing R), v_i is positive, that is, in the direction of I.
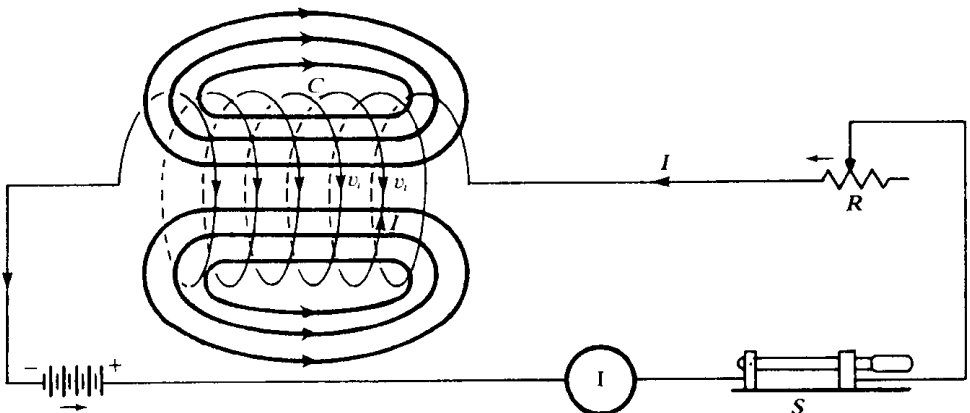
FIGURE 11.5   An example of a self-inductance coil. (*From Wells and Slusher,*[2] *p. 272.*)

Assuming C_0 in empty space and no magnetic material nearby, the flux established, and thus the flux linkage, is directly proportional to I, that is,

$$N\Phi = LI \quad \text{or} \quad L = \frac{N\Phi}{I} \tag{11.42}$$

The proportionality factor L is referred to as the *self-inductance* (or just *inductance*). The SI unit of inductance, the henry (H), was defined in the section ''Sources of the Magnetic Field.''

Differentiating Eq. (11.42) with respect to time gives for the (back-) emf in the inductor

$$v_i = -N\frac{d\Phi}{dt} = -L\frac{dI}{dt} \tag{11.43}$$

This amount of energy stored in the final magnetic field around the inductor, i.e., principally in the space enclosed by the turns of C_0, is

$$E = \tfrac{1}{2}LI^2 \tag{11.44}$$

SIMPLE ELECTRIC CIRCUITS (EXAMPLES)

1. Series R-L-C Circuit.   The type of circuit to be treated here is illustrated in Fig. 11.6. When switch S is closed, the source (a battery) begins to deliver charge q, thus establishing a current I in the circuit. Given the values of R, L, C, and emf v (assumed constant), we have to find expressions for q and I as functions of time. A more complete analysis also includes a determination of voltages v_{ab}, v_{bc}, and v_{cd}; energies stored in L and C; and the power delivered by the source—all as functions of time.

Applying Kirchhoff's voltage relation to the circuit, we have

$$v = v_{ab} + v_{bc} + v_{cd} = RI + L\frac{dI}{dt} + \frac{q}{C}$$

where the three terms on the right have been respectively obtained from relations given in previous sections. Writing $\dot{q}$ for I, we put this equation in the form
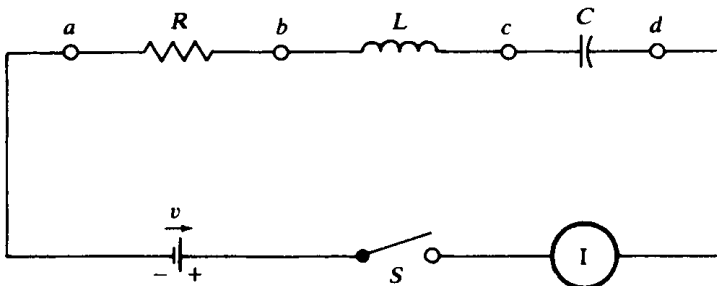
FIGURE 11.6   R-L-C circuit. (*From Wells and Slusher,*[2] *p. 285.*)

$$L\ddot{q} + R\dot{q} + \frac{1}{C}q = v \tag{11.45}$$

A solution of this linear second-order differential equation, subject to appropriate initial conditions, gives q as a function of t, and from this all other results can be obtained.

2. Series AC Circuit.   Referring to the circuit in Fig. 11.7, assume that a sinusoidal voltage, $u_a = v_a \sin \omega t$ is maintained between points a and b by an ac generator ("alternator"), where u_a is the instantaneous value of the voltage, v_a is the maximum value or amplitude of the voltage wave, and $\omega = 2\pi f$, where f is frequency in hertz (Hz).

The instantaneous values of voltage across resistor R, inductor L, and capacitor C are expressed in terms of instantaneous current i and instantaneous charge q by

$$u_R = Ri = R\dot{q} \qquad u_L = L\frac{di}{dt} = L\ddot{q} \qquad u_C = \frac{q}{C}$$

These must sum to u_a, which gives as the differential equation of the circuit

$$L\ddot{q} + R\dot{q} + \frac{1}{C}q = v_a \sin \omega t \tag{11.46}$$

After solving the last equation,

$$i = \frac{v_a}{Z} \sin (\omega t - \varphi) \equiv I \sin (\omega t - \varphi)$$

$$u_R = IR \sin (\omega t - \varphi) \equiv v_R \sin (\omega t - \varphi)$$

$$u_L = IX_L \cos (\omega t - \varphi) \equiv v_L \cos (\omega t - \varphi)$$

$$u_C = -IX_C \cos (\omega t - \varphi) \equiv -v_C \cos (\omega t - \varphi)$$

where the amplitudes of the four waves have been indicated as $I \equiv v_a/Z$, $v_R \equiv IR$, $v_L \equiv IX_L$, and $v_C \equiv IX_C$. Also,
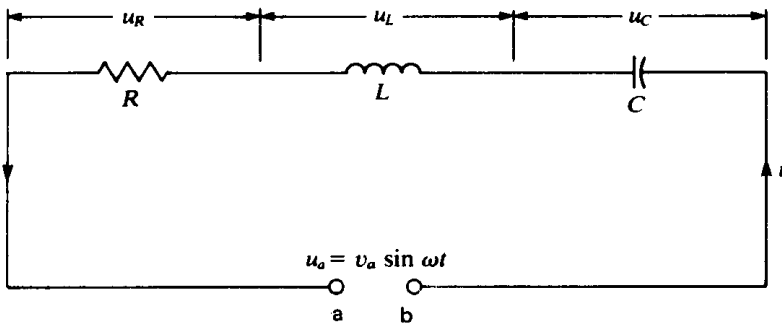
FIGURE 11.7   AC circuit. (*From Wells and Slusher,*[2] *p. 292.*)

$$X_C \equiv (\omega C)^{-1}$$

$$X_L \equiv \omega L$$

$$Z \equiv \sqrt{R^2 + (X_L - X_C)^2}$$

$$\tan \varphi \equiv \frac{X_L - X_C}{R}$$

WAVES

1. Wave Phenomena. A *wave* is a periodic disturbance that transports energy from one place to another. By *a periodic disturbance,* we mean certain changes that repeat themselves again and again in time. Examples where wave motion occurs include:

a) water waves, such as are produced when a stone is thrown into a still pool of water

b) waves on strings

c) waves on stretched springs

d) sound waves

e) light waves

f) radio waves

g) infra-red waves, which are emitted by hot bodies

h) ultra-violet waves, which are emitted by very hot bodies and some gas discharge lamps

i) x-ray waves, which are emitted by metals when they are bombarded by high speed electrons

j) gamma-rays, which are emitted by radioactive elements.

Examples a) to d) are mechanical waves and they require a medium (such as air or water) in order to move. Examples e) to j) are electromagnetic waves and do not require any medium—they can pass through a vacuum.

A wave $y(x, t)$ traveling along x with speed v is given by

$$y = f(x \pm vt) \tag{11.47}$$

where the minus and plus signs refer to wave propagation in the positive and negative x directions, respectively. This function satisfies the one-dimensional wave equation, that is,

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 y}{\partial t^2} \tag{11.48}$$

The speed of a wave traveling on a stretched uniform string, causing small transverse displacement $y(x, t)$, is given as

$$v = \sqrt{\frac{F}{\rho}} \tag{11.49}$$

where F is the tension in the string and ρ is the density of the string. The instantaneous power transmitted by the wave is given by

$$P = -F \frac{\partial y}{\partial x} \frac{\partial y}{\partial t} \tag{11.50}$$

The sinusoidal traveling wave

$$y = A \cos (kx - \omega t) \tag{11.51}$$

has a wave speed

$$v = \frac{\omega}{k}$$

where period T and wavelength λ are given by

$$T = \frac{2\pi}{\omega} \tag{11.52}$$

$$\lambda = \frac{2\pi}{k} \tag{11.53}$$

or

$$v = \frac{\lambda}{T} = \lambda f \tag{11.54}$$

where f is frequency.

2. Longitudinal and Transverse Waves. There are two very different kinds of waves whose motion satisfies Eq. (11.54). The first of these is a *longitudinal wave* in which the periodic motion of the particles producing the wave is parallel to the direction of propagation of the wave (see Fig. 11.8*a*). For example, in a sound wave in air, the air molecules vibrate back and forth in the direction in which the wave is moving and pass on energy to other molecules by collisions. In this way energy is transported through the air. Since the result of this vibrational motion is to produce a series of high- and low-pressure regions of space in the direction in which the wave is traveling, a sound wave is sometimes called a *compressional wave.*

A second, very different kind of wave is one in which the particles involved in the wave move at right angles to the direction of propagation. This we call a *transverse wave* (see Fig. 11.8*b*). A good approximation of a transverse wave is a water wave on the surface of a pond. The water molecules execute a periodic motion up and down at their fixed positions on the surface of the pond. The wave, however, moves along the surface of the pond at right angles to the up-and-down motion of the water molecules.
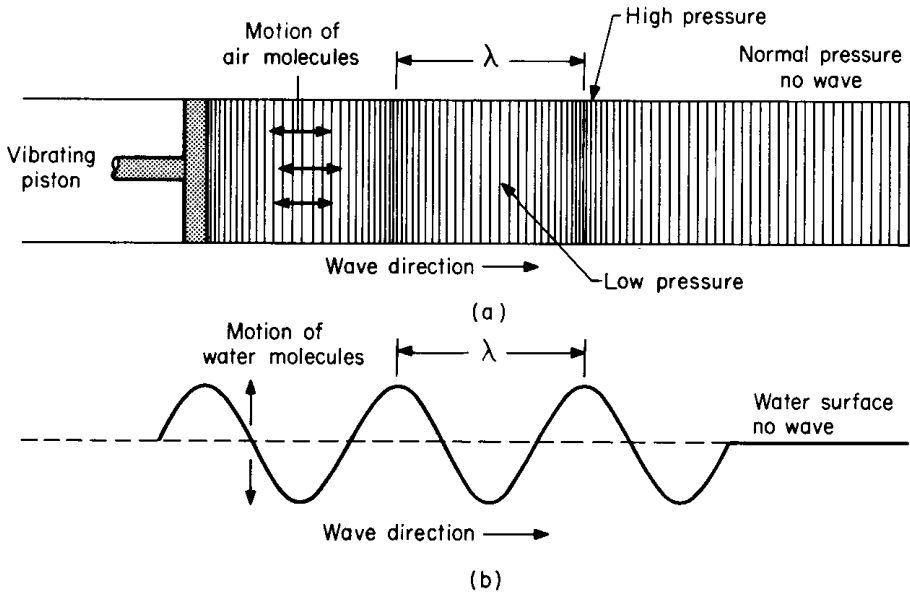
FIGURE 11.8 Longitudinal sound wave (*a*) and transverse water wave (*b*). (*From Mulligan,*[3] *p. 288.*)

3. Sound Waves. Sound waves are longitudinal waves in a medium. The longitudinal displacement l of the medium and the excess pressure p_e are related by

$$p_e = -E_0 \frac{\partial l}{\partial x} \tag{11.55}$$

where E_0 is the bulk modulus of the medium. Note that p_e and l satisfy the wave equation [Eq. (11.48)]. The wave speed is given by

$$v = \sqrt{\frac{E_0}{\rho_0}} \tag{11.56}$$

where ρ_0 is the density of the medium when it is in equilibrium. If the medium is a gas,

$$v = \sqrt{\frac{\gamma p_0}{\rho_0}} = \sqrt{\frac{\gamma K T}{m}} \tag{11.57}$$

where p_0 = undisturbed gas pressure
γ = specific heat ratio
K = Boltzmann's constant (1.38×10^{-23} J/K)
m = mass of the molecule of the gas

If γ is constant, then the speed of sound in a given gas varies as the square root of the absolute temperature.

Doppler Effect.   Suppose that a source emitting sound waves of frequency f_s and an observer move along the same straight line. Then the observer will hear sound of frequency

$$f_o = f_s \frac{v + v_o}{v - v_s}$$

where v_o and v_s are the velocities of the observer and source relative to the transmitting medium, and v is the sound speed in this medium. Velocity v_o is positive if the observer is heading toward the source, and v_s is positive if the source is heading toward the observer.

Transverse waves (light) also exhibit a *Doppler effect,* but in the case of light, the effect depends only on the relative velocity between source and observer, there being no transmitting medium.

4. Electromagnetic Waves.   In the case of mechanical waves, such as sound waves or water waves, there is always a material medium through which the wave passes, the molecules of a gas for a sound wave and the molecules of water for a water wave. Without these molecules, there would be no wave. However, there is another highly important class of waves that require no material medium for their propagation through space. These are the electromagnetic waves produced, for example, by electrons oscillating back and forth in the transmitting antenna of a radio station. These electromagnetic waves travel through a perfect vacuum just as easily as they move through air.

Electromagnetic waves are transverse waves consisting of periodic variations of electric and magnetic fields in space, and these fields can exist even where no matter is present. Sound waves in gases have different speeds depending on the properties of the gas, and these speeds often change with frequency. Electromagnetic waves in a vacuum (sometimes called *free space*), on the other hand, have speeds determined only by the properties of space. All electromagnetic waves in free space, therefore, travel with the same speed, no matter what the frequency or wavelength. This speed is found from experiments to be

$$c = 3.00 \times 10^8 \text{ m/s}$$

Since $c = \lambda f$, from Eq. (11.54), if the frequency of an electromagnetic wave is known, its wavelength is easily determined, and vice versa.

All electromagnetic waves (i.e., all waves produced by changing electric and magnetic fields) are of the same basic nature and travel at the same speed in a vacuum. They differ only in their frequency and wavelength. Waves of different frequencies interact differently with matter, however, and often appear to have very dissimilar properties. This is one example of a frequently occurring situation in physics in which quantitative changes (in this case, in wavelength and frequency) lead to important qualitative differences. Figure 11.9 shows the complete electromagnetic spectrum, with the wavelengths, frequencies, and energies corresponding to each part of that spectrum. It stretches all the way from long radio waves with wavelengths longer than 10^6 m to gamma rays with wavelengths shorter than 10^{-12} m.

Energy (J)	Frequency (Hz)	Wavelength	Type of radiation
10^{-30}	10^{3}	1000 km	Long radio waves
10^{-29}	10^{4}	100 km	
10^{-28}	10^{5}	10 km	
10^{-27}	10^{6}	1 (mil) 1 km	AM broadcast
10^{-26}	10^{7}	100 m	Radio
10^{-25}	10^{8}	10 m	
10^{-24}	10^{9}	1 m	
10^{-23}	10^{10}	10 cm	FM, TV broadcast
10^{-22}	10^{11}	(1 in) 1 cm	Shortwave radio
10^{-21}	10^{12}	1 mm	
10^{-20}	10^{13}	$\frac{1}{10}$ mm	Infrared
10^{-19}	10^{14}	$\frac{1}{100}$ mm	
10^{-18}	10^{15}	$1\ \mu m = 10^{-6}$ m	Visible
10^{-17}	10^{16}	$\frac{1}{10}\ \mu m$	Ultraviolet
10^{-16}	10^{17}	$\frac{1}{100}\ \mu m$	
10^{-16}	10^{18}	$10^{-3}\ \mu m$	
10^{-14}	10^{19}	10^{-10} m	X ray
10^{-13}	10^{20}	10^{-11} m	
		10^{-12} m	Gamma ray

FIGURE 11.9 The complete electromagnetic spectrum. (*From Mulligan,*[3] *p. 291.*)

5. Light. As we have seen from Fig. 11.9, visible light occupies that portion of the electromagnetic spectrum consisting of wavelengths from 0.4 μm (that is, 0.4×10^{-6} m) to 0.7 μm. From a physical point of view, the ultraviolet region, with wavelengths shorter than 0.4 μm, blends continuously into the visible region, which in turn blends continuously into the infrared region, with wavelengths longer than 0.7 μm. (Note that the *infra* in *infrared* and the *ultra* in *ultraviolet* refer to frequency, not wavelength.) The distinguishing feature of the visible portion of the spectrum is the human eye's sensitivity to wavelengths between 0.4 and 0.7 μm, in contrast to its insensitivity to infrared and ultraviolet radiation. The human eye sees each different wavelength of the visible region as a different color; violet corresponds to 0.4 μm and red to 0.7 μm, and there is a continuous change from violet to blue-green to orange to yellow to red between these two limits.

In free space, the speed of all electromagnetic waves is the same. This is not true in a material substance such as glass, where different wavelengths have different speeds.

Reflection and Refraction. Consider a ray of light. In an isotropic medium, rays are straight lines, along which energy travels at speed $v = c/n$, where n is the refractive index of the medium. Because $n > 1$, the speed is less than the speed in

vacuum, $c = 3 \times 10^8$ m/s. The refractive index is a function of the wavelength (in vacuum) of the light.

Let a ray in medium 1 be incident on the interface with medium 2, making angle θ_1 with the normal to the interface (Fig. 11.10). Then, in general, there will be a reflected ray in medium 1 and a refracted ray in medium 2 such that

1. The three rays and the normal all lie in a common plane, the plane of incidence.
2. The angle of incidence equals the angle of reflection, that is, $\theta_1 = \theta_r$.
3. The directions of the incident and refracted rays are related by Snell's law:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

If $n_1 > n_2$ and θ_1 exceeds the critical angle θ_c, where

$$\sin \theta_c = \frac{n_2}{n_1}$$

then there will be no refracted ray—a phenomenon called *total reflection* or, from the point of view of medium 1, *total internal reflection.*

A displacement of light rays, i.e. refraction of light, is again clearly demonstrated in Fig. 11.11, where the path of a ray of light passes through a parallel-sided glass block. The incident ray AB that has an angle of incidence α enters the glass block at B. The direction of the ray changes to BC such that the angle β is less than angle α. β, called the angle of refraction. When the ray emerges from the glass at C, the direction changes to CD, angle β' being greater than α'. The final emerging ray CD is parallel to the incident ray AB.

In general, when entering a more dense medium from a less dense medium, light is refracted towards the normal and when it passes from a dense to a less dense medium it is refracted away from the normal.

Geometrical Optics. Geometrical optics is concerned with the effects on light rays of mirrors (plane, concave, convex) and lenses (converging, diverging). *Lenses* are pieces of glass or other transparent material with a spherical surface on one or both sides. When light is passed through a lens it is refracted.
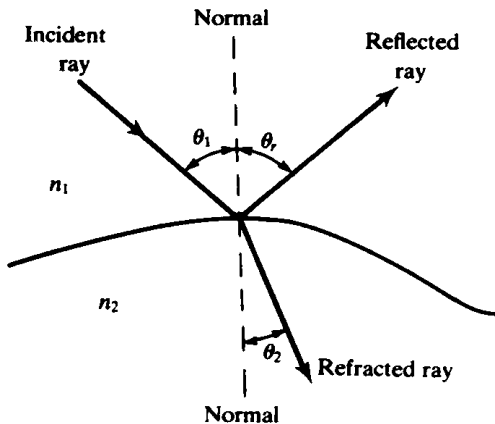
FIGURE 11.10 Illustration of the light refraction. (*From Wells and Slusher,*[2] *p. 301.*)
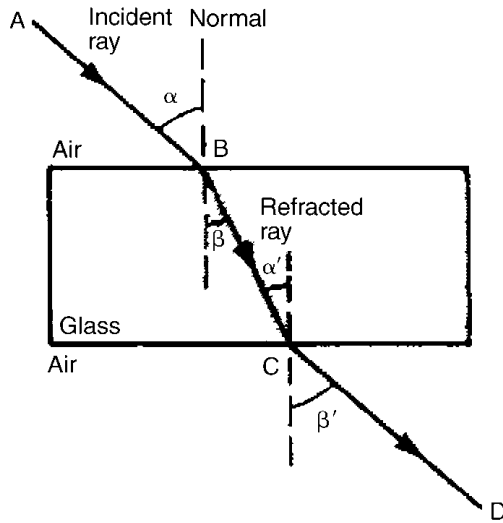
FIGURE 11.11   Displacement of light rays.

Lenses are used in spectacles, magnifying glasses and microscopes, telescopes, cameras and projectors. There are a number of different shaped lenses, and two of the most common are shown in Fig. 11.12. Figure 11.12*a* shows a bi-convex lens, so called since both its surfaces curve outwards. Figure 11.12*b* shows a bi-concave lens, so called since both of its surfaces curve inwards. The line passing through the center of curvature of the lens surface is called the principal axis. Figure 11.12*c* shows a number of parallel rays of light passing through a bi-convex lens. They are seen to converge at a point F on the principal axis. Figure 11.12*d* shows parallel rays of light passing through a bi-concave lens. They are seen to diverge such that they appear to come from a point F, which lies between the source of light and the lens, on the principal axis. In both Figure cases, F is called the principal focus or the focal point, and the distance from F to the center of the lens is called the focal length of the lens. It is assumed here that all lenses are thin (i.e., their thickness is small compared to their radii of curvature) and that all rays are paraxial (i.e., they make small angles with the axis of the optical system).

The two basic relations for both mirrors and lenses are

$$\frac{1}{d_o} + \frac{1}{d_i} = \frac{1}{f} \tag{11.58}$$

and

$$\frac{d_o}{d_i} = \frac{\ell_o}{\ell_i} \tag{11.59}$$

where f = focal length (cm)
  d_o = distance of object (cm)
  d_i = distance of image (cm)
  ℓ_o = linear size of object (cm)
  ℓ_i = linear size of image (cm)

The five quantities are conventionally given in centimeters. Each quantity carries
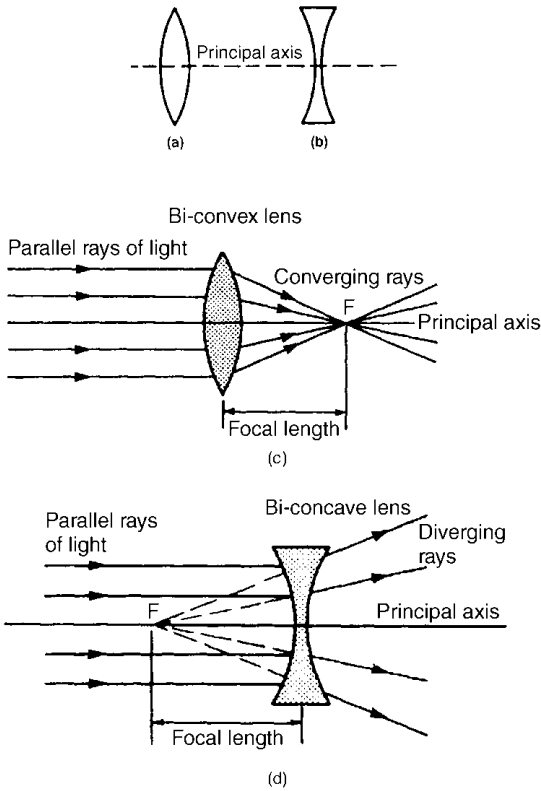
FIGURE 11.12 Lenses.

an algebraic sign, as specified in Table 11.1. The light rays diverge from a real object and converge toward a virtual object. Contrarily, light rays converge toward a real image (which could not be projected directly on a screen).

EXAMPLE 11.4 A simple periscope arrangement is shown in Fig. 11.13. A ray of light from O strikes a plane mirror at an angle of 45° at point P. Since from the laws of reflection the angle of incidence α is equal to the angle of reflection β,

TABLE 11.1

	Sign	
Quantity	+	−
f	converging lens, concave mirror	diverging lens, convex mirror
d_o	real object	virtual object
d_i	real image	virtual image
ℓ_o	erect object	inverted object
ℓ_i	erect image	inverted image

Source: From Wells and Slusher,[2] p. 309.
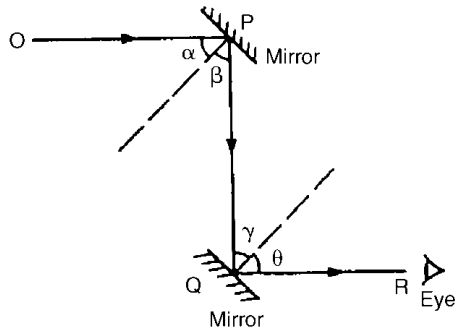
FIGURE 11.13   Periscope.

then $\alpha = \beta = 45°$. Thus angle OPQ $= 90°$ and the light is reflected through 90°. The ray then strikes another mirror at 45° at point Q. Thus $\gamma = \theta = 45°$, angle PQR $= 90°$ and the light ray is again reflected through 90°. Thus the light from O finally travels in the direction QR, which is parallel to OP, but displaced by the distance PQ. The arrangement thus acts as a **periscope.**

EXAMPLE 11.5   Describe the image formed by a concave mirror when a real object is situated outside the center of curvature C (Fig. 11.14). In Fig. 11.11, real rays a, b, c are drawn from a point P_1 on the real object P_1P_2. For convenience, a is drawn parallel to the optical axis, b is drawn through F, and c is drawn through C. From the geometry of the mirror and the law of reflection, reflected ray a' passes through F, b' is parallel to the optical axis, and c' returns through C along the path of c. The intersection of the real rays o', b', and c' at P_1' locates one point on the real image $P_1'P_2'$. Rays a, b, and c may seem very special. However, any (paraxial) ray from P_1, after reflection, passes through P_1', as may be shown by an application of the law of reflection. Note that the image is inverted.
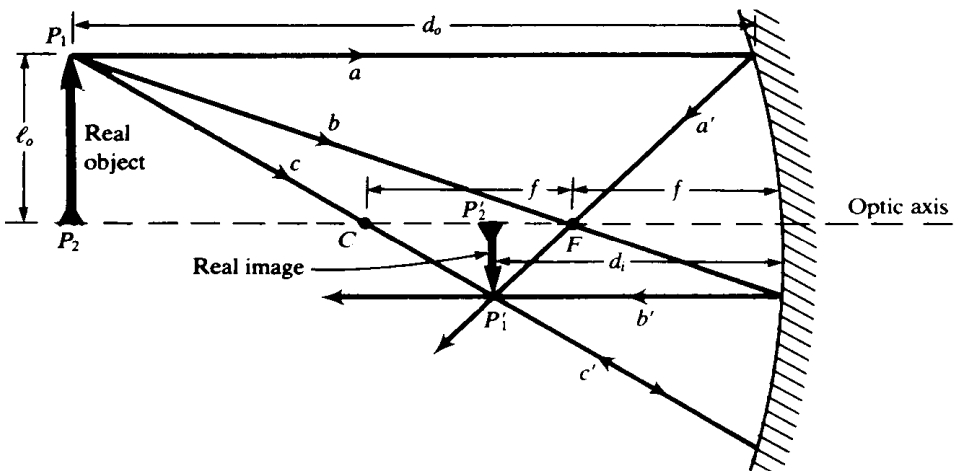
FIGURE 11.14   Sketch of the concave mirror image. (*From Wells and Slusher,*[2] *p. 311.*)
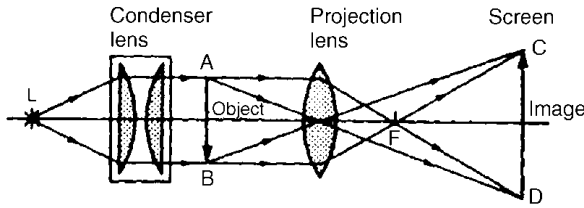
FIGURE 11.15   Projector.

Let $f = -20$ cm, $d_o = +45$ cm, and $\ell_o = +5$ cm. Then Eq. (11.58) gives

$$\frac{1}{45} + \frac{1}{d_i} = \frac{1}{20} \qquad \text{or} \qquad d_i = +36 \text{ cm}$$

(a real image), and Eq. (11.59) gives

$$\frac{45}{36} = -\frac{5}{\ell_i} \qquad \text{or} \qquad \ell_i = -4 \text{ cm}$$

(an inverted, minified image).

EXAMPLE 11.6   A simple *projector* arrangement is shown in Fig. 11.15 and consists of a source of light and two-lens system. L is a brilliant source of light, such as a tungsten filament. One lens system, called the condenser (usually consisting of two converging lenses as shown), is used to produce an intense illumination, of the object AB, which is a slide transparency or film. The second lens, called the projection lens, is used to form a magnified, real, upright image of the illuminated object on a distant screen CD.

LASERS

Lasers are commonly thought of as a kind of "death ray." Light from a laser differs in two respects from light from a normal source. First, it is absolutely monochromatic; it consists of light of only one frequency. Second, it is coherent. All light (because it is a form of a electromagnet radiation) is wavelike in nature. Light from conventional sources is emitted with random phase, as shown in Fig. 11.16a, so even if it is monochromatic some cancellation will occur. All the light from a laser, however, is exactly in phase, as in Fig. 11.16b, and reinforces rather than cancels.

If an atom absorbs energy, say from being heated, electrons move out to larger radius orbits. The atom is then said to be in an excited state. Eventually the electrons return to a lower orbit, releasing energy in the form of a packet of light called a photon. Because the electron orbits are fixed, only certain energy gains and losses are allowed.

Possible energy states for hydrogen are shown in Fig. 11.17. A minimum of 10 eV is required to lift it to state 1. Frequency of the emitted light depends on the energy change and is given by:

$$E_2 - E_1 = hf$$

where E_1 and E_2 are the energy *states,* f is the frequency and h is Planck's constant.
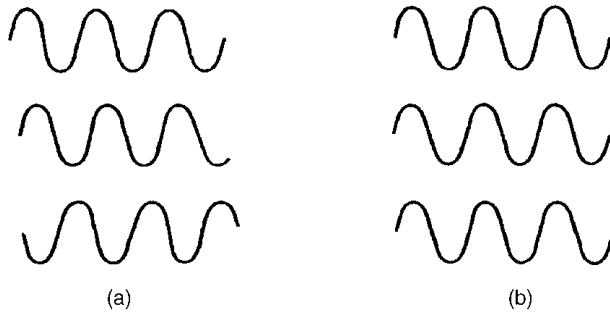
FIGURE 11.16 Coherent and incoherent light: (*a*) single-frequency, incoherent light; (*b*) coherent light.
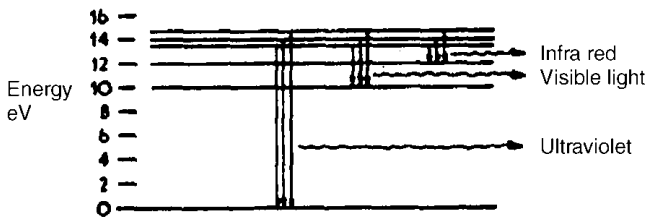
FIGURE 11.17 Energy states of a hydrogen atom.

Where a substance is heated, atoms are continually excited into higher states and falling back. Each transition, from higher state to lower state, and from high states to base state, emits a different wavelength. We perceive the combination of all these as the object glowing. Laser light, however, arises from just one transition—hence, its monochromatic nature. Over the past years, lasers have been emerging from the laboratories to become a useful tool for industry.

Laser applications are based on a laser beam's monochromaticity, its tight beam (a laser beam spreads by less than 0.001 radians) and its high concentration of power (because a laser beam is very narrow, the energy is concentrated on a small area). Typical uses include accurate level-setting in civil engineering, and precision cutting. Precision distance measurement can be made by interferometry techniques because of the monochromatic nature of laser light. Distances from a few millimeters to thousands of kilometers can be measured to a high degree of accuracy. A "typical" CO_2 laser configuration is given in Fig. 11.18.

These lasers, available in output powers ranging from approximately 1,000 to 15,000 W, have been used to demonstrate specific welding accomplishments in a variety of metals and alloys. In addition, several high-power laser systems have been designed, produced, and delivered for actual commercial applications. Substantial advances in laser technology made possible the production of fully automated multikilowatt industrial laser systems that can be operated on a continuous production basis. These systems can be used for a variety of development programs and on-line production applications.

Lasers can be classified according to different schemes. A common system is to use the form and substance of the laser: gas, solid state. Another method of classifying lasers is according to the safety standards. The laser device may be both a fire hazard and hazard to eyes and skin from both direct and specular reflected beams, and sometimes also from a diffuse reflected beam.
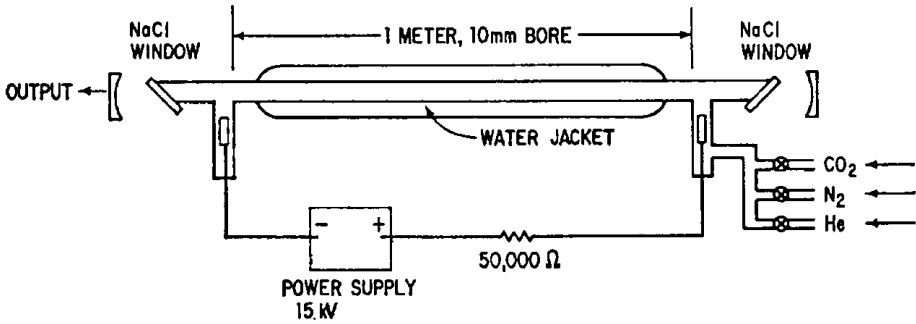
FIGURE 11.18 Schematic diagram of a CO_2 laser.

FIBER OPTICS

The *optical fiber* is a light guide. It is applied for either image or data transmission. It works on the following optic principles.

When a light beam passes from a less dense medium (such as air) to a more dense medium (such as glass) it is bent towards the vertical, as shown before in Fig. 11.19. This effect is known as refraction. Light passing to a less dense medium (e.g. from glass to air) is also bent, as in Fig. 11.20, but as the angle increases total reflection takes place beyond a certain critical angle. For a glass-to-air transition the conical angle is about 40 degrees.
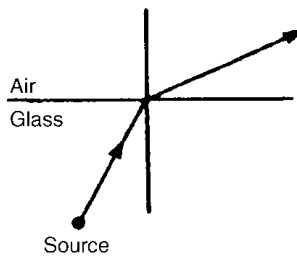
FIGURE 11.19 Refraction of light.
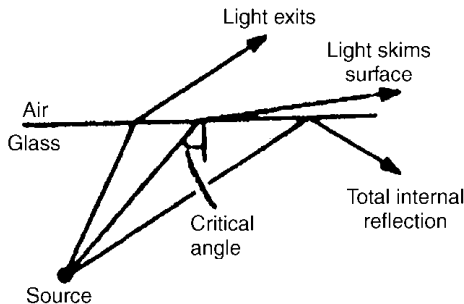
FIGURE 11.20 Critical angle and total internal reflection.

In Fig. 11.21, light is entering a glass rod. As the beam passes down the rod it strikes the edges, but because the angle of incidence is larger than the critical angle, internal reflection occurs and the light beam is conveyed without loss (although there is some attenuation caused by scattering off inevitable flaws in the glass). This is the simple basis of data transmission by light signals. All that is now required is a modulated light source, a transparent conductor arranged to provide internal reflection and a light-sensitive receiver.

In practice, very small optical fibers (of glass or polymer material) are used in place of the glass rod. This gives a flexible 'cable' and lower losses than the simple arrangement of Fig. 11.21. The technique is called fiber optic transmission. Two types of fiber are commonly used: *step index,* which operates as in Fig. 11.21 with reflection taking place at a boundary; and *graded index,* where density varies in a uniform manner across the fiber, giving a gentler reflection, as in Fig. 11.22. Graded index fiber has a lower transmission loss but is more expensive to manufacture.

Light-emitting diodes are usually employed as transmitters, although low powered lasers are often employed on long distance links. Photo avalanche diodes are commonly used as receivers. Digital encoding and transmission techniques are used. Elements of a typical system are shown in Fig. 11.23.

There are a number of advantages in using fiber optic transmission. In theory, very large bandwidths are available; up to 10,000 times higher than the highest achievable radio frequency. Fiber optic cables are physically much smaller than conventional low loss coaxial cables. There is no electromagnetic interference from the fibers, and the signal is unaffected by external interference. Finally, if a fiber cable is damaged or broken there is no risk of fire or sparking. This latter charac-
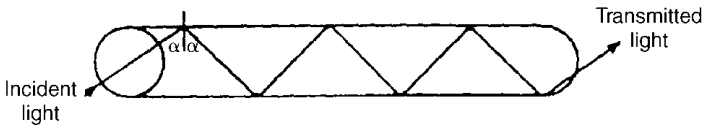
FIGURE 11.21   Fiber optic rod (step index).
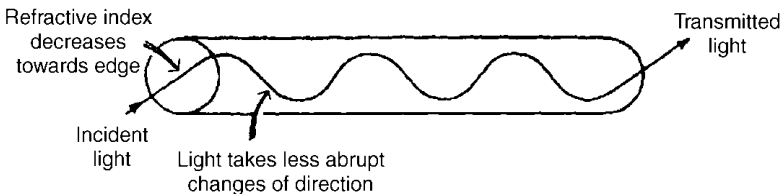
FIGURE 11.22   Graded index fiber optic cable.
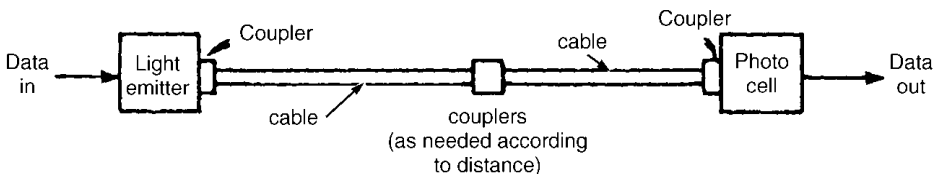
FIGURE 11.23   Fiber optic elements of data transmission.

teristic makes fiber optic cable particularly attractive for data transmission through hazardous areas in petrochemical plants and similar sites.

For transmission over any distance the attenuation or loss of the light intensity is also an important property, and is caused principally by light absorption and scattering. Every material has some fundamental absorption due to its constituents. In addition, the presence of impurities can cause the absorption of light at specific wavelength. Fluctuations in a material on a molecular scale cause intrinsic scattering of light, while fiber core diameter variations or the presence of defects such as bubbles can also cause scattering light loss. Because high-silica-content glasses as materials have very low intrinsic absorption and scattering at the near-infrared wave lengths where light sources (lasers and light-emitting diodes) operate, fibers made from such glasses are most suitable for transmission applications requiring very low attenuation. Very specialized glassmaking techniques, however, are necessary in order to eliminate light loss due to defects or to impurities such as iron and water. Plastics, whose inherent attenuation is higher, can be used for many image transfer applications.

Many other fiber properties, such as strength, dimensional control, and chemical durability, are also important and dictated by the particular application.

NOMENCLATURE

Symbol = Definition, SI Units (*U. S. Customary Units*)

A = area, m^2 (ft^2)

B, **B** = magnetic flux density (magnitude and vector), T

C = capacitance, F

E, **E** = electric field (magnitude and vector), N/C

E = energy (in ''Ohm's law'' and ''Inductance'' sections), J (Btu)

e = electric charge, C

F, **F** = force (magnitude and vector), N (lb_f)

F = tension (in ''Waves'' section), N/m^2 (lb_f/ft^2)

I = electric current, A

i = alternating electric current, A

J = electric current density, A/m^2

K_e = kinetic energy, N · m (lb_f · ft)

L, **L** = length (magnitude and vector), m (ft)

L = inductance (in ''Inductance'' section), H

P = electric power, W

Q = electric charge, C

q = electric charge, C

q_1 = electric charge, C

q_2 = electric charge, C

R = resistance, Ω

R_0 = internal resistance of a source, Ω

r = distance, m (ft)

r, **r** = displacement (magnitude and vector), m (ft)

S = surface area, m^2 (ft^2)

s, **s** = length (magnitude and vector), m (ft)

T = temperature, K (°R)

t = time, s

U = electric potential energy, N · m (lb_f · ft)

u, **u** = velocity (magnitude and vector), m/s (ft/s)

v = voltage, V; velocity (in "Waves" section), m/s (ft/s)

Greek

θ = angle, rad, deg.

ψ = electric flux, N · m^2/C

ϕ = electric potential, V

Φ = magnetic flux, Wb

ρ = resistivity, Ω · m; density (in "Waves" section), kg/m^3 (lb/ft^3)

μ_0 = permeability of empty space, H/m

ACKNOWLEDGMENTS

REFERENCES

1. S. P. Parker, *McGraw-Hill Encyclopedia of Physics,* McGraw-Hill, New York, 1983.

2. D. A. Wells and H. S. Slusher, *Physics for Engineering and Science,* McGraw-Hill, New York, 1983.

3. J. F. Mulligan, *Practical Physics,* McGraw-Hill, New York, 1980.

4. E. U. Condon and H. Odishaw, *Handbook of Physics,* 2d ed., McGraw-Hill, New York, 1967.

5. D. E. Gray (Coordinating Editor), *American Institute of Physics Handbook,* 3d ed. McGraw-Hill, 1972.

6. J. Bird, *Newnes Engineering Science Pocket Book,* Newnes, 2001.

7. N. S. Kapany, *Fiber Optics Principles and Applications,* McGraw-Hill, 1967.

8. S. E. Miller and A. G. Chynoweth (eds.), *Optical Fiber Telecommunications,* McGraw-Hill, 1979.

9. T. Carr and K. Brindley, *Newnes Electronics Engineer's Pocket Book,* 2d ed., Newnes, 2000.

10. D. Christiansen, *Electronics Engineer's Handbook,* 4th ed., McGraw-Hill, 1997.

CHAPTER 12
AUTOMATIC CONTROL

INTRODUCTION

The purpose of an automatic control on a system is to produce a desired output when inputs to the system are changed. Inputs are in the form of commands, which the output is expected to follow, and disturbances, which the automatic control is expected to minimize.

The various aspects of automatic control can best be described by use of an example. Consider a process such as that shown in Fig. 12.1. The flowing liquid is to be heated to a desired temperature by steam flowing through heating coils. The temperature of the exit flow is affected by factors (process variables) such as the temperature of the incoming liquid, the flow rate of the liquid, the temperature of steam, the flow rate of steam, heat capacities of the fluids, heat loss from the vessel, and mixer speed.

1. Open- and Closed-Loop Systems*. The system shown in Fig. 12.1 is normally classified as ''open-loop.'' Open-loop control systems are those in which information about the controlled variable (in this case, temperature) is not used to adjust any of the system inputs to compensate for variations in the process variables. The term *open loop* is often encountered in discussions of control systems to indicate that the uncontrolled process dynamics are being studied.

A closed-loop control system implies that the controlled variable is measured and that the result of this measurement is used to manipulate one of the process variables, such as steam flow.

2. Feedback Control. In the closed-loop control system, information about the controlled variable is fed back as the basis for control of a process variable; hence the designation *closed-loop feedback control.* This feedback can be accomplished by a human operator (manual control) or by use of instruments (automatic control).

For manual control, referring to Fig. 12.1, an operator periodically measures the temperature; if this temperature, for example, is below the desired value, the operator increases the steam flow by opening the valve slightly. For automatic control, a temperature-sensitive device is used to produce a signal (electrical, pneumatic,

---

*Sections 1 to 3 taken in part from *Perry's Chetnical Engineers' Handbook,* 6th ed., by R. H. Perry, D. W. Green, and J. O. Maloney (eds.). Copyright © 1984. Used by permission of McGraw-Hill, Inc. All rights reserved. Updated by the Editors in 2002.
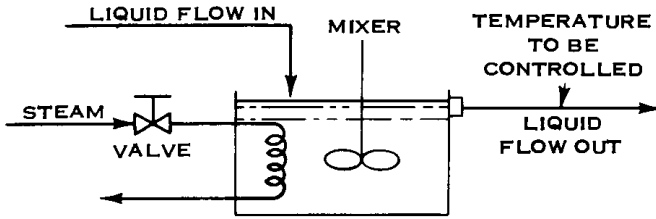
FIGURE 12.1   Simple heat-exchange process. (*From Perry et al.*[1])

etc.) proportional to the measured temperature. This signal is fed to a controller, which compares it with a preset desired value, or set point. If a difference exists, the controller changes the opening of the steam-control valve to correct the temperature, as in Fig. 12.2.

3. Feedforward Control.   Feedforward control is becoming widely used. Process disturbances are measured and compensated for without waiting for a change in the controlled variable to indicate that a disturbance has occurred. Feedforward control is also useful when the final controlled variable cannot be measured.

   The general theories and definitions of automatic control have been developed to aid the designer to meet primarily three basic specifications for the performance of the control system, namely, stability, accuracy, and speed of response.

4. Nomenclature*.   The nomenclature of automatic control has been reviewed by both the American Society of Mechanical Engineers (ASME) and the American Institute of Electrical Engineers (AIEE) in recent years. Many terms adopted by the ASME naturally tend somewhat toward the vocabulary of the process control
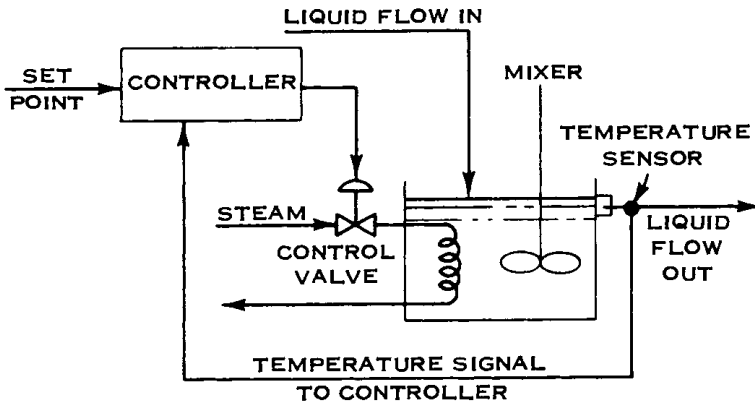
FIGURE 12.2   Automatic feedback control of heat-exchange process. (*From Perry et al.*[1])

engineer and in many cases are not sufficiently broad for general control application. The following terms and definitions have been selected to assist the reader and to serve as reference to a complex area of technology whose breadth crosses several professional disciplines.

Automatic Regulator. An apparatus which measures the value of a quantity or condition which is subject to change with time, and operates to maintain within limits this measured value

Controlled Variable (*Types*). That quantity or condition of the controlled system which is directly measured or controlled

Throttling Range. That range of values through which the variable must change to cause the final control element to move from one extreme position to the other

Set Point. The value of the controlled variable that it is desired to maintain

Deviation. The difference at any time between the controlled variable and the set point

Corrective Action. A change in the flow of the control agent initiated by the measuring means of the automatic controller

Control Agent. Process energy whose flow is directly varied by the control element

Self-Regulation. That operating characteristic which inherently assists the establishment of equilibrium

Two-Position Controller Action. That in which the final control element is moved immediately, from one extreme to the other of its stroke, at predetermined values of the variable

Proportional-Position Controller Action. That in which there is a continuous linear relation between the values of the controlled variable and the rate of motion of a final control element

Floating Controller Action. That in which there is a predetermined relation between the values of the controlled variable and the rate of motion of a final control element

Proportional-Speed Floating Controller Action. That in which there is a continuous linear relation between the rate of motion of the final control element and the deviation of the controlled variable

Derivate Controller Action. That in which there is a predetermined relation between a derivative function of the controlled variable and the position of a final control element

Proportional Plus Floating Controller Action (*Proportional + Reset*). That in which proportional-position and proportional-speed floating actions are additively combined

Proportional Band. The range of scale values through which the controlled variable must pass in order that the final control element be moved through its entire range

Floating Speed. The rate of movement of a final control element corresponding to a specified deviation

Command Signal. The input which is established or varied by some means external to, and independent of, the feedback control system under consideration

Disturbance. A signal which tends to affect the value of the controlled variable

Response Time. The time required for the controlled variable to reach a specified value after the application of a step input of disturbance

Peak Time. The time required for the controlled variable to reach its first maximum following the application of a step input

Rise Time. The time required for the controlled variable to increase from one specified percentage of the final value to another, following the application of a step input

Settling Time. The time required for the absolute value of the difference be-
tween the controlled variable and its final value to become and remain less than a
specified amount following the application of a step disturbance

Compensation. A method of changing or maintaining the state of a system by
employing means to offset the effects of disturbances without causal relationship
between the error in the state of the system and the action of the compensating
means.

BASIC AUTOMATIC-CONTROL SYSTEM*

The general components of a basic automatic-control system are shown in Fig.
12.3. Each block in the diagram represents a function which must be performed by
the control. The operation may be explained as follows: (1) A command signal θ_i
is applied to the input and compared with the instantaneous position of the output
θ_o. (2) The result of this comparison ϵ, representing an error, is amplified by a
controller and used to position a power element. (3) The power device in turn
further amplifies the error signal to supply large amounts of power to the output or
load to reduce the difference between θ_i, and θ_o.

ANALYSIS OF CONTROL SYSTEM

The stability, accuracy, and speed of response of a control system are determined
by analyzing the steady-state and the transient performance. It is desirable to
achieve the steady state in the shortest possible time, while maintaining the output
within specified limits. Steady-state performance is evaluated in terms of the ac-
curacy with which the output is controlled for a specified input. The transient per-
formance, i.e., the behavior of the output variable as the system changes from one
steady-state condition to another, is evaluated in terms of such quantities as maxi-
mum overshoot, rise time, and response time (Fig. 12.4).

An automatic control normally has only two places where disturbances can be
expected: at the input or at the load. For a purely mechanical system the input
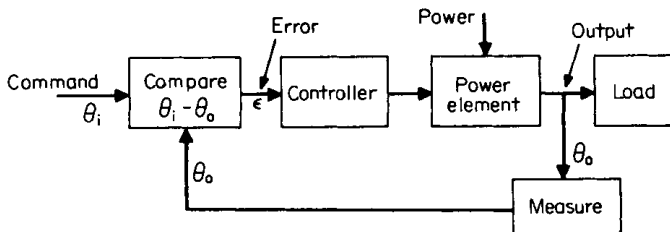
FIGURE 12.3 Functional diagram of an automatic-control system.
(*From Baumeister et al.*[3])
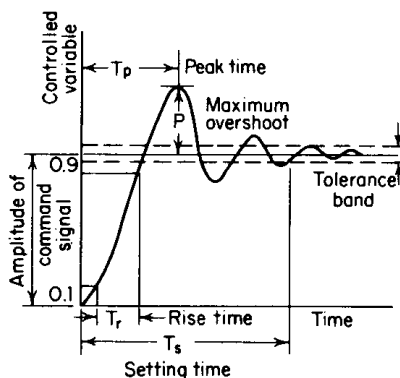
FIGURE 12.4 System response to a unit step-function command. (*From Avallone and Baumeister.*[2])

disturbance may take the form of a periodic oscillation, a displacement, a velocity, or an acceleration. Disturbances at the output are usually load changes expressed as a torque or force quantity. Nonmechanical systems have disturbances expressed in different quantities; however, they are directly analogous to the mechanical system.

1. Laplace Transforms*.   The steady-state and dynamic behavior of a system can be determined by solving the differential equation representing that system. This maybe a long and tedious task, especially if there are many elements in the system. One technique for solving such differential equations uses the Laplace transformation. Here the problem is stated in terms of a second variable which allows the problem to be solved algebraically. Then, by transformation back to the original independent variable, the solution to the original differential equation is obtained (see Chap. 3).

Laplace transforms, as useful as they are, can be employed only for linear differential equations. Such equations describe a linear system, one in which the rules of superposition apply. That is, if the time response of the system is $y_1(t)$ when a forcing function $x_1(t)$ is applied and the response is $y_2(t)$ when $x_2(t)$ is applied, then if $x_1(t) + x_2(t)$ is applied, the response $y_1(t) + y_2(t)$. A linear differential equation of this type is

$$p_n(t) \frac{d^n y(t)}{dt^n} + p_{n-1}(t) \frac{d^{n-1}y(t)}{dt^{n-1}} + \cdots + p_0(t)y(t) = x(t)$$

where the coefficients $p_i(t)$ are not functions of the dependent variable $y(t)$ or any of its derivatives. Generally, the solution of such an equation with time-varying coefficients is difficult. Most chemical operations are nonlinear over a wide range, but the assumption of linearity over a small region near the operating point can usually be justified (linearization). The coefficients of this linearized differential

---

* This section taken in part from *Perry's Chemical Engineers' Handbook,* 6th ed., by R. H. Perry, D. W. Green, and J. O. Maloney (eds.). Copyright © 1984. Used by permission of McGraw-Hill, Inc. All rights reserved.

equation are usually independent of time. Therefore, for process control the basic process differential equation is usually of the form

$$m_n(t) \frac{d^n y(t)}{dt^n} + m_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \cdots + m_0 y(t) = x(t)$$

which can be solved routinely with Laplace transforms.

2. Transfer Functions*.   To illustrate the convenience of the Laplace transformation, consider a differential equation given by

$$T \frac{dc(t)}{dt} + c(t) = RMu(t) \qquad c(0) = 0 \tag{12.1}$$

where T represents the time constant of the system, and $Mu(t)$ describes a step input to the system of magnitude M. We wish to find the response of the output $c(t)$. The Laplace transforms needed are taken from tables. We have

$$\mathcal{L}\left(\frac{cd(t)}{dt}\right) = sC(s) - c(o^+)$$

$c(0^+) = 0$ since the initial conditions were zero. Thus

$$\mathcal{L}[c(t)] = C(s)$$

$$\mathcal{L}[Mu(t)] = M\left(\frac{1}{s}\right)$$

The Laplace transform of Eq. (12.1) is then

$$TsC(s) + C(s) = R\left(\frac{M}{s}\right)$$

$$C(s) = \frac{RM}{s(Ts + 1)} \tag{12.2}$$

To transform this equation back into the time domain, we look up inverse transform $\mathcal{L}^{-1}$ and find that

$$c(t) = \mathcal{L}^{-1}[C(s)] = RM(1 - e^{-t/T})$$

In a system of many elements the transformed equation corresponding to Eq. (12.2) may be quite complicated. However, it can usually be manipulated into a form for which the inverse transform can be found in a table.

In process-control work Laplace transforms are used to determine responses to disturbances. Steady-state or constant terms will usually drop out of the solutions of the differential equations because initial conditions will usually be assumed to be zero.

The transfer function is defined as the ratio of the Laplace transform of the responding variable (output) to the Laplace transform of the disturbing variable (input). From Eq. (12.2), we get for the transfer function $KG(s)$

$$KG(s) = \frac{\text{output}}{\text{input}} = \frac{C(s)}{M/s} = \frac{R}{Ts + 1} \qquad (12.3)$$

The convention for designating transfer functions in a control diagram is the expression $KG(s)$. Capital letters are used when the functions are in the s domain, and small letters are used in the time domain. $G(s)$ represents the dynamic portion of transfer function, and K is related to the steady-state gain through an element. In the transfer function, Eq. (12.3), $K = R$ and $G(s) = 1/(Ts + 1)$.

It is common practice in drawing block diagrams in the s domain to omit the (s) from the $F(s)$'s. Instead only the capital-letter designation is used to represent the s-domain transforms.

Combining Transfer Functions.   Fluid- and thermal-process systems exhibit many different dynamic characteristics, but many systems may be described by combinations of five transfer functions:

Proportional element: $\qquad\qquad\qquad\qquad\qquad\qquad K$

Capacitance element: $\qquad\qquad\qquad\qquad\qquad\qquad \dfrac{1}{Ts}$

First-order element: $\qquad\qquad\qquad\qquad\qquad\qquad \dfrac{1}{Ts + 1}$

Second-order element: $\qquad\qquad\qquad\qquad\qquad \dfrac{1}{T^2s^2 + 2\xi Ts + 1}$

Dead-time element: $\qquad\qquad\qquad\qquad\qquad\qquad e^{-Ls}$

Transfer functions are important tools in the analysis of control systems. Each block or element of the control system has its own characteristic transfer function. If the s-domain transfer function notation $KG(s)$ is used for each block, the system elements can be combined by algebraic procedures into an overall expression for the entire control system.

3. Block Diagrams*.   A useful representation of the mathematical relationships defining the flow of information and energy through the control system is by means of a block diagram. In the diagram the components of the control system are considered as functional blocks in series and parallel arrangements according to their position in the actual control system. Each component is represented by its transfer function, the ratio of the Laplace transform of the output variable to the input variable with all initial conditions taken as zero.

The block diagram of a single-loop feedback-control system subjected to a command input $R(s)$ and a disturbance $U(s)$ is shown in Fig. 12.5.
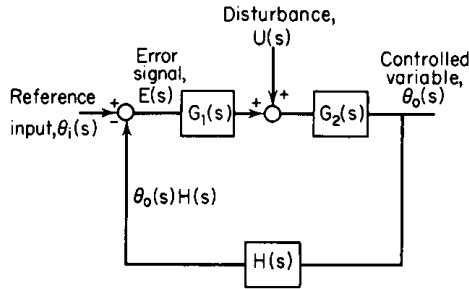
FIGURE 12.5 Single-loop feedback control system. (*From Avallone and Baumeister.*[2])

When $U(s) = 0$ and the input is a reference change, the system may be reduced as follows:

$$E(s) = \theta_i(s) - \theta_o(s) H(s)$$

$$\theta_o(s) = E(s) [G_1(s)G_2(s)]$$

Therefore

$$\frac{\theta_o(s)}{\theta_i(s)} = \frac{G_1(s)G(s)}{1 + G_1(s)G_2(s)H(s)} \qquad (12.4)$$

When $\theta_i(s) = 0$ and the input is a disturbance, the system may be reduced as follows:

$$E(s) = -\theta_o(s)H(s)$$

$$[E(s)G_1(s) + U(s)] G_2(s) = \theta_o(s)$$

$$\frac{\theta_o(s)}{U(s)} = \frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)} \qquad (12.5)$$

Equations (12.4) and (12.5) are in the form

$$\frac{\text{Response function}}{\text{Excitation function}} = \text{system function}$$

The system function is expressible as the ratio of two polynomials, $A(s)/B(s)$. The equation $B(s) = 0$ is the characteristic equation. When the excitation function is specified, inverse transformation of $\theta_o(s)$ yields $\theta_o(t)$, the transient response.

4. Examples*

1. *First-Order Lag Element* (*Time-Constant Element*): A first-order lag element may be illustrated by the liquid level in the tank shown in Fig. 12.6, where the

---

*This section taken in part from *Perry's Chemical Engineers' Handbook,* 6th ed., by R. H. Perry, D. W. Green, and J. O. Maloney (eds.). Copyright © 1984. Used by permission of McGraw-Hill, Inc. All rights reserved.
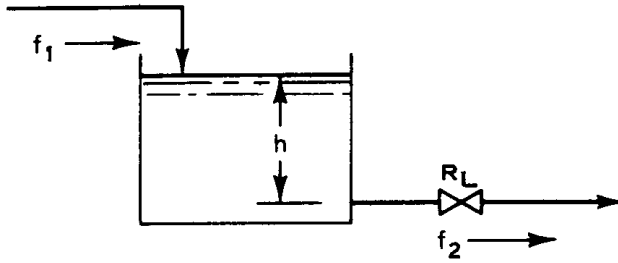
FIGURE 12.6 First-order lag, liquid-flow element. (*From Perry et al.*[1])

flow out of the tank is assumed proportional to the level in the tank. The dynamics are described by

$$C_L \frac{dh}{dt} = f_1 - f_2 \quad \text{and} \quad f_2 = \frac{h}{R_L}$$

Therefore,

$$R_L C_L \frac{dh}{dt} + h = R_L f_1 \tag{12.6}$$

The product RC appears often in analyses of automatic control systems. By analogy from electric-circuit theory, it is called the *time constant* and is usually designated as T. Time constants are probably the most used indicators of process-element characteristics. Substituting $T = R_L C_L$ Eq. (12.6) in the s domain gives

$$(Ts + 1)H(s) = R_L F_1(s)$$

$$H(s) = \frac{R_L}{Ts + 1} F_1(s) \tag{12.7}$$

The transfer function for this first-order lag element is $R_L/(Ts + 1)$; see Fig. 12.7. R_L is a magnitude factor arising here because head or potential was process element output. First-order lag elements are characterized by the expression $1/(Ts + 1)$.

The time response of a first-order lag can be obtained by solving Eq. (12.6) for a step change of inflow with the tank initially empty.

Rearranging terms,

$$\frac{dh}{R_L f_1 - h} = \frac{dt}{T} \quad f_1 = \text{constant}$$

Integrating between the limits $h_0 = 0$ gives
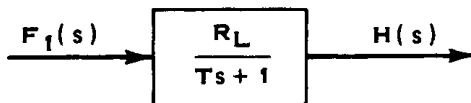
FIGURE 12.7 Block diagram of a liquid-flow, first-order lag process. (*From Perry et al.*[1])

$$- \ln (R_L f_1 - h)|_h^{h_o} = \frac{t}{T}$$

$$h(t) = R_L f_1 (1 - e^{-t/T})$$

A plot of the process-element output for a unit-step increase in input in Fig. 12.8 illustrates that the time constant is the time required to reach $1 - e^{-1}$, 63.2 percent of the final value. The final value of the output will be R, where $R = R_L f_1$.

2. *Second-Order Lag Element* (*Quadratic or Oscillatory Element*): The second-order system is important in the study of automatic control systems because the actual response of a controlled system is often compared with that of an oscillatory second-order system. Few open-loop chemical processes exhibit oscillatory characteristics; however, the closed-loop control system is often tuned to have characteristics similar to the second-order oscillatory system.

Second-order-system characteristics are illustrated by the spring mass, and damper system shown in Fig. 12.9. The mass is acted on by four forces: (1) the spring-displacement force kx, (2) the velocity dependent damping force $p\, dx/dt$, (3) the acceleration-dependent inertial force $(W/g)\,(d^2x/dt^2)$ (g is gravitational acceleration), and (4) the applied force f. The output of this element is the displacement x. The equation relating the forces is
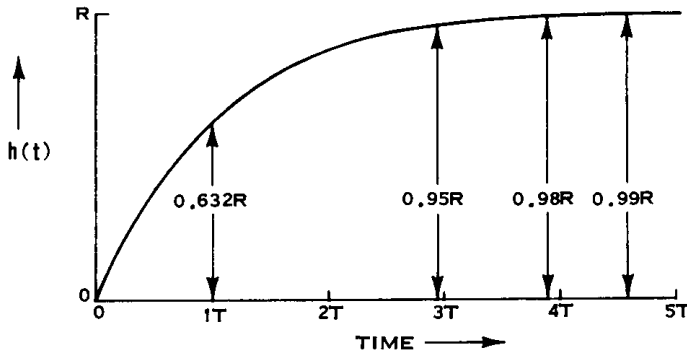
FIGURE 12.8   Time response of a first-order lag resulting from a unit-step increase in input at time equals zero. (*From Perry et al.*[1])
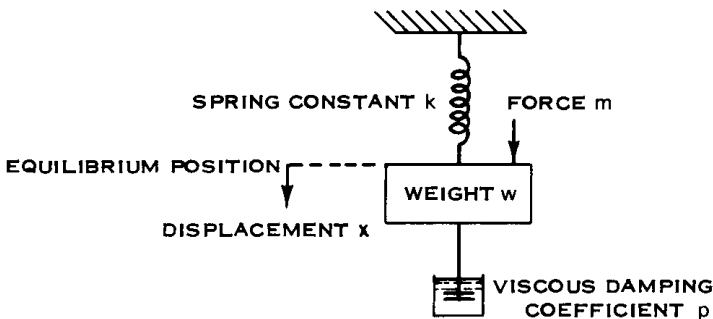
FIGURE 12.9   Mechanical representation of the oscillatory element. (*From Perry et al.*[1])

$$\frac{W}{g}\frac{d^2x}{dt^2} + p\frac{dx}{dt} + kx = f$$

If the initial displacement and velocity are zero, the Laplace transformations are

$$\mathcal{L}[x(t)] = X(s) \qquad \mathcal{L}\left(\frac{dx}{dt}\right) = sX(s) \qquad \mathcal{L}\left(\frac{d^2x}{dt^2}\right) = s^2X(s)$$

After rearranging, the transformed equation is

$$\frac{X(s)}{F(s)} = \frac{1/k}{(W/gk)s^2 + (p/k)s + 1}$$

where $F(s)$ and $X(s)$ are the system input and output, respectively.

This transfer function is characteristic of all quadratic lags. For discussion, it is convenient to write this transfer function in a standard form:

$$KG(s) = \frac{X(s)}{F(s)} = \frac{K}{T_c^2 s^2 + 2\xi T_c s + 1} \qquad (12.8)$$

where K = system gain = $1/k$
T_c = characteristic time = $\sqrt{W/gk}$
ξ = damping factor = $(p/2k)(1/T_c)$

The significance of T_c and ξ can be seen from the response of the system to a unit-step change in the input shown in Fig. 12.10. The response depends on the value
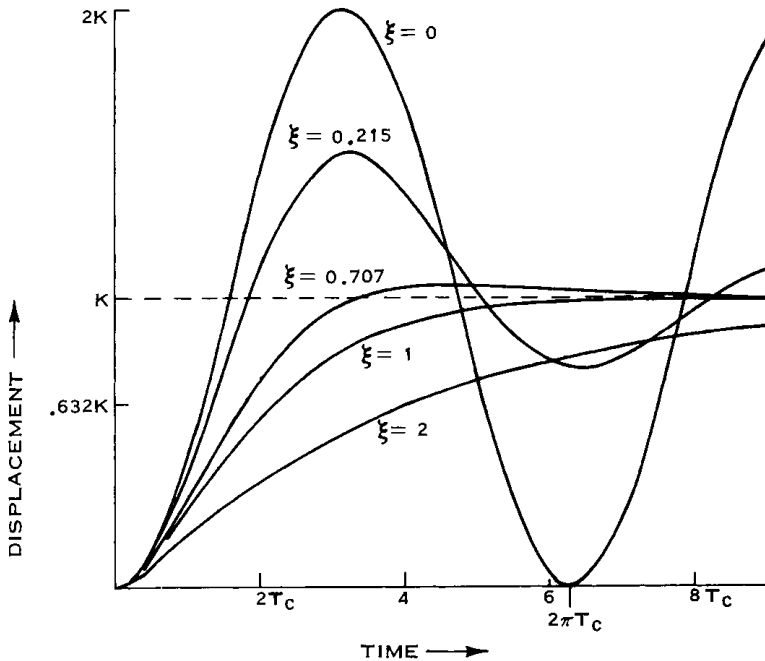
FIGURE 12.10 Time response of an oscillatory, or second-order, system to a unit-step disturbance for various values of damping factor. (*From Perry et al.*[1])

of ξ. For $\xi < 1$ is oscillatory, while for $\xi > 1$ it is nonosciflatory. Systems for which $\xi > 1$ are termed *overdamped;* for $\xi < 1$, they are termed *underdamped;* and for $\xi = 1$, the systems are *critically damped.*

For $\xi < 1$:

$$x(t) = K\left[1 - \frac{1}{\sqrt{1 - \xi^2}}\, e^{-\xi(t/T_c)} \sin(\omega_r t + \phi)\right] \tag{12.9}$$

where $\qquad \omega_r = \dfrac{\sqrt{1 - \xi^2}}{T_c} \qquad$ and $\qquad \phi = \tan^{-1}\dfrac{\sqrt{1 - \xi^2}}{\xi} = \cos^{-1}\xi$

For $\xi = 1$:

$$x(t) = K\left[t - \left(1 + \frac{1}{T_c}\right) e^{-(t/T_c)}\right]$$

For $\xi > 1$:

$$x(t) = K\left[1 - \frac{1}{\sqrt{\xi^2 - 1}}\, e^{\xi(t/T_c)} \sinh(\omega_d t + \phi)\right]$$

where $\qquad \omega_d = \dfrac{\sqrt{\xi^2 - 1}}{T_c} \qquad$ and $\qquad \phi = \cosh^{-1}\xi$

The underdamped response $\xi < 1$ is commonly used for describing control system performance. From Eq. (12.9) some specific relationships can be noted. The frequency of oscillation f is a function of both ξ and T_c:

$$f = \frac{\omega_n}{2\pi} = \frac{\sqrt{1 - \xi^2}}{2\pi T_c}$$

The undamped natural frequency f_n for $\xi = 0$,

$$f_n = \frac{\omega_n}{2\pi} = \frac{1}{2\pi T_c} \tag{12.10}$$

defines the characteristic time T_c.

Note that damping is a property of the system which opposes a change in the output variable.

The immediately apparent features of an observed transient performance are (1) the existence and magnitude of the maximum overshoot, (2) the frequency of the transient oscillation, and (3) the response time.

When an automatic-control system is underdamped, the output variable overshoots its desired steady-state condition and a transient oscillation occurs. The first overshoot is the greatest, and it is the effect of its amplitude which must concern the control designer. The primary considerations for limiting this maximum overshoot are (1) to avoid damage to the process due to excessive excursions of the controlled variable beyond that specified by the command signal, and (2) to avoid the excessive settling time associated with highly underdamped systems. Obviously, exact quantitative limits cannot generally be specified for the magnitude of this overshoot. However, experience indicates that satisfactory performance can generally be obtained if the overshoot is limited to 30 percent or less.

An undamped system oscillates about the final steady-state condition with a frequency of oscillation which should be as high as possible in order to minimize the response time. The designer must, however, avoid resonance conditions where the frequency of the transient oscillation is near the natural frequency of the system or its component parts.

Although these quantities ξ, T_c, ω_n, . . . are defined for a second-order system, they may be useful in the early design states of higher-order systems if the response of the higher-order system is dominated by roots of the characteristic equation near the imaginary axis.

3. *Distance-Velocity Lag* (*Dead-Time Element*): The dead-time element, commonly called *distance-velocity lag* or *true time delay,* is often encountered in process systems. For example, if a temperature-measuring element is located downstream from a heat exchanger, a time delay occurs before the heated fluid leaving the exchanger arrives at the temperature-measurement point. If some element of a system produces a dead time of L time units, then any input $f(t)$ to that element will be reproduced at the output as $f(t - L)$. Transforming to the s domain gives

$$\mathscr{L}[f(t)] = F(s) = \text{input}$$

$$\mathscr{L}[f(t - L)] = e^{-Ls}F(s) = \text{output}$$

and

$$KG(s) = \frac{\text{output}}{\text{input}} = \frac{e^{-Ls}F(s)}{F(s)} = e^{-Ls}$$

Putting this into block-diagram notation gives Fig. 12.11.

FREQUENCY RESPONSE*

Although it is the time response of the control system that is of major importance, study of the effect on transient response of changes in system parameters, either in the process or controller, is more conveniently made from a frequency-response analysis of the system. The frequency response of a system is the steady-state, output of the system to input sinusoids of varying frequency. The output for a linear system can be completely described in terms of the amplitude ratio of the output sinusoid to the input sinusoid to the phase of the output sinusoid to the input sinusoid. The amplitude ratio or gain, and phase, are functions of the frequency of
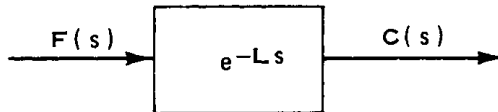
FIGURE 12.11   Block diagram of a dead-time process element. (*From Perry et al.*[1])

the input sinusoid. For purposes of system analysis the frequency response function is more useful than that of the closed loop. Means for obtaining the closed-loop frequency response and evaluating transient performance from the open-loop frequency response are discussed in Ref. 2.

The frequency response can be obtained analytically from the transfer functions of the components, or system, by replacing s, the Laplace operator, with $j\omega$. Table 12.1 shows the frequency responses for some common control-system elements.

The frequency response can also be obtained experimentally for systems not readily amenable to mathematical analysis by subjecting the system to input sinusoids of varying frequency.

STABILITY AND PERFORMANCE OF AN AUTOMATIC CONTROL

An automatic-control system is stable if the amplitude of transient oscillations decreases with time and the system reaches a steady state. The stability of a system can be evaluated by examining the roots of the differential equation describing the system. The presence of positive real roots or complex roots with positive real parts dictates an unstable system. Any stability test utilizing the open-loop transfer function or its plot must utilize this fact as the basis of the test.

1. The Nyquist Stability Criterion*. The $KG(j\omega)$ locus for a typical single-loop automatic-control system plotted for all positive and negative frequencies is shown in Fig. 12.12. The locus for negative values of ω is the mirror image of the positive ω locus in the real axis. To complete the diagram, a semicircle (or full circle if the locus approaches $-\infty$ on the real axis) of infinite radius is assumed to connect in a positive sense, the $+$ locus at $\omega \rightarrow 0$ with the negative locus at $\omega \rightarrow -0$. If this locus is traced in a positive sense from a $\omega \rightarrow \infty$, to $\omega \rightarrow 0$, around the circle at ∞, and then along the negative-frequency locus, the following may be concluded: (1) If the locus does not enclose the $-1 + j0$ point, the system is stable; (2) if the locus does enclose the $-1 + j0$ point, the system is unstable. The Nyquist criterion can also be applied to the log magnitude of $KG(j\omega)$ and phase-vs.-log ω diagrams. In this method of display, the criterion for stability reduces to the requirement that the log magnitude of $KG(j\omega)$ must cross the O-dB axis at a frequency less than the frequency at which the phase curve crosses the $-180°$ line. Two stability conditions are illustrated in Fig. 12.13. The Nyquist criterion not only provides a simple test for the stability of an automatic-control system but also indicates the degree of stability of the system by indicating the degree to which the $KG(j\omega)$ locus avoids the $-1 + j0$ point.

The concepts of phase margin and gain margin are employed to give this quantitative indication of the degree of stability of an automatic-control system. Phase margin is defined as the additional negative phase shift necessary to make the phase angle of the transfer function $-180°$ at the frequency where the magnitude of the $KG(j\omega)$ vector is unity. Physically, phase margin can be interpreted as the amount by which the unity KG vector has to be shifted to make a stable system unstable.

TABLE 12.1 Frequency-Response Equations for Some Common Control-System Elements

Description	Transfer function $G(s)$	Frequency response $G(j\omega)$	Magnitude ratio	Phase angle
1. Dead time	$\epsilon^{-T_L s}$	$\epsilon^{-j\omega T_L}$	1	$-\omega T_L$ radians
2. First-order lag	$\dfrac{1}{Ts+1}$	$\dfrac{1}{j\omega T+1}$	$\dfrac{1}{\sqrt{\omega^2 T^2+1}}$	$-\tan^{-1}(\omega T)$
3. Second-order lag	$\dfrac{1}{(Ts+1)(aTs+1)}$	$\dfrac{1}{-a\omega^2 T^2+j(1+a)\omega T+1}$	$\dfrac{1}{\sqrt{(1-a\omega^2 T^2)^2+(1+a)^2\omega^2 T^2}}$	$-\tan^{-1}\left[\dfrac{(1+a)\omega T}{1-aT^2\omega^2}\right]$
4. Quadratic (underdamped)	$\dfrac{1}{\left(\dfrac{s}{\omega_n}\right)^2+\dfrac{2\zeta}{\omega_n}s+1}$	$\dfrac{1}{-\left(\dfrac{\omega}{\omega_n}\right)^2+j2\zeta\dfrac{\omega}{\omega_n}+1}$	$\dfrac{1}{\sqrt{\left(1-\dfrac{\omega^2}{\omega_n^2}\right)^2+4\zeta^2\left(\dfrac{\omega}{\omega_n}\right)^2}}$	$-\tan^{-1}\dfrac{2\zeta\dfrac{\omega}{\omega_n}}{1-\left(\dfrac{\omega}{\omega_n}\right)^2}$
5. Ideal proportional controller	K	K	K	0
6. Ideal proportional-plus-reset controller $T_i=\dfrac{1}{r}$ r = reset rate	$K\left(1+\dfrac{1}{T_i s}\right)$ or $K\dfrac{T_i s+1}{T_i s}$	$K\left(1+\dfrac{1}{j\omega T_i}\right)$ or $K\dfrac{j\omega T_i+1}{j\omega T_i}$	$K\sqrt{1+\left(\dfrac{1}{\omega T_i}\right)^2}$	$-\tan^{-1}\left(\dfrac{1}{\omega T_i}\right)$
7. Ideal proportional-plus-rate controller	$K(1+T_d s)$	$K(1+j\omega T_d)$	$K\sqrt{1+\omega^2 T_d^2}$	$\tan^{-1}(\omega T_d)$
8. Ideal proportional-plus-reset-plus-rate controller	$K\left(1+T_d s+\dfrac{1}{T_i s}\right)$	$K\left(1+j\omega T_d+\dfrac{1}{j\omega T_i}\right)$ or $K\dfrac{j\omega T_i-\omega^2 T_d T_i+1}{j\omega T_i}$	$\dfrac{K\sqrt{(\omega T_i)^2+(1-\omega^2 T_d T_i)^2}}{\omega T_i}$	$\tan^{-1}\left(\omega T_d-\dfrac{1}{\omega T_i}\right)$

Source: From Avallone and Baumeister.[2] Can also be found in Considine.[4]

12.15
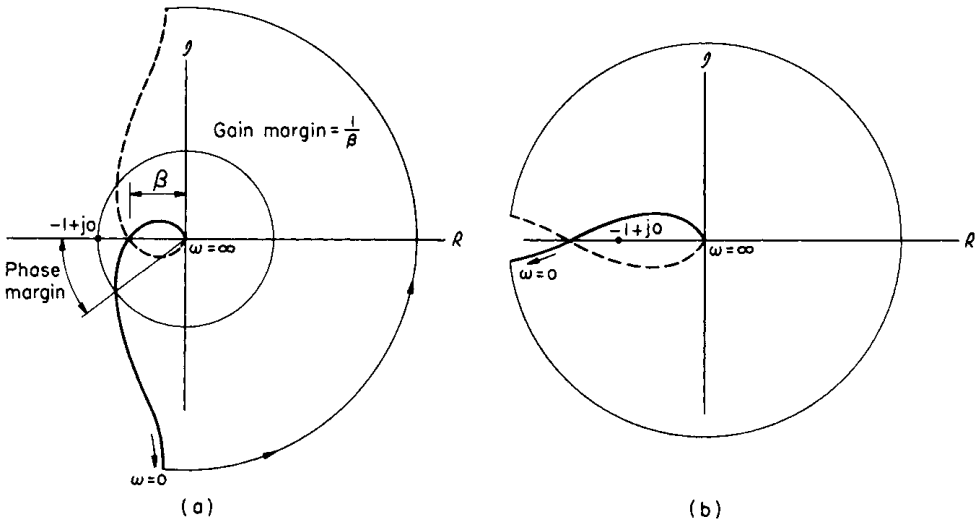
FIGURE 12.12 Typical $KG(j\omega)$ loci illustrating application of Nyquist's stability criterion: (a) stable; (b) unstable. (*From Avallone and Baumeister.*[2])

In a similar manner, gain margin is defined as the reciprocal of the magnitude of the KG vector at $-180°$. Physically, gain margin is the number by which the gain must be multiplied to put the system to the limit of stability. Satisfactory results can be obtained in most control applications if the phase margin is between 40 and 60° while the gain margin is between 3 and 10 (10 to 20 dB). These values will ensure a small transient overshoot with a single cycle in the transient. The margin concepts are qualitatively illustrated in Figs. 12.12 and 12.13.

2. Routh's Stability Criterion. The frequency-response equation of a closed-loop automatic control is

$$\frac{\theta_o}{\theta_i} = \frac{KG(j\omega)}{1 + KG(j\omega)}$$
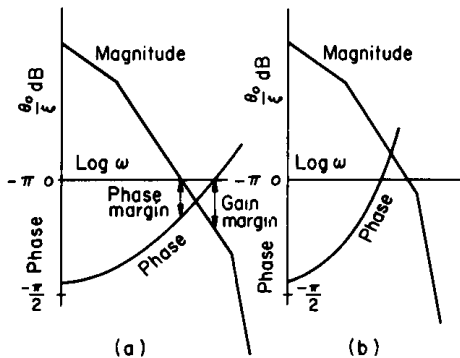
FIGURE 12.13 Nyquist stability criterion in terms of log magnitude $KG(j\omega)$ diagrams: (a) stable; (b) unstable. (*From Avallone and Baumeister.*[2])

The characteristic equation obtained therefrom has the algebraic form

$$A(j\omega)^n + B(j\omega)^{n-1} + C(j\omega)^{n-2} + \cdots = 0$$

The purpose of Routh's method is to determine the existence of roots of this equation which are positive or which are complex with positive real parts and thus identify the resulting instability. To apply the criterion, the coefficients are written alternately in two rows as

$$A\ C\ E\ G$$

$$B\ D\ F\ H$$

This array is then expanded to

$$A\ C\ E\ G$$

$$B\ D\ F\ H$$

$$\alpha_1\ \alpha_2\ \alpha_3$$

$$\beta_1\ \beta_2\ \beta_3$$

$$\gamma_1\ \gamma_2$$

where α_1, α_2, α_3, β_1, β_2, β_3, γ_1, and γ_2 are computed as

$$\alpha_1 = \frac{BC - AD}{B} \qquad \beta_1 = \frac{D\alpha_1 - B\alpha_2}{\alpha_1}$$

$$\alpha_2 = \frac{BE - AF}{B} \qquad \beta_2 = \frac{F\alpha_1 - B\alpha_3}{\alpha_1}$$

$$\alpha_3 = \frac{BG - AH}{B} \qquad \beta_3 = \frac{H\alpha_1 - B_0}{\alpha_1}$$

When the array has been computed, the left-hand column (A, B, α_1, β_1, γ_1) is examined. If the signs of all the numbers in the left-hand column are the same, there are no positive real roots. If there are changes in sign, the number of positive real roots is equal to the number of changes in sign. It should be recognized that this is a test for instability; the absence of sign changes does not guarantee stability.

3. Examples*

1. A servo system with a loop delay is represented by the block diagram in Fig. 12.14 and the function e^{-sT}. Find the maximum value of T for a stable system.
Find the loop gain first. The loop gain is

$$GH(s) = G(s)H(s) = G_1(s)(G_2(s)G_3(S)e^{-sT}$$

Neglecting the phase shift, we obtain $GH'(s)$:
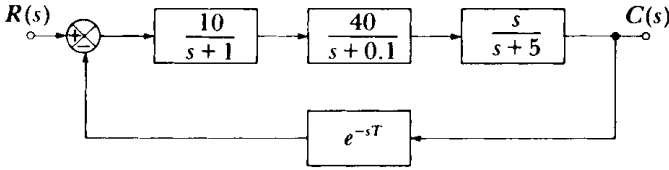
FIGURE 12.14 Block diagram of servo system.

$$GH'(s) = \frac{10}{s + 1} \frac{40}{s + 0.1} \frac{s}{s + 5} = \frac{400s}{(s + 0.1)(s + 1)(s + 5)}$$

$$GH'(j\omega) = \frac{400j\omega}{(0.1 + j\omega)(1 + j\omega)(5 + j\omega)}$$

Plot the function next. For a starting point, find $G(j\omega)$ far from a breakpoint (for accuracy purposes), and plot the function. Pick $\omega = 0.01$ rad/s, and solve for $GH(j\omega)$:

$$GH(j0.01) = \frac{4j}{(0.1)(1)(5)} = 8j$$

$$|GH(j0.01)| = 20 \log_{10} 8 = 18.6 \text{ dB}$$

Now plot the straight-line approximation as shown in Fig. 12.15. (*Note:* One octave separates any two frequencies which are in ratio 2 or ½.)

Determine the phase at the crossover point: It appears as though the plot crosses 0 dB at 20 rad/s. The phase at this crossover point is

$$GH(j20) = \frac{400(j20)}{(0.1 + j20)(1 + j20)(5 + j20)} \approx \frac{400}{(20\underline{/87.1°})(20.6\underline{/78°}}$$

$$GH(j20) = 0.97\underline{/- 163.1°}$$
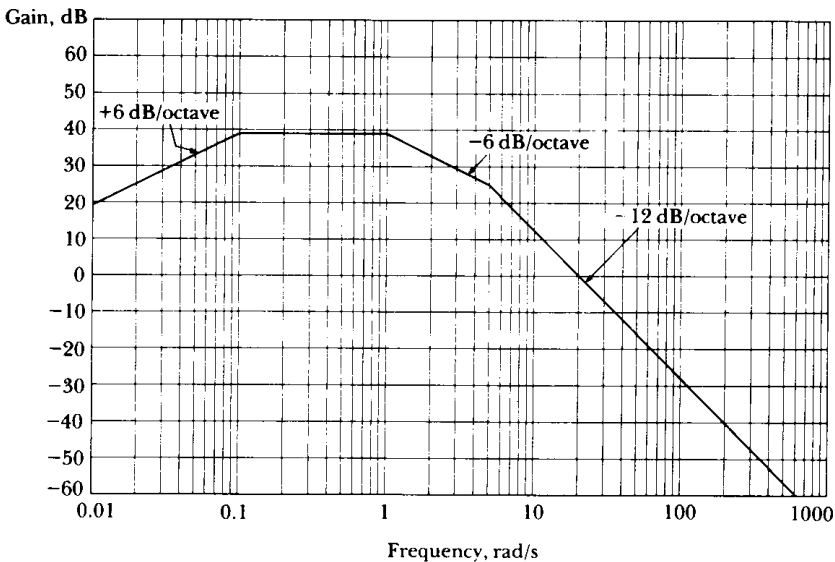
FIGURE 12.15 Bode plot straight-line approximation for $GH(s)$.

The gain checks since it is close to 1. The phase margin is only 16.9° (= 180° − 163.1°) and is already potentially unstable. It will surely be unstable if another 16.9° is added. Therefore, at $\omega = 20$ rad/s

$$\omega T = 16.9° \times \frac{2\pi \text{ rad}}{360°}$$

$$\omega T = 0.294 \text{ rad}$$

$$T = \frac{0.294}{20} = 14.75 \text{ ms}$$

2. The servo system shown in Fig. 12.16 is to be used in an industrial application. Determine the range of K for the system to be stable.

To solve, use the Routh criterion. Using the Routh criterion gives

$$G(s) = G_1(s)G_2(s) = \frac{K}{(s + 2)(s^2 + 4s + 20)} = \frac{K}{D}$$

We set $1 + G(S)H(s) = 0$:

$$1 + \frac{K}{(s + 2)(s^2 + 4s + 20)} \frac{10}{s} = 0$$

$$s(s + 2)(s^2 + 4s + 20) + 10K = 0$$

$$s^4 + 6s^3 + 28s^2 + 40s + 10K = 0$$

Set up the Routh array:

s^4	1	28	$10K$
s^3	6	40	0
s^2	$\dfrac{(6)(28) - (1)(40)}{6} = 21.33$	$\dfrac{60K - 1(0)}{6} = 10K$	0
s^1	$\dfrac{(21.33)40 - 60K}{21.33} = 40 - 2.813K$	0	
s^0	$10K$		

Solve for the range of K:
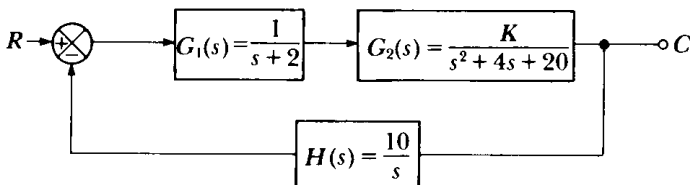
FIGURE 12.16 Servo system block diagram.

$$K = 0 \qquad \text{(for the } s^0 \text{ term)}$$

$$40 - 2.813K = 0 \qquad \text{(for the } s^1 \text{ term)}$$

$$2.813K = 40$$

$$K \geq 14.22$$

Therefore K must lie between 0 and 14.22.

3. An angular-position servo system uses potentiometer feedback and has these gain-transfer functions:

Amplifier:

$$G_1(s) = \frac{100}{(s + 5)} \qquad \frac{V}{V}$$

Motor mechanical transfer function:

$$G_2(s) = \frac{1}{40s} \qquad \frac{\text{rad}}{A}$$

Motor electrical transfer function:

$$G_3(s) = \frac{1}{s + 4} \qquad \frac{A}{V}$$

Potentiometer gain constant:

$$G_4(s) = K \qquad \frac{V}{\text{rad}}$$

Draw a block diagram for this system, and determine the open-loop transfer function and the closed-loop transfer function. If $K = 100$, is the system stable?

Write the open-loop transfer function. The block diagram is drawn as shown in Fig. 12.17. The open-loop transfer function is

$$G(s) = G_1(s)G_2(s)G_3(s) = \frac{100}{40s(s + 5)(s + 4)} = \frac{2.5}{s(s + 5)(s + 4)}$$

Write the closed-loop transfer function. The closed-loop transfer function requirements suggest that the system be put in the canonical form as
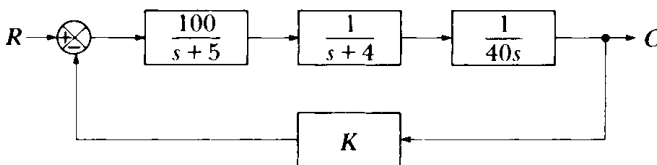
FIGURE 12.17 Angular-position servo system.

$$\frac{C}{R} = \frac{G(s)}{1 + G(s)H(s)} = \frac{2.5}{(s)(s + 5)(s + 4) + 2.5K}$$

Determine whether the system is stable. The characteristic equation for this system is

$$(s)(s + 5)(s + 4) + 2.5K = 0$$

$$s^3 + 9s^2 + 20s + 2.5K = 0$$

Set up a Routh array to find the gain range of K for a stable system:

s^3	1	20	0
s^2	9	2.5K	0
s^1	$\dfrac{180 - 2.5K}{9}$	0	
s^0	2.5K	0	

Setting the s^1 and s^0 terms in the first column ≥ 0, solve for the range of K that makes the system stable. Thus,

For $180 - 2.5K \geq 0$: $\qquad\qquad$ K $\leq$ 72

For $2.5K \geq 0$: $\qquad\qquad\qquad$ $K \geq 0$

So the range of K for a stable system is $0 \leq K \leq 72$ V/rad. So the system is not stable for $K = 100$.

SAMPLED-DATA CONTROL SYSTEMS*

Definition. Sampled-data control systems are those in which continuous information is transformed at one or more points of the control system into a series of pulses. This transformation may be performed intentionally, e.g., the flow of information over long distances is transformed to preserve the accuracy of the data during the transmission, or it may be inherent in the generation of the information flow, e.g., radiating energy from a radar antenna which is in the form of a train of pulses, or the signals developed by a digital computer during a direct digital control of machine-tool operation.

Methods of analysis analogous to those for continuous-data systems have been developed for the sampled-data systems.

Sampling. The ideal sampler is a simple switch (Fig. 12.18), which is closed only instantaneously and opens and closes at a constant frequency. The switch, which may or may not be a physical component in a sampled-data feedback system,

---
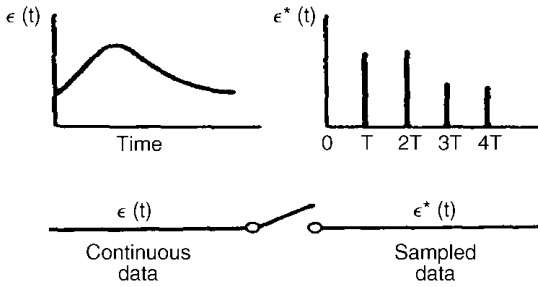
FIGURE 12.18 Ideal sampler, showing continuous input and sampled output.

indicates a sampled signal. Such a sampled-data feedback system is shown in Fig. 12.19. The error signal in continuous form is $\epsilon(t)$, and the sampled error signal is $\epsilon^*(t)$. Fig. 12.18 shows the relationship between these signals in a graphical form.

The stability of sampled-data systems can be demonstrated utilizing frequency-response methods, which have been discussed in previous sections.

Transformation.   In the analysis of continuous-data systems, it has been shown, that the Laplace transformation can be used to reduce ordinary differential equations to algebraic equations. For sampled-data systems, an operational calculus, the z transform, can be used to simplify the analysis of such systems (see Ref. 2).

STATE FUNCTIONS CONCEPT IN CONTROL*

State-space methods permit system analysis and design by study of a set of first-order differential equations rather than a single higher-order equation. This is convenient for solution by numerical methods, using the digital computer, and especially useful for systems with nonlinearities, time-varying characteristics, and multiple inputs and outputs.

For a system previously described as
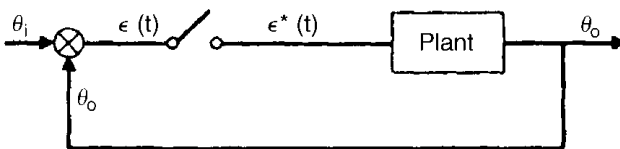
FIGURE 12.19 Sampled-data feedback system.

$$A_n \frac{d^n c}{dt^n} + A_{n-1} \frac{d^{n-1} c}{dt^{n-1}} + \cdots + A_0 c = u(t)$$

where $c(t)$ is the output and $u(t)$ the input, allowing $x_1 = c$, $x_2 = \dot{c}$, $x_3 = \ddot{c}$, ... , $x_n = c^{(n-1)}$ yields the state-variable representation

$$\dot{\mathbf{x}} = \mathbf{Px} + \mathbf{Bu}$$

$$\mathbf{c} = \mathbf{Lx}$$

where $\mathbf{x}$ is state vector and

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1} \\ x_n \end{bmatrix} \qquad \mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -Z_0/A_n & -A_1/A_n & -A_2/A_n & \cdots & A_{n-1}/A_n \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1/A_n \end{bmatrix} \qquad \mathbf{L} = [1, 0, 0, \ldots, 0]$$

The vector $\mathbf{P}$ is called the *companion matrix.*

Transition Matrix. The transition matrix relates the transition of the system state at time $t_0 = 0$ to the state at some later time t.

From $\dot{\mathbf{x}} = \mathbf{Px} + \mathbf{Bu}$ can be derived

$$\mathbf{x}(t) = \phi(t)\mathbf{x}^{(0+)} + \int_0^1 \phi(t - \tau) \mathbf{Bu}(\tau) \, d\tau$$

which is called the *state transition equation of the system.*
$\phi(t)$ is the transition matrix, calculable from

$$\phi(t) = \mathcal{L}^{-1} [(\mathbf{sI} - \mathbf{P})^{-1}]$$

$\phi(t)$ has the properties

$$\phi(0) = \mathbf{I}$$

$$\phi(t_2 - t_0) = \phi(t_2 - t_1)\phi(t_1 - t_0)$$

$$\phi(t + \tau) = \phi(t)\phi(\tau)$$

$$\phi^{-1}(t) = \phi(- t)$$

Since $\mathbf{c}(t) = \mathbf{Lx}(t)$, the following is the system output in terms of the transition matrix:

$$\mathbf{c}(t) = \mathbf{L}\boldsymbol{\phi}(t)\mathbf{x}(0^+) + \int_0^t \mathbf{L}\boldsymbol{\phi}(t - \tau)\mathbf{B}\mathbf{u}(\tau) \, d\tau$$

See Ref. 6 for digital-computer methods for determining $\boldsymbol{\phi}(t)$, given **P.** *Note:* Ref. 6 has complete coverage of the state functions concept in control engineering.

MODELING OF PHYSICAL SYSTEMS

The behavior of real physical systems which engineers must design, analyze, and understand is controlled by the flow, storage, and interchange of various forms of energy. In almost all cases, real systems are extremely complex and may involve several interacting energy phenomena.

The analysis of a dynamic system always involves the formulation of a conceptual model made up of basic building blocks (lumped model elements) that are idealizations of the essential physical phenomena occurring in real systems. An adequate conceptual model of a particular physical device or system will behave approximately like the real system. The best system model is the simplest one which yields the information necessary for engineering decision making.

In modeling system elements it is convenient to determine certain functional relationships between variables of a system which are measurable characteristics of a system and may change with time. A lumped system element is usually described by a relationship between two physical variables, a *through-variable,* which has the same value at the two terminals or ends of the element, and an *across-variable,* which is specified in terms of a relative value or difference between the terminals.

Force, torque, current, fluid flow, and heat flux are through-variables. Velocity, angular velocity, voltage, pressure difference, and temperature difference are across-variables.

A convenient symbol for the relationship between through and across variables is the *linear graph* (a line segment). The two ends or terminals of this graph indicate the across-variables for the element, and the line between these terminals represents the continuity of the through-variables in the element. Linear graphs for ideal system elements are shown in Table 12.2. This table includes also the constitutive relationships and ideal elemental equations for the system elements. These system elements are for the purpose of notation classified as:

A type = the mass, inertia, and capacitance store energy by virtue of their across-variables—velocity and voltage

T type = springs and inductance store energy by virtue of their through-variables

D type = dampers and resistance dissipate energy elements
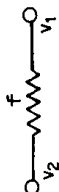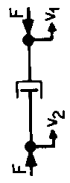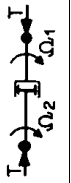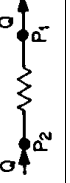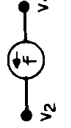
Equations for the automatic control analysis of a system made up of several system elements can be derived by combining linear graphs of elements and using their constitutive relationships (see Refs. 6 and 7).

An illustrative example of step-by-step modeling of a rather complex dynamic system (modeling of an automobile suspension system) is shown in Fig. 12.20. Note that all necessary relationships for the elements on Fig. 12.20 are given in Table 12.2.

TABLE 12.2 Summary of Ideal System Elements*

Type of element	Physical element	Linear graph	Diagram	Constitutive relationship	Energy or power function	Ideal elemental equation	Ideal energy or power
T-type energy storage $\varepsilon \geq 0$ Pure: $x_{21} = f(f)$, $\varepsilon = \int_0^f f\,dx_{21}$ Ideal: $x_{21} = Lf$, $\varepsilon = \frac{1}{2}Lf^2$	Translational spring	v_{21}, F, k	F; v_1, x_1; v_2, x_2	$x_{21} = f(F)$	$\varepsilon = \int_0^F F\,dx_{21}$	$v_{21} = \frac{1}{k}\frac{dF}{dt}$	$\varepsilon = \frac{1}{2}\frac{F^2}{k}$
	Rotational spring	Ω_{21}, T, K	T; Ω_1, Θ_1; Ω_2, Θ_2	$\Theta_{21} = f(T)$	$\varepsilon = \int_0^T T\,d\Theta_{21}$	$\Omega_{21} = \frac{1}{K}\frac{dT}{dt}$	$\varepsilon = \frac{1}{2}\frac{T^2}{K}$
	Inductance	v_{21}, i, L	i; v_1, λ_1; v_2, λ_2	$\lambda_{21} = f(i)$	$\varepsilon = \int_0^i i\,d\lambda_{21}$	$v_{21} = L\frac{di}{dt}$	$\varepsilon = \frac{1}{2}Li^2$
	Fluid inertance	P_{21}, Q, I	Q; P_1, Γ_1; P_2, Γ_2	$\Gamma_{21} = f(Q)$	$\varepsilon = \int_0^Q Q\,d\Gamma_{21}$	$P_{21} = I\frac{dQ}{dt}$	$\varepsilon = \frac{1}{2}IQ^2$
A-type energy storage $\varepsilon \geq 0$ Pure: $h = f(v_{21})$, $\varepsilon = \int_0^{v_{21}} v_{21}\,dh$ Ideal: $h = Cv_{21}$, $\varepsilon = \frac{1}{2}Cv_{21}^2$	Translational mass	v_{21}, F, m	F, p; $v_1 = \text{const}$; v_2	$p = f(v_2)$	$\varepsilon = \int_0^{v_2} v_2\,dp$	$F = m\frac{dv_2}{dt}$	$\varepsilon = \frac{1}{2}mv_x^2$
	Inertia	Ω_{21}, T, J	T, h; $\Omega_1 = \text{const}$; Ω_2	$h = f(\Omega_2)$	$\varepsilon = \int_0^{\Omega_2} \Omega_2\,dh$	$T = J\frac{d\Omega_2}{dt}$	$\varepsilon = \frac{1}{2}J\Omega_2^2$
	Electrical capacitance	v_{21}, i, C	i, q; v_1; v_2	$q = f(v_{21})$	$\varepsilon = \int_0^{v_{21}} v_{21}\,dq$	$I = C\frac{dv_{21}}{dt}$	$\varepsilon = \frac{1}{2}Cv_{21}^2$
	Fluid capacitance	P_{21}, Q, C_f	Q, V; $P_1 = \text{const}$; P_2	$V = f(P_2)$	$\varepsilon = \int_0^{P_2} P_2\,dV$	$Q = C_f\frac{dP_2}{dt}$	$\varepsilon = \frac{1}{2}C_f P_2^2$
	Thermal capacitance	θ_{21}, q, C_t	q, H; $\theta_1 = \text{const}$; θ_2	$\mathcal{H} = f(\theta_0)$	$\varepsilon = \int_0^{\theta_2} q\,dt = \mathcal{H}$	$q = C_t\frac{d\theta_2}{dt}$	$\varepsilon = C_t\theta_2$

TABLE 12.2 Summary of Ideal System Elements* (*Continued*)

Type of element	Physical element	Linear graph	Diagram	Constitutive relationship	Energy or power function	Ideal elemental equation	Ideal energy or power
D-type energy dissipators $\mathcal{P} \geq 0$	Translational damper			$F = f(v_{21})$	$\mathcal{P} = Fv_{21}$	$F = bv_{21}$	$\mathcal{P} = bv_{21}^2$
	Rotational damper			$T = f(\Omega_{21})$	$\mathcal{P} = T\Omega_{21}$	$T = B\Omega_{21}$	$\mathcal{P} = B\Omega_{21}^2$
	Electrical resistance			$I = f(v_{21})$	$\mathcal{P} = Iv_{21}$	$I = \dfrac{1}{R} v_{21}$	$\mathcal{P} = \dfrac{1}{R} v_{21}^2$
Pure $f = f(v_{21})$ $\mathcal{P} = v^{21}f(v_{21})$ **Ideal** $f = \dfrac{1}{R} v_{21}$ $\mathcal{P} = \dfrac{1}{R} v_{21}^2$ $= Ff^2$	Fluid resistance			$Q = f(P_{21})$	$\mathcal{P} = QP_{21}$	$Q = \dfrac{1}{R_f} P_{21}$	$\mathcal{P} = \dfrac{1}{R_f} P_{21}^2$
	Thermal resistance			$q = \text{f}(\theta_{21})$	$\mathcal{P} = q$	$q = \dfrac{1}{R_t} \theta_{21}$	$\mathcal{P} = \dfrac{1}{R_t} \theta_{21}$
Energy sources $\mathcal{P} \lessgtr 0$ $\mathcal{E} \lessgtr 0$	A-type across-variable source			$v_{21} = f(t)$	$\mathcal{P} = fv^{21}$		
	T-type through-variable source			$f = f(t)$	$\mathcal{P} = fv_{21}$		

12.26

FIGURE 12.20 Successive models of automobile and suspension system. (*a*) Automobile and suspension system; (*b*) lumped-mass model; (*c*) mass, spring, and damper model; (*d*) model including tire stiffness and wheel and axle mass; (*e*) two-dimensional model. (*From Shearer et al.*[7])

GENERAL DESIGN PROCEDURE*

The initial performance specifications for an automatic control generally prescribe such quantities as the range of operation of the input variable and its derivatives, the maximum acceptable value of the steady-state error, and possibly other quantities, such as maximum settling time and peak overshoot. With preliminary knowledge of the nature of the input variable and the load, the designer integrates the components of the basic automatic-control system, develops the open-loop transfer function of this basic system, and examines its $G(s)$ locus. The system gain K is then adjusted to satisfy the steady-state error requirements and the resulting locus $KG(s)$ is again examined for stability. If instability exists at the required gain, the $KG(s)$ locus is reshaped through the use of derivative or integral compensation by means of a phase-lead or phase-lag component to display acceptable phase and gain margins. A detailed discussion of gain adjustment and phase compensation may be found in the references.

---

*This section taken in part from *Marks' Standard Handbook for Mechanical Engineers,* 9th ed., by E. A. Avallone and T. Baumeister III (eds.). Copyright © 1987. Used by permission of McGraw-Hill, Inc. All rights reserved.

COMPUTER CONTROL

Digital computers are being used with increasing frequency in the control of diverse processes. All but a few of the digital applications are supervisory or optimizing in nature; the computer, programmed to a model of the process, accepts measured data from conventional instruments, calculates optimum control settings for conventional controllers, and corrects them automatically. The computer need not be concerned only with optimizing the variables of a process for physical stability and quality but can also be used for economic optimization.

A second use of the digital computer is direct digital control (DDC), in which conventional automatic-control instruments are directly displaced by a special purpose digital computer time-shared among many control loops. The advantages of such a system are higher accuracy, more flexibility in incorporating advanced control techniques, and savings in control room costs because of compactness.

Applications of supervisory control computers are found in the electric utility industry, where they are applied to load frequency control and automatic dispatch as well as closed-loop control; in the steel industry, where they are applied to rolling mills; and in the chemical industry, where they are used for closed-loop process control.

*DATA ACQUISITION FOR SENSORS AND CONTROL SYSTEMS**

The input signals generated by sensors can be fed into an interface board, called an I/0 board. This board can be placed inside a PC-based system.

To create a data acquisition system for sensors and control systems that really meets the engineering requirements, same knowledge of electrical and computer engineering is required. The following key areas are fundamental in understanding the concept of data acquisition for sensors and control systems:

Real-world phenomena

Sensors and actuators

Signal conditioning

Data acquisition for sensors and control hardware

Computer systems

Communication interfaces

Software

Real-World Phenomena

Data acquisition and process control systems measure real-world phenomena, such as temperature, pressure, and flow rate. These phenomena are sensed by sensors,

---

*Section adopted from Ref. 10.

and are then converted into analog signals, which are eventually sent to the computer as digital signals.

Some real-world events, such as contact monitoring and event counting, can be detected and transmitted as digital signals directly. The computer then records and analyzes this digital data to interpret real-world phenomena as useful information.

The real world can also be controlled by devices or equipment operated by analog or digital signals that are generated by the computer (Fig. 12.21).

Sensors and Actuators

A sensor converts a physical phenomenon such as temperature, pressure, level, length, position, or presence or absence, into a voltage, current, frequency, pulses, etc.

For temperature measurements, some of the most common sensors include thermocouples, thermistors, and resistance temperature detectors (RTDs). Other types of sensors include flow sensors, pressure sensors, strain gauges, load cells, and optical sensors.

An actuator is a device that activates process control equipment by using pneumatic, hydraulic, electromechanical, or electronic signals. For example, a valve actuator is used to control fluid rate for opening and closing a valve.

Signal Conditioning

A signal conditioner is a circuit module specifically intended to provide signal scaling, amplification, linearization, cold junction compensation, filtering, attenuation, excitation, common mode rejection, etc. Signal conditioning improves the quality of the sensor signals that will be converted in to digital signals by the PC's data acquisition hardware.

One of the most common functions of signal conditioning is amplification. Amplifying a sensor signal provides an analog-to-digital (A/D) converter with a much stronger signal and thereby increases resolution. To acquire the highest resolution during A/D conversion, the amplified signal should be equal to approximately the maximum input range of the A/D converter.
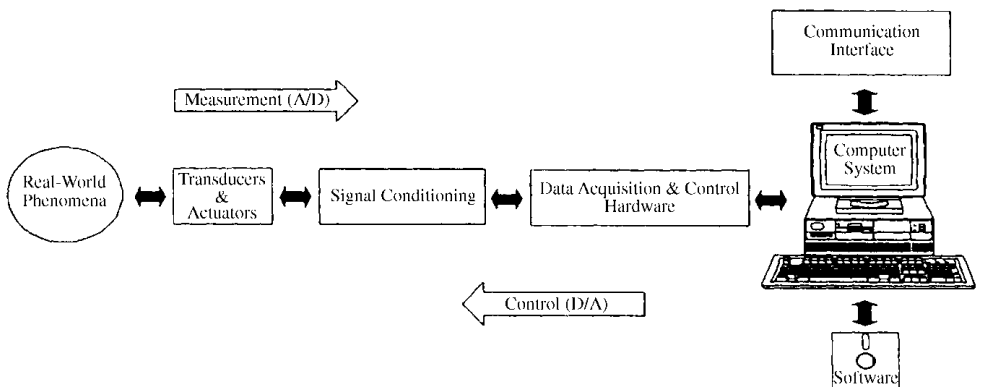
FIGURE 12.21 Integration of computer-controlled devices. (Ref. 10)

Data Acquisition for Sensors and Control Hardware

In general, data acquisition for sensors and control hardware performs one or more of the following functions:

1. Analog input
2. Analog output
3. Digital input
4. Digital output
5. Counter/timer

Analog Input (A/D).   An analog-to-digital converter produces digital output directly proportional to an analog signal input, so that it can be digitally read by the computer (Fig. 12.22.)
   The most significant aspects of selecting A/D hardware are:

1. Number of input channels
2. Single-ended or differential input
3. Sampling rate (in samples per second)
4. Resolution (in bits)
5. Input range (specified as full-scale volts)
6. Noise and nonlinearity

Analog Output (D/A).   A digital-to-analog (DIA) converter changes digital information into a corresponding analog voltage or current. This conversion allows the computer to control real-world events.
   Analog output may directly control equipment in a process that is then measured as an analog input. It is possible to perform a closed loop or proportional integral-differential (PID) control with this function. Analog output can also generate waveforms in a function generator (Fig. 12.23.)

Digital lnput and Output.   Digital input and output are useful in many applications, such as contact closure and switch status monitoring, industrial ON/OFF control, and digital communication (Fig. 12.24.)
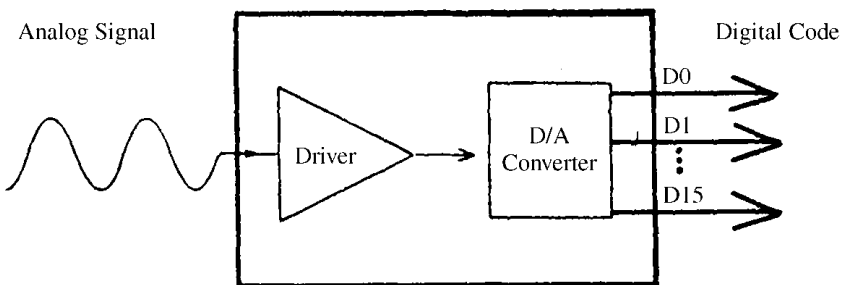
FIGURE 12.22   Analog-to-digital converter.
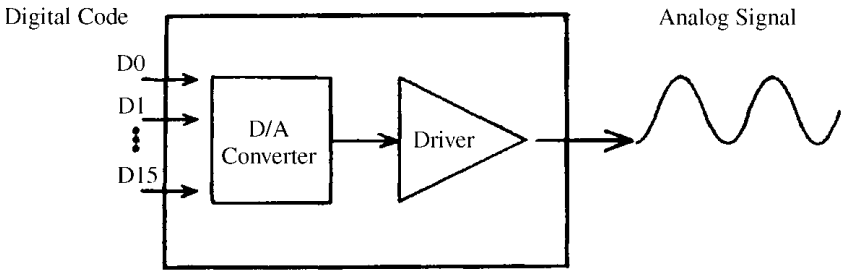
FIGURE 12.23   Analog output.
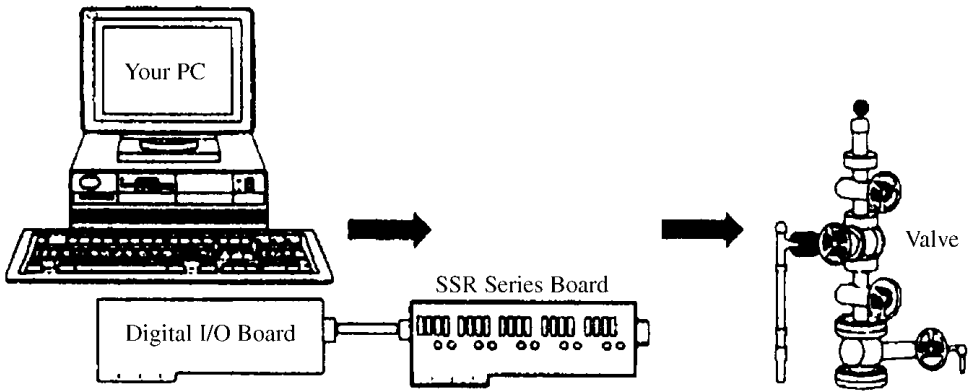
FIGURE 12.24   Application of digital input/output. (Ref. 10)

Counter/Timer.   A counter/timer can be used to perform event counting, flow meter monitoring, frequency counting, pulse width and time period measurement, etc.

Most data acquisition and control hardware is designed with the multiplicity of functions described above on a single card for maximum performance and flexibility. Multifunction data acquisition for high-performance hardware can be obtained through PC boards specially designed by various manufacturers for data acquisition systems.

Computer System

Today's rapidly growing PC market offers a great selection of PC hardware and software.

Hardware Considerations.   Different applications require different system performance levels. Currently, there are Pentium, Pentium Pro, and Pentium II P.IV CPV which will allow a PC to run at benchmark speeds from 300 up to 800 MHz. Measurements and process control applications usually require 80286 systems. But for applications that require high speed, real-time data analysis, an 80386 or 80486 system will be much more suitable.

Industrial PCs. An industrial PC (IPC) is designed specifically to protect the system hardware in harsh operating environments. IPCs are rugged chasses that protect system hardware against excessive heat, dust, moisture, shock, and vibration. Some, IPCs are even equipped with power supplies that can withstand temperatures from -20 to $+85°C$ for added reliability in harsh environments.

Passive Backplane and CPU Card. More and more industrial data acquisition for sensors and control systems are using passive backplane and CPU card configurations. The advantages of these configurations are reduced mean time to repair (MTTR), ease of upgrading the system, and increased PC-bus expansion slot capacity.

A passive backplane allows the user to plug in and unplug a CPU card without the effort of removing an entire motherboard in case of damage or repair.

Communication Interfaces

The most common types of communication interfaces used in the past in PC-based data acquisition for sensor and control system applications are RS-232, RS-422/485, and the IEEE-488 general-purpose interface bus (GPIB).

The RS-232 interface is the most widely used interface in data acquisition for sensors and control systems. However, it is not always suitable for distances longer than 50 m or for multidrop network interfaces. The RS-422 protocol has been designed for long distances (up to 1200 m) and high-speed (usually up to 56,000 bit/s) serial data communication. The RS-485 interface can support multidrop data communication networks. Advancements in these areas are given in References 8 and 10.

Software

The driving force behind any data acquisition for sensors and control systems is its software control. Programming the data acquisition for sensors and control systems can be accomplished by one of three methods:

1. *Hardware-level programming* is used to directly program the data acquisition hardware's data registers. In order to achieve this, the control code values must determine what will be written to the hardware's registers. This requires that the programmer use a language that can write or read data from the data acquisition hardware connected to the PC. Hardware-level programming is complex, and requires significant time-time that might be prohibitive to spend. This is the reason that most manufacturers of data acquisition hardware supply their customers with either driver-level or package-level programs.

2. *Driver-level programming* uses function calls with popular programming languages such as C, PASCAL, C++, Visual C++, and BASIC, thereby simplifying data register programming.

3. *Package-level programming* is the most convenient technique of programming the entire data acquisition system. It integrates data analysis, presentation, and instrument control capabilities in to a single software package. These programs offer a multitude of features, such as pull-down menus and icons, data logging and analysis, and real-time graphic displays.

The introduction of computers increased the complexity of algorithms that can be used to control processes. Furthermore, as processors executing algorithms be-
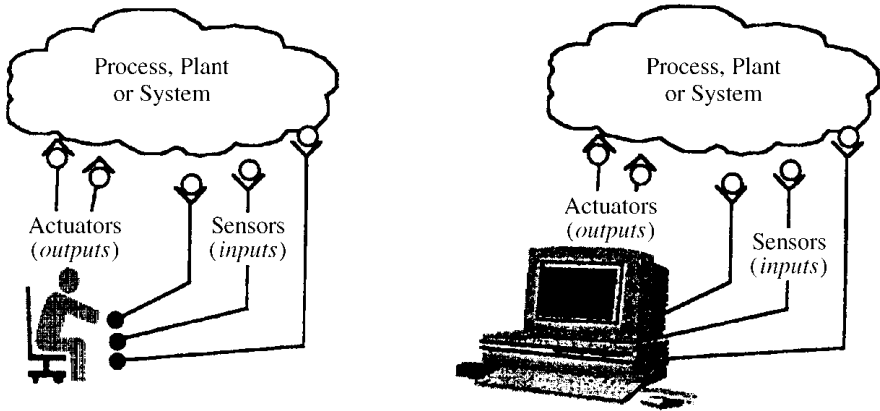
FIGURE 12.25   A centrally located computer eliminates the need for human operator. (Ref. 10)

come a part of sensors and are placed close to the controlled processes, a notion of a truly distributed control system takes on a new meaning. This migration of intelligence toward sensors lead to new opportunities in the design of control systems. At the same time, the proliferation of new applications has created a demand for technologies to lower development costs.

Initially, due to the high cost of electronic components, computers served as hubs of control systems. In these systems, Fig. 12.25, sensors and actuators are directly connected to the central computer as a steady replacement of human operator. Analog sensors are linked to computers using A/D.

NOMENCLATURE

Symbol = Definition, SI units (U.S. Customary units)

C = capacitance, F

C_f = fluid capacitance, m^{5}N (ft^5/lb$_f$)

C_t = thermal capacitance, J/°C (Btu/°F)

E = energy, J (lb$_f$ · ft)

F = force, N (lb$_f$)

f = real variable function

H = heat, J (Btu)

I = fluid inertance, N · s^2/m^5 (lb$_f$ · s^2/ft^5)

i = current, A

L = inductance, W/A

m = mass, kg (lb)

P = power, W (lb$_f$ · ft/s)

p = translational momentum, N · s (lb$_f$ · s)

p = pressure, N/m^2 (lb$_f$/ft^2)

Q = fluid flow rate, m^3/s (ft^3/s)

$\quad q$ = heat flow rate, J/s (Btu/s)

$\quad q$ = charge, C

$\quad R$ = resistance, V/A (or Ω for ohms)

$\quad R_f$ = fluid resistance, $N \cdot s/m^5$ ($lb_f \cdot s/ft^5$)

$\quad R_t$ = thermal resistance, $°C \cdot s/J$ ($°F \cdot s/Btu$)

$\quad s$ = complex variable

$\quad T$ = torque; $N \cdot m$ ($lb_f \cdot ft$)

$\quad t$ = time, s

$\quad V$ = fluid volume displacement, m^3 (ft^3)

$\quad v$ = translational velocity, m/s (ft/s)

$\quad v$ = voltage, V

$\quad x$ = translational displacement, m (ft)

Greek

θ = temperature, °C (°F)

ϕ = angular displacement, rad (deg)

λ = flux linkage, w

Γ = pressure momentum, $N \cdot s$ ($lb_f \cdot s$)

Ω = angular velocity, rad/s

Subscripts

$\quad 1$ = at point 1

$\quad 2$ = at point 2

21 = $2 - 1$ (e.g., $x_{21} = x_2 - x_1$)

Note. Other symbols are defined in the text.

*REFERENCES**

1. R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Perry's Chemical Engineers Handbook,* 6th ed., McGraw-Hill, New York, 1984, Sec. 22.

2. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers,* 9th ed., McGraw-Hill, New York, 1987, Sec. 16.

3. H. A. Rothbart, *Mechanical Design and Systems Handbook,* 2d ed., McGraw-Hill, New York, 1985.

3. T. Baumeister, E. A. Avallone, and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers,* 8th ed., McGraw-Hill, New York, 1987.

*Those references listed above but not cited in the text were used for comparison of different data sources, clarification, clarity of presentation, and, most important, reader's convenience when further interest in the subject exists.

4. D. Considine, *Process Instruments and Controls Handbook,* McGraw-Hill, New York, 1974.

5. T. G. Hicks, *Standard Handbook of Engineering Calculations,* 2d ed., McGraw-Hill, New York, 1985, Sec. 7.

6. D. G. Schultz and J. L. Melsa, *State Functions and Linear Control Systems,* McGraw-Hill, New York, 1967.

7. J. L. Shearer, A. T. Murphy, and H. H. Richardson, *Introduction to System Dynamics,* Addison-Wesley, Reading, Mass., 1971.

8. C. H. Houpis and G. B. Lamont, *Digital Control Systems; Theory, Hardware, Software,* 2d ed., McGraw-Hill, 1992.

9. R. B. Northrop, *Introduction to Instrumentation and Measurements,* CRC Press, 1997.

10. S. Solomon, *Sensors Handbook,* McGraw-Hill, 1999.

CHAPTER 13
MECHANICAL ENGINEERING*

MECHANICAL DESIGN ENGINEERING

PRINCIPLES OF MECHANISM

1. Definitions and Concepts. A *machine* is a combination of bodies which have definite for, and the strength to maintain that form. These bodies or *machine members* are so shaped and connected that they can move upon each other, but only in a particular way. Freedom for relative motion of parts differentiates the machine from the structure; constraint to a definite relative motion makes it more than a random assemblage of connected bodies.

Regularly one member (more extensive than the others) is fixed or basal as the *frame* of the machine, upon or within which the moving members function. Commonly one of these receives an input of driving force and acts as driver; this force, or the work of which it is a factor, is transmitted through the machine, with change in force and velocity and some loss due to friction; and by a final or driven member, at the place of output, force or work is applied against a useful resistance or to a useful effect.

2. Typical Examples. The members of the engine outlined in Fig. 13.1 are frame or bed 1, piston and slide 2, connecting rod 3, and crank and shaft or crank shaft 4. The piston is driver, receiving total gas pressure *P*. Through the connecting rod, by means of its total lengthwise stress *S*, the driving effect is carried over to the crank pin, becoming active in tangential component *T*. The shaft is the output member, turning against whatever load is imposed upon it.

The hoist gearing or hoisting-gear train of Fig. 13.2 starts with motor *M* and ends with hoist drum *H*. It has four turning members, numbered at their shaft axes and connected by the gear pairs, *AB*, *CD*, and *EF*. The effect of this machine is to change by successive steps from the input condition of a small force at high speed to the output condition of a big resistance at low speed.

3. Composition of Machines. A machine must have at least two members, definitely connected and with one fixed. A loose prying lever, such as a crowbar or canthook, is merely a tool or appliance, as is also a roller placed beneath a heavy body to help move it.

Examples of machines having but one moving member are:

A lever jack, for automobiles and trucks

A simple windlass, without gearing

---

*From *Engineering Manual,* 3d ed., by R. H. Perry (ed.). Copyright © 1976. Used by permission of McGraw-Hill, Inc. All rights reserved. Updated and metricated by the editors, 2002.
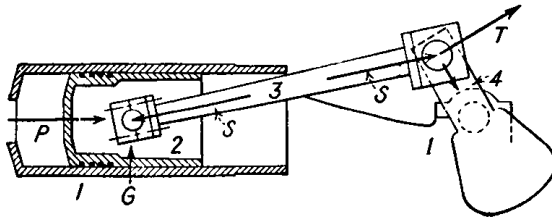
FIGURE 13.1 Outline of engine mechanism.
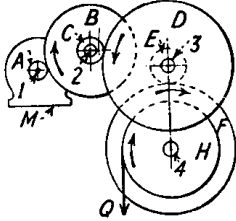
FIGURE 13.2 A hoisting-gear train.
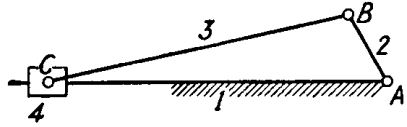
FIGURE 13.3 The engine linkage.

A waterwheel or turbine, a centrifugal pump or blower, an electric motor or generator; and, with greater completeness of function, such a combination as turbine and generator or motor and pump

4. Units of Mechanism. The distinctive meaning of the word *mechanism* is well shown by reducing the engine to its bare geometrical outline (Fig. 13.3); this *kinematic skeleton* shows all of form and dimension that is needed for motion analysis and determination. To this essential form for motion the name *mechanism* is given.

Figures 13.3 and 13.4 are typical *linkage mechanisms,* composed of cranks or arms, rods, and slides. The first has three pin or turning joints and one sliding joint, the second is the four-pin linkage. The regular linkage unit contains four members; but the special wedge mechanism in Fig. 13.5, with only sliding joints, has but three members. The *screw mechanism* (Fig. 13.6) is really a derivative of the wedge unit, substituting long-range for short-range action. Note the interchange of screw-and-nut joint and end-thrust turning joint as between the two cases.

A very common three-member unit, shown in three forms in Fig. 13.7, consists of a pair of gears and the frame carrying their shaft bearings. A solid disk or rim with teeth on the outside is called a *spur gear.* A pair of these, pair 2–3 in Fig. 13.7*a*, has opposite directions of turning. By the use of an *internal gear,* with teeth
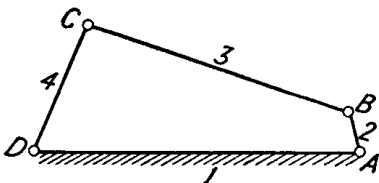
FIGURE 13.4 The four-pin linkage.
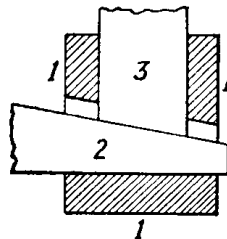
FIGURE 13.5 The wedge mechanism.
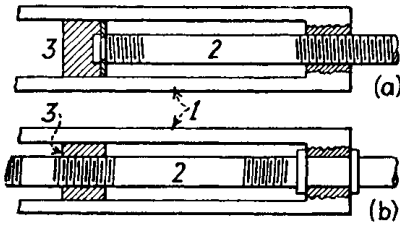
FIGURE 13.6   Screw mechanisms.
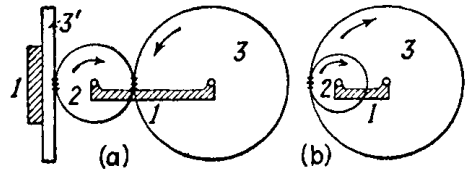
FIGURE 13.7   Gear pairs.

on the inside of the rim as indicated in Fig. 13.7*b*, like directions are obtained. The straight toothed bar 3′ in Fig. 13.7*a* is called a *rack*. A gear rim of large radius, made in segments to be bolted to the base of a gun mount or turret or of a turntable, may be called a *circular rack.*

Another form of wheel unit contains a pair of pulleys or sprocket wheels with a belt or chain as flexible connector (Fig. 13.8). The dotted lines indicate the crossed belt which gives reversal of turning, but the chain can of course drive only directly.

5. The Cam—Principle and Action.   The *eccentric* is a circular disk which can function as an enlarged crank pin (Fig. 13.9) or from which motion can be taken by a line-contact follower as in Figs. 13.10 to 13.12. In the latter use the eccentric has become a *cam,* a machine member which commonly serves to produce some form of reciprocating or oscillating motion.

6. Various Cams.   The powerful shear outlined in Fig. 13.13 is actuated by a cam so profiled from *A* to *B* as to give a slow cutting "stroke" at uniform velocity, followed by a quick return *BC* and a dwell *CA.* This is an unusually heavy service
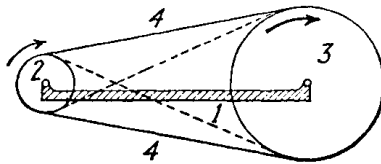
FIGURE 13.8   Belt-and-pulley pairs.
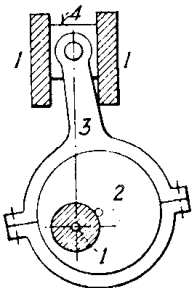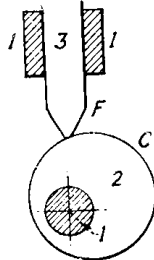
FIGURE 13.9   **FIGURE 13.10**   **FIGURE 13.11**   **FIGURE 13.12**
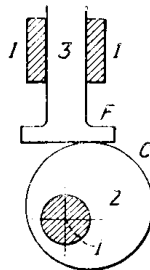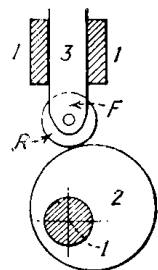
FIGURES 13.9 to 12   Various cams and followers.
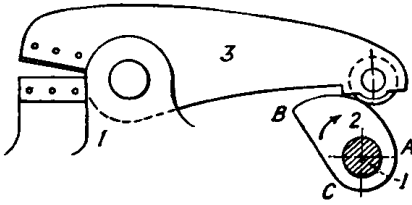
FIGURE 13.13   Heavy cam-driven shear.
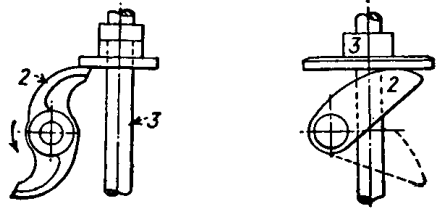
FIGURE 13.14   Trip and oscillating cams.

for a cam drive; but the trip cam in Fig. 13.14 also belongs to a machine of the rough-and-ready type: this is a lift-and-release mechanism regularly used in ore stamps for crushing a hard ore, such as gold-bearing quartz.

7. Positive Cams.   One way to make cam drive positive is to form a slot in which moves a roller or pin on the follower. This is seen in Figs. 13.15 and 13.16, both from machine-tool feed drives and producing the same motion.

8. Classification of Machines.   Any classification of mechanisms that may be made has to do primarily with the internal constitution of the machine, between input and output points. Machines may also be classified as to service or function, or on the basis of external relationship. Important and obvious classes are as follows:

1. Prime movers or power-generating machines, which receive a natural driving force and deliver power to drive other machines; this class includes all kinds of engines and motors, driven by fluid pressure or by electromagnetic attraction.

2. Reversed prime movers, or pumps, compressors, and electric generators, which work against the same forces that drive prime movers.

3. Transmission machinery, between prime movers and working machines. Such an aggregation of shafting, pulleys, belts, etc., is an alternative to electrical transmission with a motor at and for each machine.

4. Transportation machinery of all kinds—locomotives, self-propelled cars, propelling equipment of ships and aircraft.
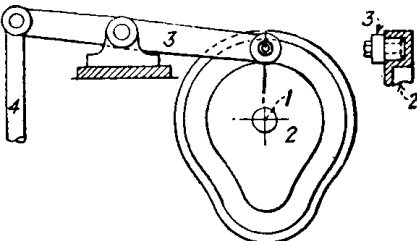
FIGURE 13.15   Positive disk cam.
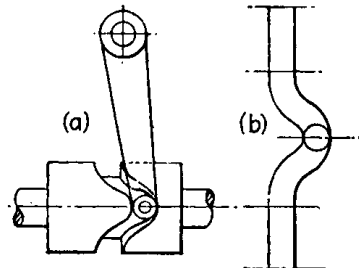
FIGURE 13.16   A drum cam and its profile.

5. Hoisting and conveying machines or machinery, including various types of excavators.

6. Many groups of process machines, which work upon various materials to convert or shape or fabricate them.

FORCE AND WORK RELATIONS

9. Work and Power.   When a force P drives a body through a distance S against an equal and opposite resistance Q it performs the work $P \times S$ or $Q \times S$. With P in lb (N) and S in ft (m) the work is in ft · lb (N · m). If this is done in time t, so that the velocity (constant or mean) is $V = S/t$, the work rate is PV or QV. Velocity V is often that of a point in a circle path, due to the rotation of a radius R at N turns or r/min, so that $V = 2\pi RN$. With V in ft/min, the expressions for power are

$$\text{Power} = \frac{PV}{33,000} = \frac{P \times 2\pi RW}{33,000} = \frac{PRN}{5252.1} \quad \text{hp} \quad (1 \text{ hp} = 746 \text{ W}) \quad (13.1)$$

In the closer study of machines it is convenient to express rate of turning as angular velocity in radians per second. The *radian* is the angle subtended by an arc equal in length to the radius; there are 2π or 6.2832 such angles, of the value 57.295°, in one turn or revolution. With $N/60 = n$ r/s, and ω for angular velocity,

$$\omega = 2\pi n = 2\pi \frac{N}{60} = 0.10472N \qquad (13.2)$$

In Eq. (13.1) substitute ω for $2\pi n$. Then

$$\text{Power} = \frac{PRW}{550} \quad \text{hp} \quad (1 \text{ hp} = 746 \text{ W}) \qquad (13.3)$$

the expression $PR\omega$ giving ft · lb/s (N · m/s) as the product of moment in ft · lb (N · m) and angular velocity in radians/s. Emphasizing moment or torque PR and distinguishing PR' in lb · ft (N · m) and PR'' in lb · in (N · m)

$$PR' = 5252.1 \frac{\text{Power (hp)}}{N} \quad (1 \text{ hp} = 746 \text{ W}) \qquad (13.4)$$

$$PR'' = 63,025 \frac{\text{Power (hp)}}{N} \quad (1 \text{ hp} = 746 \text{ W})$$

10. Pressure Work.   Let a pressure p lb/in² (kPa) act upon a piston of area A in² (m²), the piston making N strokes of length S ft/min (m/min). Then $pA = P$ is total pressure or driving force and NS is velocity V; therefore the power developed is

$$\text{Power} = \frac{pANS}{33,000} \quad \text{hp} \quad (1 \text{ hp} = 746 \text{ W}) \qquad (13.5)$$

This formula shows the method of calculating the power of an engine, pressure p

or p_m being the *mean effective pressure* determined by the indicator diagram. If the crank shaft makes N r/min, this is also the number of working strokes of each piston face in the ordinary double-acting engine.

Consider pressure p on area A, pushing this area through S ft (m). The work done is

$$W = pA \times S = P \times S \tag{13.6}$$

Change the grouping of factors and convert area A to ft² (m²). Then

$$W = 144p \times \frac{AS}{144} = 144p \times (v_2 - v_1) \tag{13.7}$$

in which $(v_2 - v_1)$ is the volume in ft³ (m³) displaced or "swept through" by the piston. This second definition or determination of work makes it the product of pressure in lb/ft² (kPa) by volume change in ft³ (m³).

11. Transmission of Work and Power.   As the simplest type of example—because conditions stay constant at any one place in the mechanism—consider a gear train like Fig. 13.2. The transmission process starts with the input work rate P_iV_i and runs through to the output rate P_oV_o; and at each gear contact there is a change in the two factors, force P increasing as velocity V decreases. Also a progressive decrease occurs in the value of PV, due to machine frictions.

In the case of turning member 1 in Fig. 13.2 there are no definite factors P_i and V_i, but rather a moment or torque PR on the armature of the motor and an angular velocity ω. With M as a single symbol for moment the alternative work rate is $M\omega$.

For the overall effect of a machine let P be the input force and Q the output resistance, with the work rates $U_p = PV_p$ and $U_q = QV_q$. The ideal or frictionless machine would have no internal loss, making

$$PV_p = QV_q \qquad Q_0 = P \times r_v \qquad P_0 = Q \div r_v \tag{13.8}$$

Here r_v is the velocity ratio V_p/V_q, taken from input toward output and expressing the kinematic or motional effect of the mechanism; Q_0 is the ideal load for a given driving force, greater than actual Q; P_0 is the ideal driving force for a given load, less than actual P.

12. Friction and Efficiency.   Let U_f stand for the work rate of friction (work absorption) and e for mechanical efficiency. Then in the actual machine

$$PV_p = QV_q + U_f \qquad e = \frac{QV_q}{PV_p} \tag{13.9}$$

Efficiency is primarily a ratio of works or work rates; but by r_v for V_p/V_q and by Eq. (13.8) it becomes

$$e = \frac{Q}{Pr_v} = \frac{Q}{Q_0} = \frac{P_0}{P} \tag{13.10}$$

making efficiency a ratio of forces

From the first form of Eq. (13.10) come the actual relations,

$$Q = P \times er_v \qquad P = Q \div er_v \qquad (13.11)$$

in contrast with the ideal relations in Eq. (13.8). Velocity ratio r_v is often called the *mechanical advantage* of the mechanism. It is now seen that r_v is an ideal and er_v an actual or realized value of the ratio of Q to P.

13. Hoisting Work. Let the unit of load be one tone of 2000 lb (908 kg) and consider the height through which it can be lifted by one 1 hp (746 W) in 1 min. Without friction this would be

$$h_o = 33,000 \div 2000 = 16.5 \text{ ft } (5.0 \text{ m})$$

With an efficiency ranging from 0.6 to 0.8 the velocity of lift will be from 10 to 13 ft (3.0 to 3.9 m) by 1 input hp/ton (746 W/1016 kg).

14. The Wheel Pair. Consider the pair of gears A and B in Fig. 13.17, turning on fixed axes 1 and 2. Their relative speeds or turning rates are determined by the fact of a common linear, tangential velocity at the place of circle contact, at the ends of radii a and b, so that

$$aN_1 = bN_2 \qquad N_2 = N_1 \frac{a}{b} = \frac{N_1}{r_n} \qquad (13.12)$$

Turning ratio r_n is taken in the same direction as r_v in Sec. 11, from driver to follower and usually from greater to less. Note that speeds N_1 and N_2 are inverse to radii a and b, the smaller wheel running faster, the large one more slowly.

15. Gear-Train Relations. In the gear train outlined in Fig. 13.18 the first thing to be noted is the relative directions of turning. A single pair of ordinary spur gears (outside contact) reverses direction. With two or any even number of such pairs the follower and driver turn alike; with an odd number of pairs there is overall reversal. Second arrangement b in Fig. 13.18 shows the introduction of an idle wheel or idler X into pair AB; this does not affect velocity or force transmission, but merely reverses direction.

For the turning ratio of a train take first the successive single pairs, then combine them, thus,

$$\frac{N_2}{N_2} = \frac{b}{a} \qquad \frac{N_2}{N_3} = \frac{d}{c} \qquad \frac{N_1 N_2}{N_2 N_3} = \frac{N_1}{N_3} = r_n = \frac{bd}{ac} \qquad (13.13)$$

In words, *the turning speed of the first driving wheel or shaft is to the speed of the*
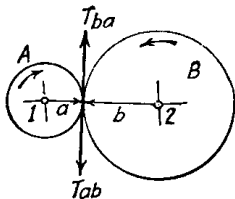
FIGURE 13.17   Gear relations.
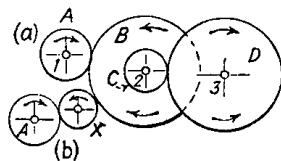
FIGURE 13.18   Directions of turning.

last driven shaft as the continued product of (the radii of) all the followers is to the product of all the drivers.

Figure 13.19 puts the gear train into power-transmission service, by applying the driving moment PR_p to member 1 and the resisting moment QR_q to member 3. The no-friction relation would be

$$PR_p \times r_n = QR_q \tag{13.14}$$

while with friction it is

$$PR_p \times er_n = QR_q \tag{13.15}$$

CONSTRUCTIVE ELEMENTS OF MACHINES

16. Fastenings—Riveted Joints. Fastenings, such as rivets, bolts, screws, keys, and even forced or shrunk fits, are elements of machine structures, whether in the frame or in moving members.

Rivets and riveted joints find little place in actual machines, but in one class of mechanical structure, namely, existing older boilers and pressure tanks, they are very important. The button-head rivet show in Fig. 13.20a is the form most used; but there is also the "pan" head (a truncated cone), and the countersunk head (Fig. 13.20b) which is used where a smooth surface is desired, especially in ship construction.

17. Bolts and Screws. The three typical forms of bolts appear in Figs. 13.21, 13.22, and 13.23, namely:
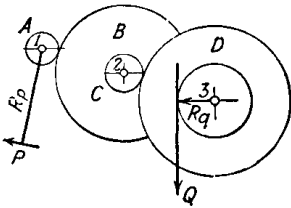
FIGURE 13.19   Force relations.
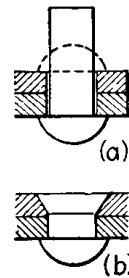
FIGURE 13.20   Rivets.
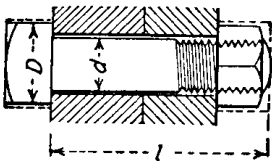
FIGURE   13.21   Machine bolt.
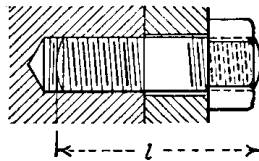
FIGURE 13.22   Stud bolt.
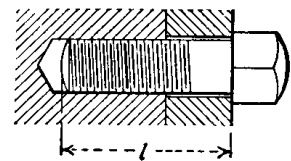
FIGURE 13.23   Cap screw.

Figure 13.21, bolt, through bolt, or machine bolt

Figure 13.22, stud bolt or stud

Figure 13.23, tap bolt or cap screw

The nominal length of the bolt is shown by dimension l in each case. Note that regularly the bolt is loose in its hole, with a diameter difference of $\frac{1}{16}$ in (1.59 mm) for small sizes to $\frac{1}{4}$ in (6.35 mm) for large sizes.

The stud bolt is supposed to be screwed in tightly and permanently; this is better (especially with cast iron) than cap-screw action when the joint has to be opened at intervals, as in the case of a cylinder head.

Screw-thread profiles for bolts and machine screws are shown in Figs. 13.24, 13.25, and 13.26.

18. Bolt and Screw Proportions.   The most important of these are established by the American Standards Association and the American Society of Automotive Engineers.

19. Various Bolts.   The machine bolt and stud are shown in flange joints in Fig. 13.27. An alternative to the stud, found in older practice and in rougher construction, is the T-head bolt in Fig. 13.28. The form of eyebolt in Fig. 13.29 is especially effective on chemical digesters or other pressure tanks that must be opened or closed frequently; the nuts need only be slacked off a few threads, then the bolts can be swung out of the way, making much easier the removal and replacement of the cover; and the screw threads are safe from being battered and spoiled.
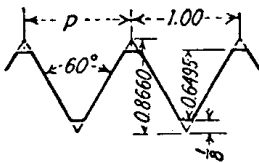
FIGURE   13.24   American thread.
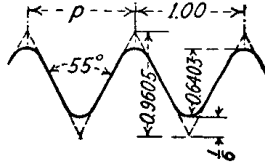
FIGURE   13.25   Whitworth thread.
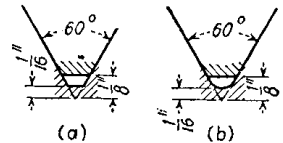
FIGURE 13.26   Groove details.
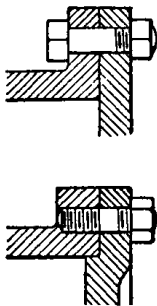
FIGURE   13.27   Flange joints.
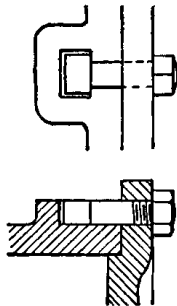
FIGURE 13.28   Tee bolt.
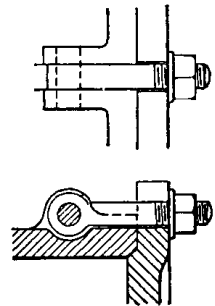
FIGURE 13.29   Eyebolt.

Foundation bolts, of which a highly developed example is shown in Fig. 13.30, are usually double-ended, with special large washers for bearing under the concrete.

20. Screws.   Machine screws are like cap screws in general form and service, but have a range of smaller sizes. The two ranges overlap, however, and the real distinction is that cap screws have diameters in binary fractions of the inch (mm), while machine-screw diameters are in a numbered series. Further, machine screws of a given number come with several numbers of threads per in (mm).

Forms of machine-screw heads are shown in Fig. 13.31, as follows:

1. *Fillister* or *cylindrical head,* here resting on flat surface
2. *Oval* or rounded head
3. *Flat fillister head,* sunk in counterbore, the preferable arrangement
4. *Countersunk head,* either flat or oval (dotted profile)

For more complete information see manufacturers' catalogs of screws and of taps and dies.

21. Set Screws.   The function of the set screw is to exert pressure in the direction of its axis, with the major purpose of preventing sidewise motion of the piece beneath it. If the purpose is merely to hold that piece in a certain position the set screw becomes a clamping or adjusting screw, as in Figs. 13.33 to 13.35. Set screws are alike in being threaded over the whole length of shank, but they differ in heads and points.
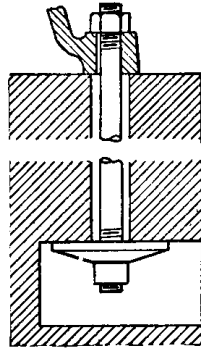
FIGURE 13.30   Foundation bolt.
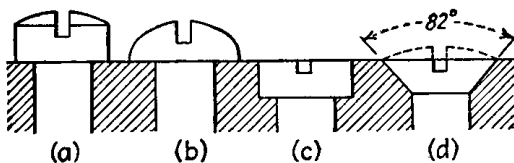
FIGURE 13.31   Machine-screw heads.
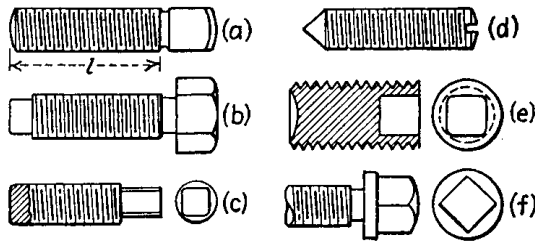
FIGURE 13.32   Set screws, points, and heads.
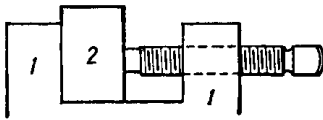
FIGURE 13.33   Clamp screw.
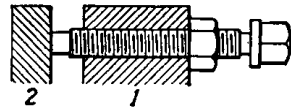
FIGURE 13.34   Screw with lock nut.

FIGURE 13.35   Loose nut.
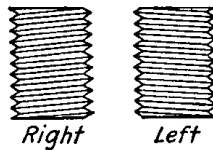
FIGURE 13.36   Right-and-left threads.

FIGURE 13.37   Turnbuckle.

22. Various Screw Threads.   Bolt and screw threads (or fastening threads in general) are single, or have but one helical coil winding around a cylinder; compare Fig. 13.36 with Fig. 13.38.

All ordinary threads are right hand (see Fig. 13.36); seldom is there need for a left-hand screw in a fastening, but the turnbuckle in Fig. 13.37 is a conspicuous example of the right-and-left screw.

A double and a triple thread are shown in Fig. 13.38 (motion screws). Here are distinguished pitch p of thread profile and lead l of the screw, or its advance per turn.

23. Pipe and Pipe Threads.   The American standard pipe thread is profiled in Fig. 13.39. This thread is tapered or is cut on a long cone, the taper being 1 in 16 in the diameter or 1 in 32 on a side.
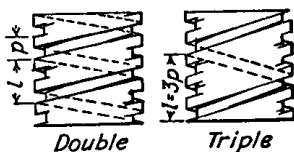
FIGURE 13.38   Multiple threads.
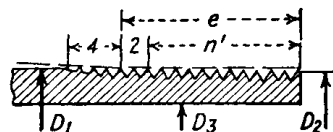
FIGURE 13.39   Pipe thread.

24. Small Pins and Cotters.   Small cotter pins and cotters are shown in Fig. 13.40. Taper pins are rated by the big-end diameter, according to the following scale:

No.	00	0	1	2	3	4	5	6	7	8	9	10
d, in	0.141	0.156	0.172	0.193	0.219	0.250	0.289	0.341	0.409	0.492	0.591	0.706
d, mm	3.6	3.9	4.4	4.9	5.6	6.4	7.3	8.7	10.4	12.5	15.0	17.9

The taper is ¼ in (6.4 mm) to the foot (0.3 m); special reamers for these pins are available. Lengths range from ½ to 2 in (12.7 to 25.4 mm) for size 00 and from 1½ to 6 in (38.1 to 152.4 mm) for size 10.

Spring cotters or split pins, of half-round wire and with easy fit in a hole of their normal size, come in the diameters ³⁄₃₂ to ⁷⁄₃₂ in (2.4 to 5.6 mm) by 64ths (0.4 mm), ¼ to ½ in (6.4 to 12.7 mm) by 16ths (1.6 mm), and ⅝ (15.9 mm), with a range of lengths like that for taper pins.

The flat spring key (Fig. 13.40c) is used to hold rigging together—as the brake rigging of a railway car—with the advantage over the round pin that it is thinner, hence does not so much weaken a rod or bolt.

25. Fits in General.   The several grades of loose fits in the ASME scheme are shown in Fig. 13.41; for the particular diameter [4 in (102 mm)], the sketches show allowance, tolerances, and tightest and loosest fits in each class, vertical dimensions being in mils or thousandths of an inch.
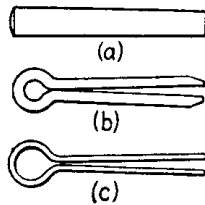
FIGURE 13.40   Cotter pins.
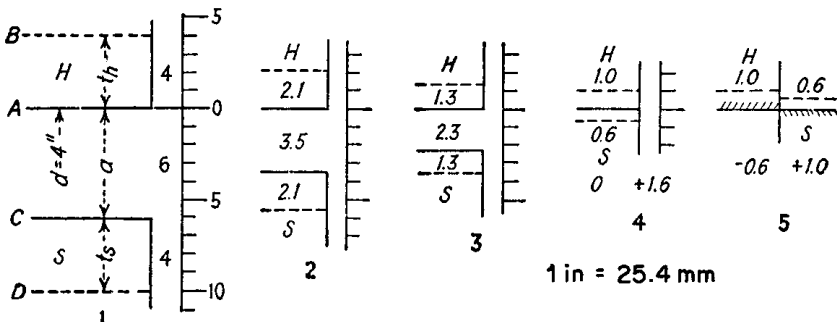
FIGURE 13.41   Allowances and tolerances.

The heights of lines as lettered for case 1 in Fig. 13.41 are as follows:

Line *A*, minimum diameter of hole, equal to nominal diameter *d* of the fit

Line *B*, maximum diameter of hole, above *A* by hole tolerance t_h, which is always +

Line *C*, maximum diameter of shaft or plug, below *A* by allowance *a*; in general

$$a = d_h \text{ min} - d_{s \text{ max}} \tag{13.16}$$

Line *D*, minimum diameter of shaft, below *C* by shaft tolerance t_s, always −

The first four classes are interchangeable fits, the parts to be made within the range of tolerance and assembled at random. Then in any class the looseness ranges from *a* to *a* + 2*t*. The classes are as follows:

Class 1— Loose Fit. Large Allowance. Considerable freedom, for cases where accuracy is not essential; agricultural and mining machinery; textile, rubber, candy, and bread machinery; general machinery of similar grade.

Class 2—Free Fit, Liberal Allowance. Running fits with speeds of 600 r/min and over, journal pressures of 600 lb/in² (4.1 MPa) and over; dynamos and engines, many machine-tool parts, some automotive parts.

Class 3—Medium Fit, Medium Allowance. Running fits under 600 r/min, journal pressure less than 600 lb/in² (4.1 MPa), and sliding fits; the more accurate machine-tool and automotive parts.

Class 4—Snug Fit, Zero Allowance. Closest fit that can be assembled by hand, calls for high precision; used where no perceptible shake is permissible; motion not under load, as in instruments.

Class 5—Wringing Fit, Zero to Negative Allowance. Metal to metal fit, assembly should be selective, to avoid interference shown as one extreme of conditions, in Fig. 13.41, for this class.

The variation of allowance and tolerances with size or diameter is shown in Fig. 13.42, for class 1. This shows steps by 0.001 in (0.025 mm). These dimensions come from formulas which for all the classes of fits are given in Table 13.1.

26. The Shrunk Fit. Simple average values of the two most important coefficients of linear expansion by heat, per degree Fahrenheit (1.8°C), are

For iron and steel, $\qquad a = 0.0000060 \pm 3 \qquad\qquad (13.17a)$

For brass and bronze, $\qquad a = 0.0000100 \pm 6 \qquad\qquad (13.17b)$
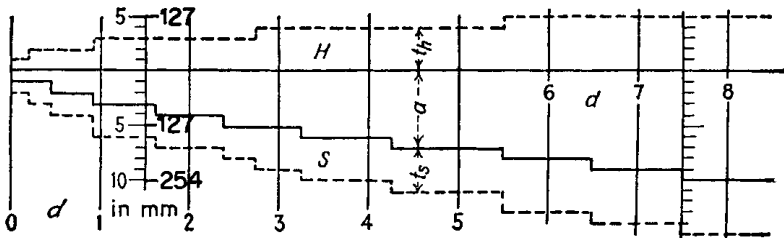
FIGURE 13.42 Variation with diameter.

TABLE 13.1 Recommended Allowances and Tolerances

Fit	Allowance	Hole tolerance	Shaft tolerance
1. Loose	$+0.0025\sqrt[3]{d^2}$	$+0.0025\sqrt[3]{d}$	$-0.0025\sqrt[3]{d}$
2. Free	$+0.0014\sqrt[3]{d^2}$	$+0.0013\sqrt[3]{d}$	$-0.0013\sqrt[3]{d}$
3. Medium	$+0.0009\sqrt[3]{d^2}$	$+0.0008\sqrt[3]{d}$	$-0.0008\sqrt[3]{d}$
4. Snug	none	$+0.0006\sqrt[3]{d}$	$-0.0004\sqrt[3]{d}$
5. Wringing	none	$+0.0006\sqrt[3]{d}$	$+0.0004\sqrt[3]{d}$
6. Tight	$-0.00025d$	$0.0006\sqrt[3]{d}$	$0.0006\sqrt[3]{d}$
7. Medium force	$-0.005d$	$0.0006\sqrt[3]{d}$	$0.0006\sqrt[3]{d}$
8. Heavy force	$-0.001d$	$0.0006\sqrt[3]{d}$	$0.0006\sqrt[3]{d}$

the range of variation being, of course, in the last figure or the seventh decimal place.

For a fit of steel parts the hole member will be heated from 300 to 500°F (149 to 260°C) above room temperature, the maximum of this range being well below the beginning of red heat. The proportional increase of diameter is then from 0.0018 to 0.003, plenty for easy slipping together.

MOTIVE ELEMENTS OF MACHINES

27. Shaft Dimensions. Ordinary transmission shafting comes in diameters $\frac{1}{16}$ in (1.59 mm) scant of the quarter (6.4 mm) or half-inch sizes (12.7-mm), from $1\frac{3}{16}$ to $2\frac{15}{16}$ in (30.2 to 74.6 mm) by quarters to 6 in (152.4 mm) by halves; anything bigger than 6 in (152.4 mm) is usually special. Machine shafts are made to eighths, quarters, or halves. Lengths up to 24 ft (7.3 m) are regular, but 16-, 18-, or 20-ft (4.9-, 5.5-, or 6.1-m) sections are common, to fit spacing of hangers as set by building construction.

The steel in common shafting has an ultimate tensile strength of 60,000 to 65,000 lb/in² (414 to 448 MPa). By cold rolling this is raised to approximately 80,000 lb/in² (552 MPa); and it is usual to say that cold-rolled is to turned (hot-rolled) shafting as 5:4 in transmitting capacity.

By Eq. (13.4) (Sec. 9) with M_t for turning moment PR'' in in · lb (N · m) and with S_s for maximum working shear stress, the torsion formula for a simple solid round shaft becomes

$$M_t = \frac{\pi}{16} d^3 S_s = \frac{1}{5.1} d^3 S_s = 63{,}025 \frac{H}{N} \tag{13.18}$$

Using $S_s = 10{,}000$, and solving for d:

$$d = 0.08\sqrt[3]{M_t} = 3.18\sqrt[3]{H/N} \tag{13.19}$$

With other values of S_s, the value of d as obtained from Eq. (13.19) must be multiplied by the proper factor F_1, as given below:

S_s	F_1	S_s	F_1	S_s	F_1	S_s	F_1
3500	1.419	5000	1.260	8,000	1.077	12,000	0.941
4000	1.357	6000	1.186	9,000	1.036	14,000	0.894
4500	1.305	7000	1.126	11,000	0.969	16,000	0.855

In shafting practice stress S_s is likely to range from 4000 to 8000 lb/in² (28 to 55 MPa) for the maximum load on common steel shafting, and from 5000 to 10,000 (34 to 69 MPa) on cold-rolled steel.

If in addition to turning moment or torque M_t the shaft is also subject to a bending moment M_b, with $M_b = kM_t$, a further factor F_2 must also be applied to d as computed from Eq. (13.19); its value is

$$F_2 = \sqrt[6]{k^2 + 1} \tag{13.20}$$

and values are as follows:

k	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
F_2	1.002	1.007	1.015	1.025	1.038	1.053	1.069	1.086	1.104	1.122
k	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
F_2	1.141	1.160	1.177	1.198	1.217	1.236	1.254	1.272	1.290	1.308

28. Shaft Couplings. The simplest coupling for shaft sections is the plain *muff* or *sleeve* in Fig. 13.43; one long key might be used, driven in from one end, but separate keys, as here, can more surely be driven to a tight fit in each shaft end.

The *flange coupling* (Fig. 13.44) is a much used form.

The *Oldham coupling* (Fig. 13.45) may be used to couple shafts that are likely to be out of line but not out of parallel by more than a fraction of a degree. A tongue-and-groove connection between flange *A* and disk *B* and one between disk *B* and flange *C* (at right angles to the first) permits freedom but ensures an exact transmission of angular velocity.

To meet the condition of nonparallelism, but with little or no allowance fro misalignment, a number of types of *flexible couplings* are in use. Generally a flange is keyed to each shaft end, as in Fig. 13.44; then these are coupled together by one of the following methods:
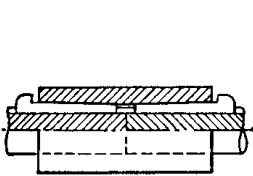
FIGURE 13.43 Simple shaft coupling.
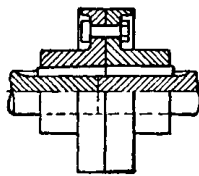
FIGURE 13.44 Flange coupling.
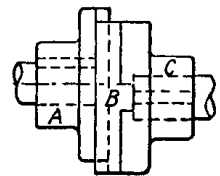
FIGURE 13.45 Oldham coupling.

1. Pins on the flange faces are connected by heavy links of leather or by some such device.
2. One flange is large, the other small, and they are connected by a ring disk of leather or light plate metal, which is fastened to them at its respective edges.
3. Flexible pins—in one case of blades of thin spring steel—are placed and held in holes in the disks.

The *separable coupling* in Fig. 13.46 introduces the class of positive or jaw clutches, but is here supposed to be used only for occasional connection or disconnection.

The *universal joint* (Fig. 13.47) has two forked shaft ends *A* and *C* engaging an intermediate cross *B*. It is highly flexible but does not transmit a uniform angular velocity. If shaft *A* (Fig. 13.48) turns uniformly, shaft *B* will oscillate back and forth with respect to uniformity. But the three-shaft arrangement, all in one plane and with equal angles θ of *A* and *C* from *B* and with the two forks on shaft *B* in the same plane, does give true and smooth motion from *A* to *C*.

29. Clutches. The *jaw clutch* (Fig. 13.46) gives powerful and positive drive but abrupt engagement, and hence can be thrown in only at low speed.

A *conical friction clutch* of heavier and rougher type is sketched in Fig. 13.49; its function is to connect hoist drum 3 to shaft 1 through disk 2.

A *disk clutch,* is shaft-coupling service, is outlined in Fig. 13.50.

The *multidisk* or *Weston clutch* in Fig. 13.51 is much used.

A *ring clutch* of the shafting-transmission kind, in the service of connecting pulley 3 to shaft 1, is shown in Fig. 13.52.

The most powerful type of friction clutch (within set limits of size and in the sense of getting a strong grip from a small gripping force) is the *coil clutch* of Fig. 13.53.

30. Bearings. The meaning of *journal* (inner, turning) and *bearing* (outer, standing) is made clear by Fig. 13.54, where the bearing might be of plain cast iron; also the simple lubricating system of oil hole and oil groove is shown. In Fig. 13.55 appears the more usual scheme of a lining or *bushing* of special material for the
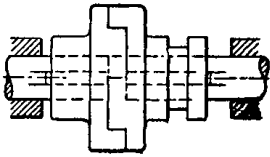
FIGURE 13.46 Jaw clutch (a separable coupling).
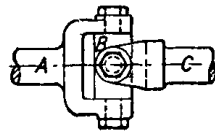
FIGURE 13.47 The universal joint.
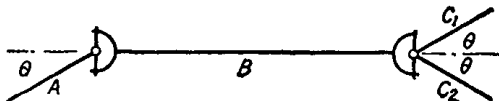
FIGURE 13.48 Diagram of action.
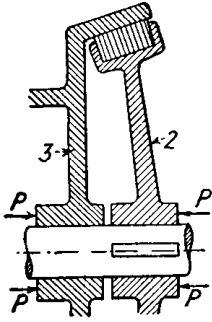
FIGURE 13.49   Cone clutch.
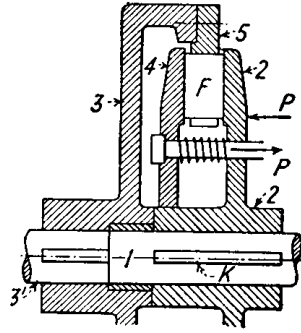
FIGURE 13.50   Disk clutch.
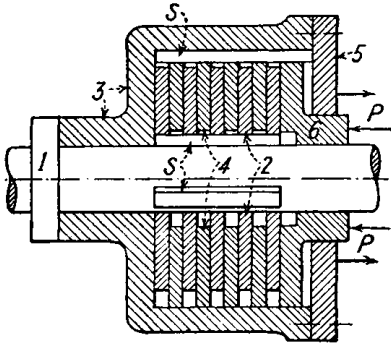
FIGURE 13.51   Multidisk or Weston clutch.
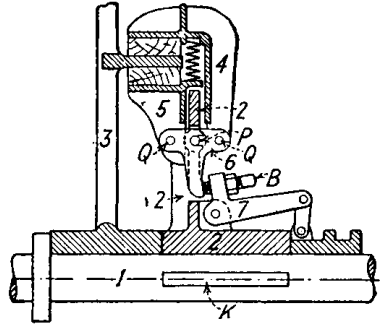
FIGURE 13.52   A ring clutch.
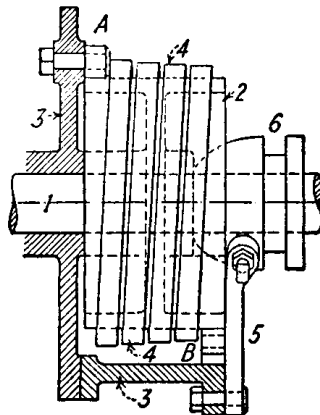
FIGURE 13.53   Coil clutch.
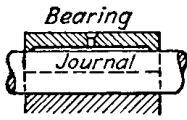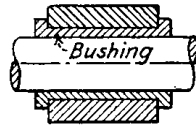
FIGURE 13.54   **FIGURE 13.55**   **FIGURE 13.56**
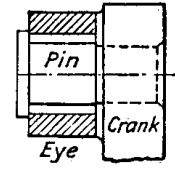
FIGURES 13.54 to 13.56   Forms of bearings, fixed and moving.

bearing. Figure 13.56 shows the case, out of linkage mechanism, in which both parts move.

31. Bearing Pressures.   The term *bearing pressure,* as regularly used with regard to cylindrical bearings, means average pressure on the projected area, or on the rectangle of diameter d and length l, and is thus a simplified representation of something that may really be very complex. If P is the total pressure or load and p the unit pressure in lb/in² (kPa),

$$p = \frac{P}{dl} \tag{13.21}$$

For a thrust bearing product dl will be replaced by the actual area of a circle, a ring, or a number of rings.

Allowable bearing pressure depends on a number of considerations, which fall under three heads, as follows:

1. Kind and degree of lubrication and character of surfaces, which determine the strength or amount of frictional resistance, or the coefficient of friction μ.
2. Velocity V of slip or sliding, which has an influence upon μ but is chiefly the second factor of the work done in overcoming friction.
3. Freedom of escape or removal of heat. A bearing may be enclosed, may be open to a free circulation of air, may have a rapid circulation of oil (cooled outside), or may be watercooled; typically, the bearings of a marine engine are backed by hollow spaces through which water can be pumped.

A useful measure of the burden on a bearing is found in the product.

$$q = pV \tag{13.22}$$

with pressure p by Eq. (13.22) and velocity V in ft/min (m/min). Multiplication by coefficient μ gives the friction work μV (ft · lb) per in²/min (N · m per cm²/min). With varying degrees of ventilation, but no special cooling arrangements, the rate of escape of heat may range from 100 to 1500 ft · lb per in²/min. (21.0 to 315.3 N · m per cm²/min) the maximum temperature being about 150°F (65.6°C) above the surrounding air. When the air is strongly agitated, as in the case of engine crank pins, the heat rate may be appreciably higher than the limit just stated.

32. Roller and Ball Bearings.   Concerning these low-resistance bearings a large amount of detailed and quantitative information is available in the technical literature issued by the manufacturers. Regularly, these bearings are made up and pur-

chased as assembled units, ready to be installed in the machine. Questions of load capacity and of mounting should in any case be taken up with the makers. In general, the mounting should support but not distort, permit and maintain alignment, retain lubricant, and exclude dirt. Lubrication is needed because of the slight friction of rolling, the incidental friction of guiding, and to prevent rusting.

33. Load Capacity of Bearings.   The crushing load on a ball is proportional to the square of its diam. d; and the working load P is either

$$P = cd^2 \qquad \text{or} \qquad P = D\,cd^2 \tag{13.23}$$

With race diam. D in fairly constant ratio to ball diam. d, the second expression is equivalent to the first. In $P = cd^2$, coefficient c will range from less than 1000 to more than 5000, depending upon material, accuracy of production methods, kind of contract (ranging from a flat race to a closely curved groove), and speed. As to speed, one rule makes P vary inversely as the cube root of the r/min or of N.

  In a radial bearing, the question arises as to the number of balls, of the total number n, which are active in carrying the load. As angle θ from the line of direct thrust is greater, the ball receives less elastic compression. Hence its pressure or reaction P is less, and of this P, only a component is effective against load Q. It can be shown that this component pressure is proportional to the $5/2$ power of cos θ. The effective number n_s of balls under full central load P is approximately

$$n_s = 0.23n \tag{13.24}$$

commonly assumed as $0.2n$; this gives

$$Q = \frac{Pn}{5} \tag{13.25}$$

for the total load Q in terms of safe ball load P.
  The most used type of formula for roller bearings is

$$Q = kdln \tag{13.26}$$

with l for length and d, n, and Q as before. This probably underrates the influence of diam. d, which should have an exponent ranging from 1.3 to 1.4.
  For definite load ratings, see makers' bulletins.
  Balls and rollers and their races are made of special alloy steels, with chromium as the ingredient most effective in providing the need combination of hardness and tough strength; requirements are most severe in the case of ball material. Very high precision is desired and attained, with such a tolerance as ± 0.1 mil or even ± 0.05 mil on diameter and sphericity.

34. Gear Teeth.   In Fig. 13.7, Sec. 4, pairs of gears have been represented by their pitch circles, which are the profiles of kinematically equivalent friction wheels, assumed to roll upon each other without slip. To make this relative motion positive and to give capacity for transmitting force and work, teeth are formed by cutting into and building out from these cylinders. The primary requirement imposed upon tooth profiles is that their motion upon each other shall produce and ensure perfect rolling of the pitch circles. If the teeth are inaccurately spaced or are of incorrect profile, the driven gear or follower will receive an irregular motion and the pair will run noisily, especially at high speed.

The parts and dimensions of gear teeth are shown in Fig. 13.57. The primary dimension is *circular pitch p*, the length of pitch-circle arc from a point on one tooth to the like point on the next. Strictly, pitch is measured or is made actual as an angle, since it is the function of the indexing mechanism of any gear-cutting machine to divide 360° of angle into any desired integral number of equal parts.

The size of teeth is regularly designated, however, not by the circular pitch *p* but by the *diametrical pitch* p_d, which is the number of teeth per inch (mm or cm) of diameter. With pitch-circle diam. *D* and the number *n* of teeth in the wheel, the two pitches are

$$p = \frac{\pi D}{n} \tag{13.27a}$$

$$p_d = \frac{n}{D} \tag{13.27b}$$

whence $pp_d = \pi$. Being the number of teeth to 1-in (25.4-mm) diam., p_d is the same number per π in (mm) of circumference. As p_d is larger, the teeth are smaller.

Under the SI metric system this designating ratio is inverted to the *module,* which is the number of millimeters of diameter per tooth, or $m = D/n$. Between our diametrical pitch and the metric module the relation is

$$mp_d = 25.4 \tag{13.28}$$

this being the number of millimeters to the inch. Using the module as a convenient linear unit in our system, its value is simply $1/p_d$ in.

The other dimensions shown on Fig. 13.57, with their values in a long-established standard system of gear teeth, are as follows:

h = *addendum,* equal to the module or $1/p_d$
d = *dedendum* or root depth, $1.157/p_d$
t = tooth *thickness,* at pitch circle
s = width of tooth *space,* also at pitch circle
b = *breadth* of gear rim
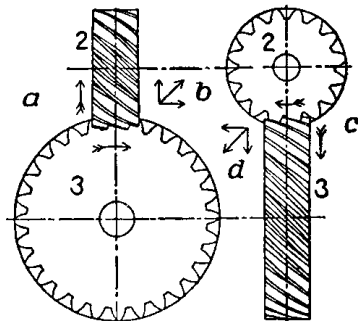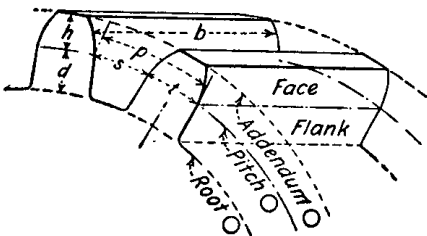
FIGURE 13.57 Parts and dimensions of teeth.

FIGURE 13.58 Helical or spiral gears.

35. Kinds of Gears. Toothed gears may connect shafts with parallel axes, with axes that intersect, or with axes that cross but do not intersect.

With axes parallel, the wheels are called *spur gears.* The teeth on a gear may be straight (parallel to axis) as in Fig. 13.57, inclined or helical, Fig. 13.58, or double helical or of "herringbone" pattern. Helical teeth give smoother running at high speed, since all portions of the pair of profiles are in contact simultaneously. With a double rim, or with two rows of teeth of opposite slant, end thrust is avoided.

When the axes intersect, the wheels become *bevel gears,* as in Figs. 13.59 and 13.60.

With axes that cross, the *worm* and *worm wheel* (Fig. 13.61), the pair of *helical* or *spiral gears* (Fig. 13.58), and the geometrically much more complex hyperboloid or "skew-bevel" pair are used.

36. Strength of Gear Teeth. Figure 13.62 indicates that the gear tooth functions as a short cantilever beam; the thrust and moment of pressure P are balanced by a combination of shearing and flexural stress at section BB. Total P may come on one tooth at its tip, especially if this tooth is on a small pinion; and in that location the tooth will have least width at root BB. Assuming different sizes of teeth to be similar in form, dimensions will be proportional to pitch p. Then with a and c as numerical coefficients, b for breadth of gear face as in Fig. 13.57, and S for the stresses due bending, the relation of load moment to resisting moment is

$$aPp = cSbp^2 \qquad \text{whence } P = kSbp \qquad (13.29)$$

with k as a comprehensive constant or factor.
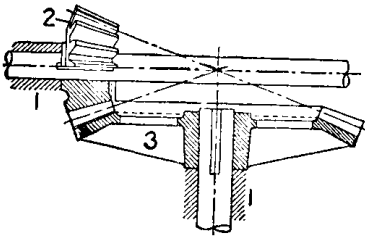
FIGURE 13.59   Bevel gear and pinion.
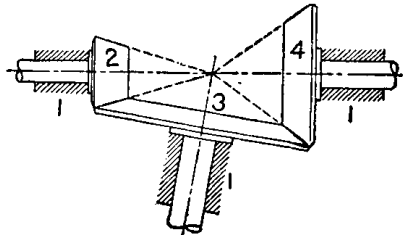
FIGURE 13.60   Gears with oblique axes.
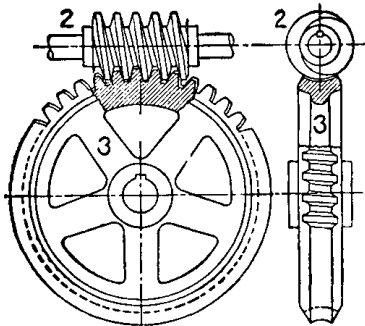
FIGURE 13.61   Worm and worm wheel.
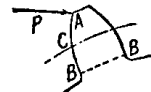
FIGURE 13.62

The much used *Lewis formula,* based on Eq. (13.29) and on a study of varying tooth proportions, uses two equations for *k*, namely,

$$k = 0.124 - \frac{0.684}{n} \qquad (14.5° \text{ obliquity}) \qquad (13.30a)$$

$$k = 0.154 - \frac{0.912}{n} \qquad (20° \text{ obliquity}) \qquad (13.30b)$$

These take account of the decrease of root width as tooth number *n* is smaller, holding down to *n* = 12. For stub teeth, *k* would be larger, but has not bee formulated.

For pitch-circle velocities not greater than 100 ft/min (30.5 m/min), values of nominal stress *S* to be used in Eq. (13.29) are

Cast iron: 8000 lb/in² (55.2 MPa)

Cast steel: 20,000 lb/in² (137.8 MPa)

Forged machinery steel: 25,000 lb/in² (172.4 MPa)

Alloy steel, up to: 40,000 lb/in² (275.8 MPa)

With rising velocity, stress *S* diminishes, rapidly at first and then more slowly, falling to about one-fourth of the preceding values at *V* = 2000 ft/min (609.6 m/min).

ENERGY ENGINEERING

PUMPS

1. Definitions.   A *pump* is a machine or device for raising a liquid, which is a relatively incompressible fluid, to a higher level or to a higher pressure. A *compressor* is a machine or device for raising a gas, which is a compressible fluid, to a higher pressure. Devices for exhausting air from closed vessels are called *air pumps,* though in reality they are air compressors working below atmospheric pressure.

A *blower,* as distinguished from a compressor, compresses a gas to a comparatively low pressure only. A *fan* is intended primarily to move large volumes of gas; the pressure developed by the fan is quite small and is secondary in importance.

2. Measurement of Head.   The head which a pump has to develop or work against is the static lift plus all the friction losses in the piping. This value may be computed, but in actual operation it would be determined in a test by measuring the pressure in the piping adjacent to the pump on both suction and discharge sides. Let *h* = total head in ft (m), *p* = pressure expressed in ft (m) of the liquid, *z* = elevation of the center of the discharge gage above the point at which the suction pressure is measured, *V* = velocity in ft/s (m/s) at the section where the gage is attached, *g* = acceleration of gravity in ft/s² (m/s²), the subscript *d* denotes discharge, and the subscript *s* denotes suction values. Then

$$h = p_d - p_s + z + \frac{V_d^2}{2g} - \frac{V_s^2}{2g} \qquad (13.31)$$

If the pressure on the intake side is below atmospheric, and if gage pressures are used in the above equation, then p_s will be negative.

3. Power. If q = rate of discharge in ft³/s (m³/s), G = gal/min (L/min), w = density of the liquid in lb/ft³ (kg/m³), the horsepower (W) delivered in the liquid, called *water horsepower*, is: Water power = $wqh/550$ hp (1 hp = 0.75 kW). In the case of water of the customary density of 62.4 lb/ft³ (8.0 kg/m³), this may be reduced to

$$\text{Water power} = \frac{qh}{8.81} = \frac{Gh}{3960} \quad \text{hp (1 hp = 0.75 kW)} \qquad (13.32)$$

For any other liquid of specific gravity s, the two expressions in the above equation, and in the one below, should be multiplied by s. If e is the overall efficiency of the pump, then the power input to the pump, often called *brake horsepower*, is

$$\text{Brake power} = \frac{qh}{e \times 8.81} = \frac{Gh}{e \times 3960} \quad \text{hp (1 hp = 0.75 kW)} \qquad (13.33)$$

4. Efficiencies Defined. *Efficiency,* sometimes called *total* or *overall efficiency,* is the ratio of the power delivered in the liquid to the power input to the pump. That is

$$e = \frac{\text{water hp}}{\text{brake hp}} = \frac{\text{water kW}}{\text{brake kW}} \qquad (13.34)$$

Hydraulic efficiency, e_h, is the ratio of the power actually delivered *in* the water to the power expended *on* the water or other liquid. These two quantities differ by the amount of the hydraulic-friction losses.

Mechanical efficiency, e_m, is the ratio of the power expended *on* the water to the power supplied to run the pump. These two differ by the amount of the mechanical-friction losses, such as friction of bearings, stuffing boxes, etc.

Volumetric efficiency, e_v, is the ratio of the amount of water actually delivered to that which would be delivered if there were no leakage losses, imperfect valve action, etc. *Slip,* in the case of a positive displacement pump, means the difference between the actual displacement and the volume of the fluid actually delivered, expressed as a percentage of the displacement. In certain combinations of a reciprocating pump and pipe line the inertia of the water causes flow to continue even while the pump is on dead center, and thus secures a discharge larger than the actual displacement volume. The relation between slip and volumetric efficiency is: slip = $100(1 - e_v)$.

The *total efficiency* is the product of the hydraulic, mechanical, and volumetric efficiencies, that is,

$$e = e_h \times e_m \times e_v \qquad (13.35)$$

Duty is another means of expressing the efficiency of steam-driven pumping engines. It is usually expressed as the foot-pounds of work done per 1000 lb (J/

454 kg) of steam supplied, but is more precisely defined as the foot-pounds of work done per million Btu (J/1.1 MJ) supplied.

5. Suction Lift.   The theoretical suction lift may be computed as follows:

$$L = b - p_v - h_f - \frac{V_s^2}{2g} \tag{13.36}$$

where L = lift, b = barometer pressure in ft (m) of the liquid, p_v = vapor pressure of the liquid in ft (m), h_f = friction losses in foot valve, suction piping, etc., and V_s = velocity at intake of pump. For water it is desirable to maintain a pressure at least about 10 ft (3 m) more than the vapor pressure. Hence the maximum allowable lift is about 10 ft (3 m) less than given by the above equation. In practice the lift is usually about 20 ft (6 m) for cold water, decreasing as the water temperature increases; above 160°F (71°C), the water should be supplied under a positive pressure.

POSITIVE DISPLACEMENT PUMPS— RECIPROCATING TYPE

6. Direct-Acting Steam Pumps.   The direct-acting steam-driven pump, such as show in Fig. 13.63, is one of the simplest types of reciprocating pumps.

Direct-acting steam pumps are very uneconomical, requiring from 100 to 300 lb (45 to 136 kg) of steam per hp · h (0.75 kWh). They are therefore, usually used only where the exhaust steam can be utilized for some heating purpose, or where economy is secondary to simplicity, ruggedness, and reliability.

If p_s and p_w denote steam and water pressures, respectively, while A_s and A_w denote the areas of the steam and water pistons, respectively, then

$$p_s A_s = m p_w A_w \tag{13.37}$$

where m is a factor greater than unity, in order to allow for friction losses, etc. For a boiler-feed pump, where the water pressure is but little greater than the steam pressure, m is given a value as high as 3 or 4 in order to provide an ample margin. In other cases it need be no larger than the reciprocal of the pump efficiency.

The dimensions of the water end may be computed from the equation
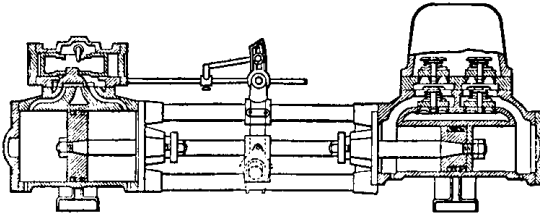
FIGURE 13.63   Single-cylinder steam pump, submerged-piston type.

$$e_v d^2 L N' = 294G \qquad (13.38)$$

where e_v is the volumetric efficiency which may range from 0.85 to 0.99, depending upon the condition of the pump; d is the diameter of the piston in in (mm); L the length of stroke in in (mm); N' the number of working strokes per min; and G is gal/min (L/min). If the pump is single-acting, then N' is the same as N, the number of "revolutions" per min. If it is double-acting, then $N' = 2N$.

7. Crank-and-Flywheel Pumps. To use steam expansively, it is necessary to use a flywheel or equivalent device to equalize between the steam and water cylinders. The addition of a flywheel necessitates a crank and connecting-rod mechanism, and thus a more complex construction than with the direct-acting type. Large triple-expansion pumping engines of the crank-and-flywheel type have given steam consumptions as low as about 10 lb (4.6 kg) of steam per hp · h (0.75 kW).

8. Power Pumps. A power pump is a positive displacement pump of the reciprocating type operated through cranks and connecting rods by power supplied to the crankshaft. The source of power is usually an electric motor which may be connected to the shaft through gears or by a belt and pulleys. The latter is more quiet.

The direct-acting steam pump is inherently a variable-speed pump. The power pump is inherently a constant-speed pump, and is thus not so well suited to a variable rate of discharge.

9. Efficiencies of Reciprocating Pumps. The overall efficiencies of reciprocating pumps, whether direct-acting or power pumps, range from about 45 percent to about 90 percent or more, depending upon the size and the condition of the pump. The higher values are usually found only in large pumps and the lower values in very small pumps. Crank-and-flywheel or power pumps are usually a few percent more efficient than direct-acting pumps of the same size.

10. Characteristics of Reciprocating Pumps. A distinguishing characteristic of the positive displacement pump is that the capacity—neglecting slip—is equal to the displacement. Thus the capacity is directly proportional to the speed of the pump, and for a constant speed the capacity is constant. If a variable quantity is desired, it is necessary to vary the speed of the pump, or to simply bypass some of the fluid, which is wasteful of power.

The slip or leakage increases as the pressure on a given pump is increased. It tends to become less in percentage value as the pump speed is increased.

The head or pressure which the positive displacement pump will develop is determined by the resistance against which the pump works, and is not fixed by any property of the pump. Its maximum or limiting value is determined by the power available to drive the pump and the strength of the various parts. Obviously such a pump should be protected against excessive pressures by some form of relief valve. The positive displacement pump is of especial value, compared with the centrifugal pump, for small capacities, and especially for very high heads.

POSITIVE DISPLACEMENT PUMPS—ROTARY TYPE

11. Types of Rotary Pumps. The principal types of rotary pumps are shown in Fig. 13.64. These are the *lobe* type, *eccentric* type, and *gear* type. In each of them the fluid is trapped within the spaces bounded by the casing and the vanes, teeth, or lobes, and is delivered to the discharge side. Such pumps have no valves. Since the leakage loss is kept within a reasonable value by nothing more than the small clearance between the various parts, such pumps deteriorate rapidly with wear. Hence they should not be used for a liquid containing grit. They are of especial value for such liquids as oils, whose physical properties produce less leakage loss than water and to provide better lubrication to such rubbing surfaces as are in contact.

The *lobe type* may be used for moderate rotative speeds. It produces a continuous but nonuniform rate of discharge and hence requires air chambers. Efficiencies as high as 85 percent are claimed for it. The *eccentric type* has blades which slide in and out and thus produce inertia forces due to their reciprocating actions, which limits the speed of such pumps. The blades can never wear to a tight fit because of the varying curvature and, furthermore, they produce considerable friction loss. There is considerable slip around the ends of the blades and the rotor, which is hard to check satisfactorily. The *gear type* is best suited to high rotative speeds, but is not quite so efficient as the lobe type. In addition to these three there is a *twin-screw type,* which consists of two parallel screws within a casing in which the thread of one fits into that of the other.

12. Characteristics of Rotary Pumps. Rotary pumps, being positive displacement pumps, are very similar to reciprocating pumps in most respects. The principal advantage is the substitution of rotation for reciprocation; this permits higher speeds, with a reduction in the size of the pump for the same capacity. Rotary pumps tend to be inefficient, however, especially with wear, which permits the leakage loss to become excessive.

CENTRIFUGAL PUMPS

13. Classifications. Centrifugal pumps may be divided into *turbine* pumps and *volute* pumps. In the former the impeller is surrounded by diffusion vanes which are so formed as to reduce the velocity of the water and efficiently convert the kinetic energy into pressure. In the latter the impeller is merely surrounded by a
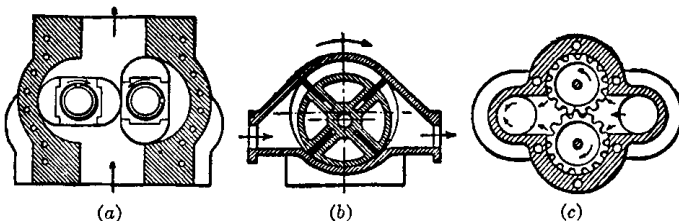
FIGURE 13.64 Rotary pumps. (*a*) Lobe type. (*b*) Eccentric type. (*c*) Gear-wheel pump.

spiral casing, called the *volute,* which is also so designed as to convert efficiently the kinetic energy into pressure (see Figs. 13.65 and 13.66).

They may also be classified as *single-stage* and *multistage,* according to whether only one impeller is used, or whether there are more than one in series. In the latter case each stage adds to the pressure produced by the preceding stage, so that the total pressure is equal to that developed by one impeller alone multiplied by the number of stages (see Fig. 13.67).

If the impeller takes water from one side only, the pump is called a *single-suction pump.* If it takes water from both sides, it is a *double-suction pump.* A single-suction impeller produces an end thrust which must be provided for, while a double-suction impeller is in hydraulic balance.

Sometimes deep-well pumps are the *axial-flow* type, in which case the impellers become very similar in appearance to screw propellers as used for ships. They are, in fact, a form of *propeller* pump. Propeller pumps are also used in large sizes for pumping large quantities of water against relatively low heads.

14. Nominal Pump Size. The size of a centrifugal pump is usually designated by the diameter in inches of the discharge-pipe connection. As the velocity of the water at that section is usually about 10 ft/s (3.0 m/s), one may thus obtain an
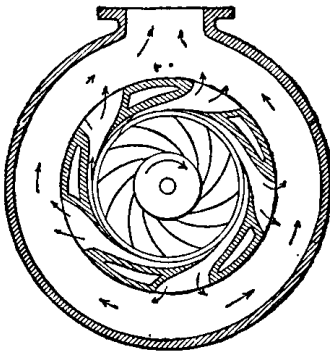
FIGURE 13.65 Turbine pump with circular case.
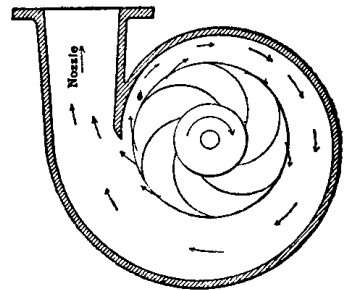
FIGURE 13.66 Volute pump.
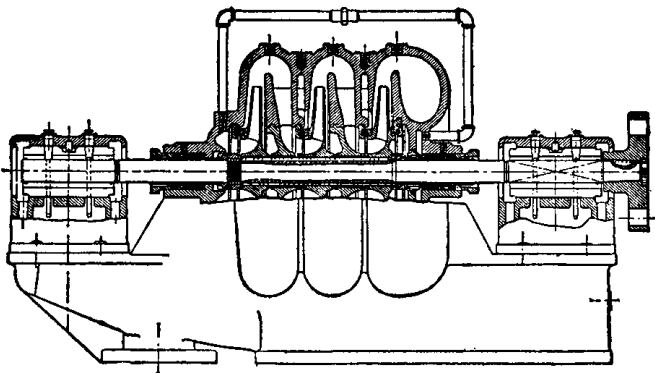
FIGURE 13.67 Multistage pump.

approximate value of the pump capacity by multiplying the area so obtained by 10. But this is subject to considerable deviation.

15. Pump Characteristics. Owing to the practical difficulty of determining the proper values of certain velocities, angles, and factors to use in the equations presented in Sec. 16, the relation between head, discharge, and pump speed is best determined by test rather than by calculation. Test curves for pumps running at constant speed are shown in Fig. 13.68.

16. General Laws and Factors. It is sometimes convenient to use factors ϕ and c such that $u_2 = \phi\sqrt{2gh}$ and $v_2 = c\sqrt{2gh}$, where u_2 = the peripheral velocity of the impeller, v_2 = the relative velocity of water at exit from impeller, both in ft/s (m/s), while h is the head in ft (m) of the fluid.

While for a given pump, values of ϕ and c may vary over a wide range, the values of especial interest are those for the head and capacity at which maximum efficiency is obtained. These values for maximum efficiency are

$$\phi = 0.90 \text{ to } 1.30$$

$$c = 0.10 \text{ to } 0.30$$

The rate of discharge of the pump in ft²/s (m³/s) is $q = a_2v_2$, where a_2 = the cross-sectional area of the streams of water at exit from the impeller, measured normal to the direction of v_2 and expressed in ft² (m²).

The radial component of the water velocity is $V_2 \sin \alpha_2$ or $v_2 \sin \beta_2$. If B and D represent impeller width and diameter in inches (mm), respectively, n the number of vanes, and t their thickness measured along the circumference, the net circumferential area in ft² (m²) is $B(\pi D - nt)/144$. Hence

$$q = v_2 \sin \beta_2 \frac{B(\pi D - nt)}{144} \quad (0.028 \text{ m}^3/\text{s}) \ (q) \quad (13.39)$$

The factor giving the radial component of the velocity is $c \sin \beta_2$. The vane angle β_2 may be anything from 10° to 80° or more, but it is usually between 20° and 30°.
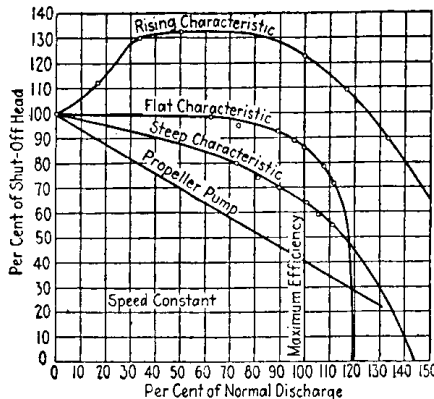
FIGURE 13.68 Head-discharge characteristics.

Thus the factor for the radial component of velocity may range from about 0.04 to 0.15 in ordinary practice.

Along any one parabola in Fig. 13.69, the factor ϕ is constant. When ϕ is constant, then c and, likewise, the hydraulic efficiency e_h are also constant. And only when ϕ and c are constant may one apply the simple ratios previously given, that

$$q \text{ varies as } N$$

$$h \text{ varies as } N^2$$

$$hp \text{ (W) varies as } N^3$$

The overall efficiency will not be quite constant, but may be assumed so for a reasonable speed range.

17. Variation with Diameter. For a series of homologous impellers, in which all angles are the same, and all ratios of dimensions are the same, but which differ from each other in actual dimensions only so that each one is simply an enlargement or a reduction of another, all areas will vary as the square of the diameter. And, for the same rotative speed in all cases, all velocities will vary as the first power of the diameter. Thus for such a series at the same rotative speed, the following will hold:

$$q \text{ varies as } D^3$$

$$h \text{ varies } D^2$$

$$hp \text{ (W) varies as } D^5$$

But if the impellers were not homologous but differed in diameter only, while the impeller width remained unchanged, it would be found that
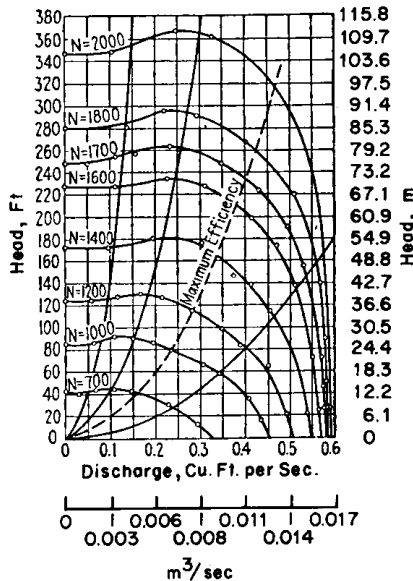
FIGURE 13.69 Relation between head and discharge at various speeds.

$$q \text{ varies as } D^2$$

$$h \text{ varies as } D^2$$

$$\text{hp (W) varies as } D^4$$

This assumes that each impeller is a separate construction such that the diameter is the only dimension that is changed, except that each diameter of impeller is in a case of suitable proportions for it.

The diameter of impeller, speed, and head may all be combined in the following useful expression:

$$DN = 1840\phi\sqrt{h} \tag{13.40}$$

For a series of homologous impellers, ϕ is constant. Thus, for that series $DN/\sqrt{h}$ is constant. In practice, for different designs of centrifugal pumps this combination may range in value from 1660 to 2400, according to the value of ϕ for maximum efficiency.

18. Specific Speed. A most useful factor is the specific speed, which may be expressed as*

$$N_s = \frac{N\sqrt{\text{gal/min}}}{h^{3/4}} \tag{13.41}$$

If preferred, cubic feet per second may be used instead of gallons per minute, but the latter will give values 21.2 times that of the former. In this expression the values for speed, capacity, and head should be those at which the maximum efficiency is obtained. [Note that this differs from specific speed as used for hydraulic turbines in which horsepower (kW) instead of capacity is employed.] The specific speed of any one pump is a constant, and thus the above may be used to obtain possible combinations of the three factors involved. But it is also a constant for a series of homologous centrifugal pumps. The value of the head to be used to obtain N_s is the head per stage in the case of a multistage pump.

The preceding expression for specific speed applies directly to single-suction pumps. For double-suction pumps the specific speed should be computed by employing one-half of the actual pump capacity, since such an impeller is equivalent to two single-suction impellers placed back to back.

Small values of N_s are found with pump impellers whose width is small compared to the diameter, while large values are found for large values of that ratio. Usually the ratio of B/D will be found to range from 2 to 70.

For the centrifugal pump, values of specific speed are usually between the limits of 500 to 9000. With propeller pumps, this value may be increased up to 16,000.

19. Efficiency of Centrifugal Pumps. The efficiency obtained with centrifugal pumps depends among other things upon the absolute size, as large pumps are more efficient than small ones. In Figs. 13.70 and 13.71 are shown probable optimum values of efficiency as a function of the pump capacity.

---

*The SI specific speed = 0.613 times the U.S.Customary System N_s when flow is expressed in L/s and head in m.
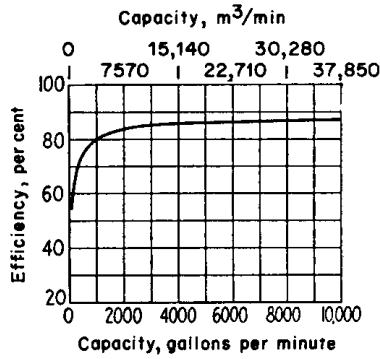
FIGURE 13.70 Optimum efficiency of turbine pumps as a function of capacity.
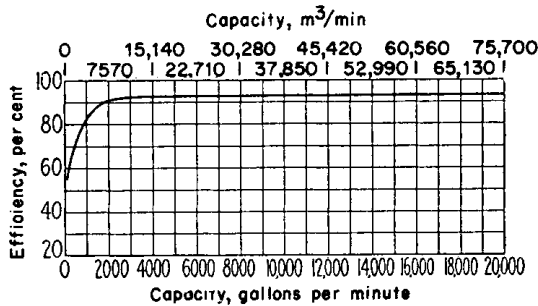
FIGURE 13.71 Optimum efficiency of volute pumps as a function of capacity.

Efficiency is also a function of specific speed, as shown by Fig. 13.72. In fact the actual efficiency depends upon both the capacity and the specific speed; and to get the highest values, the pump must be of reasonably large capacity and also of a favorable specific speed.

20. Head per Stage. Usual practice is to limit the head per stage to about 100 to 150 ft (30.5 to 45.7 m). But heads of several hundred feet (m) per stage have been employed successfully. The principal criterion is whether a favorable or even a possible value of specific speed may be obtained. Provided a favorable value for N_s is obtained, the efficiency does not seem to be materially affected by the value of the head used.

21. Operating Characteristics of Centrifugal Pumps. Unlike the positive displacement pump, the head and capacity are not independent of each other but are mutually related. Thus the head cannot be varied without changing the rate of discharge. The head developed is a function not only of the rate of discharge but also of the pump speed. On the other hand, the pump may be operated at a constant speed and the rate of discharge may be varied; this is impossible with the positive
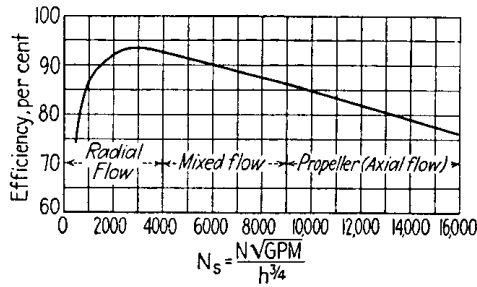
FIGURE 13.72 Optimum efficiency as a function of specific speed.

displacement pump. The centrifugal pump has the advantage of being suited to the high rotative speeds that are found also with electric motors and steam turbines.

22. Selection of Type of Pump. The selection of the type of centrifugal pump is primarily a matter of choice of specific speed. For a given installation the total head is usually fixed by the physical setting and is not subject to change. However, it may be developed in one stage or divided among a number of stages of a multistage pump. The total quantity to be pumped is also fixed but, if large, may be divided among several pumps. The rotative speed may also be given different values, but it is usually confined to a limited number of synchronous speeds if the pump is directly connected to an alternating-current motor. Hence by varying any one or all three of the terms entering into the specific speed, the latter may be given a series of values, often over a very wide range.

The factors that affect the selection of the most suitable specific speed are as follows: The higher the specific speed, the smaller the pump and electric motor and the less the cost. But Fig. 13.72 shows that after a specific speed of about 3000 is exceeded, the possible efficiency to be obtained becomes less. Also Table 13.2 shows that the higher the specific speed, the less is the allowable suction lift. In fact, unless the head per stage is low, the required pump submergence may be impracticable. Thus this last consideration may impose a limit upon the maximum specific speed that can be employed in a given case.

MISCELLANEOUS TYPES OF PUMPS

23. The Air Lift. The air lift shown in Fig. 13.73 consists of a long pipe, of which a considerable portion is submerged below the surface, into which air is admitted near the lower end. This air bubbles up through the water in the pipe and, when water is flowing out at the top, the weight of the mixture of air and water within the pipe produces a unit pressure at the lower end that is slightly less than that produced by the water surrounding the pipe.

Increasing h_s will reduce the quantity of air that is required, but will increase the pressure necessary, for the air pressure must be slightly greater than that due to the depth of water h_s. In practice h_s is from 1.0 to $4.0h_d$.

The volume of air required, expressed as cubic feet of free air, or air at atmospheric pressure, will be given by

TABLE 13.2   Values of σ as a Function of Specific Speed

Specific speed	σ	
	Maximum	Minimum
500	0.030	0.018
1,000	0.075	0.046
1,500	0.13	0.080
2,000	0.19	0.11
3,000	0.33	0.20
4,000	0.48	0.29
5,000	0.64	0.31
6,000	0.82	0.52
7,000	1.0	0.52
8,000	1.2	0.74
9,000	1.5	0.90
10,000	1.6	1.0
12,000	2.1	1.2
14,000	2.5	1.5
16,000	3.0	1.8
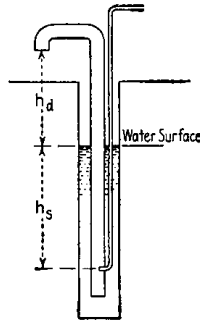
FIGURE 13.73   Air lift.

$$v = \frac{Q(h_d + h_f + h_v)}{34 \log_c (p_1/p_a)} \; (0.028v \text{ m}^3) \qquad (13.42)$$

where v = ft³ (m³) of free air, Q = ft³ (m³) of water, h_d = lift, h_f = friction head, h_v = velocity head, all in ft (m); while p_1 is pressure at depth h_s, and p_a is atmospheric pressure.

The air lift is of value in obtaining large flows from deep wells of small diameters. It is of especial value if the water contains grit or other material which would damage any working parts of a pump. This type of pump is also very useful in pumping oil from very deep wells. Instead of air, a gas obtained from the oil is usually used, and it is then known as a *gas lift.*

24. Jet Pumps.   Jet pumps are of two types: the *injector* and the *ejector.* The injector (Fig. 13.74) uses steam as the working medium and is employed to pump feed water into the boiler. Steam, discharging at high velocity through a nozzle,
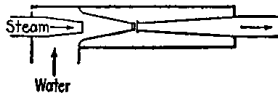
FIGURE 13.74 Injector.

FIGURE 13.75 Ejector.

mixes with cold water, which condenses the steam and, at the same time, warms the water, and the mixture then passes on into the boiler. The heat-balance equation for this is

$$0.98(H - h_2) = W(h_2 - h_1) \qquad (13.43)$$

where W = lb of water per lb of steam (kg/kg), H = total heat per lb (kg) of steam supplied, h_2 = heat of liquid of water going to boiler, h_1 = heat of liquid of cold water supplied.

The law of momentum, with a correction factor introduced to take care of losses, is

$$0.5(V_1 + WV_2) = (1 + W)V_3 \qquad (13.44)$$

where V_1 = velocity of steam jet, V_2 = velocity of cold water, usually negligible in practice, and V_3 = velocity of the mixture. In accordance with Bernoulli's theorem, the velocity V_3 is largely transformed into pressure before entrance into the boiler by means of the diverging tube as shown.

The efficiency of the injector, considered solely as a pump, is very low, being only about 2 percent. But since all the heat of the steam used is returned to the boiler, save for a small radiation loss, its thermal efficiency approaches 100 percent. The steam consumption of an injector is about 400 lb of steam per hp · h (182 kg per 0.75 kWh). The injector cannot be used to lift hot water, and it is not an economical water heater. Its principal merits are its small size and light weight.

The ejector (Fig. 13.75) is operated with either water or steam or compressed air as the working medium, and it may pump either a liquid or a gas. It usually uses a small quantity of the working medium at a high pressure and pumps a large quantity at a low pressure. The equations for the injector apply to it also, except that if steam is not used, the heat-balance equation cannot be employed. The efficiency of the ejector is very low, being from 15 to 30 percent. It may be used for capacities up to 700 gal/min (2650 L/min).

COMPRESSORS

25. Density of Air. The weight of dry air, lb/ft³ (kg/m³), is given by $w_a = 2.7p_a/T_a$, where p_a is the air pressure, in lb/in² abs (kPa abs), and T_a the temperature, in °F abs (°C abs).

The true pressure of the air is $p_a = B - hp_v$, where B is the barometer pressure (in same units as p), h the relative humidity, and p_v the vapor pressure for saturation at T_a. This is the value given in steam tables. The weight of water vapor, lb/ft³ (kg/m³), is $w_v = h/V_s$, where V_s is the specific volume of dry saturated steam at T_a.

The weight of the moist air is $w = w_a + w_v$. The quantity w is the actual weight per cubic foot (m³) that is handled by the compressor. At 65°F (18.3°C), 14.7 lb/in² (101.3 kPa) barometer, and 70 percent humidity, $w = 0.0752$.

26. Free Air. By free air is meant air at the pressure and temperature of the place from which the compressor draws its supply. Its volume may be obtained from the volume at any other set of conditions by the equation $pV/T =$ constant.

27. Gas Laws. If a perfect gas is compressed or expanded, it follows the law $pV^n =$ constant. If the process is isothermal, $n = 1$. If it is adiabatic, $n = k$, which for air is approximately 1.4. In a piston compressor, n is usually from 1.3 to 1.35, and in a centrifugal compressor very nearly 1.4.

The temperature change during compression or expansion is determined by the relation

$$\frac{T_2}{T_1} = \left(\frac{V_1}{V_2}\right)^{n-1} = \left(\frac{p_2}{p_1}\right)^{(n-1)/n} \tag{13.45}$$

28. Volumetric Efficiency. This term is really a capacity factor, as it represents the relationship between the actual capacity of a machine and its physical size, and in an air compressor has but small effect upon true efficiency. It is defined as the ratio of the volume of free air actually delivered by the compressor to its displacement volume.

In Fig. 13.76, the compressed air in the clearance space at the end of the stroke expands down to 4 before the suction valve can open. The volume of air drawn into the machine is only V_1, while the displacement volume is V_p. If $c = V_c/V_p$, or the percent of clearance divided by 100, then

$$V_1 = V_p \left[1 + c - c \left(\frac{p_2}{p_1}\right)^{1/n} \right] \tag{13.46}$$

To reduce the volume V_1 to that of free air, multiply it by $(p_1/p_a)(T_a/T_1)$. The effect of the clearance and the value of p_1/p_a can be obtained from an indicator card, but the temperature of the air in the cylinder at the end of the suction stroke is not known.

For piston compressors, clearance is usually from 2 to 6 percent, the larger values being found with smaller machines. The actual volumetric efficiencies of
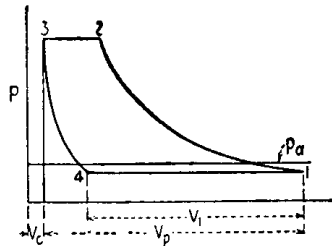
FIGURE 13.76

single-stage compressors are from 50 to 85 percent, depending upon the clearance and the pressure.

29. Multistaging. For high pressures, compressors are built in two or more stages, thus reducing the pressure range within each cylinder and hence maintaining a reasonable volumetric efficiency. If the air is also cooled, between stages, to the initial temperature, there is a saving of power as shown in Fig. 13.77, where ABC is the compression line for a single cylinder and AD is an isothermal.

For minimum work with perfect intercooling, the intermediate pressures are given by $p = \sqrt{p_1 p_2}$ for a two-stage compressor; $p' = \sqrt[3]{p_1^2 p_2}$ and $p'' = \sqrt[3]{p_1 p_2^2}$ for a three-stage machine.

If there were no pressure drops in the intercoolers, the cylinder volumes would be inversely proportional to these pressure and to p_1. Owing to the higher volumetric efficiency with the lower pressure, the volume of the low-pressure cylinder will be less than that of a single-stage compressor of the same capacity.

30. Compressor Efficiencies. The *mechanical* efficiency is the work done in the air cylinders as given by the actual indicator card divided by the indicated work in the engine cylinders, if driven by a direct-connected engine, or by the work delivered by a belt, if it is belt-driven. For piston compressors, mechanical efficiencies are about 90 percent for direct-connected machines and a few percent higher for the belt-driven types.

The *compression* efficiency is the ratio of the work required to compress and deliver the gas, as given by an ideal cycle with isothermal (or adiabatic) compression, to the work done on the air, as shown by actual indicator cards. Values of *isothermal compression* efficiency are usually from 65 to 70 percent and nearly independent of the pressure, while *adiabatic compression* efficiencies are: 64 percent for $p_2/p_1 = 1$; 76 percent for $p_2/p_1 = 2$; 87 percent for $p_2/p_1 = 3$; and about 90 percent for all higher ratios.
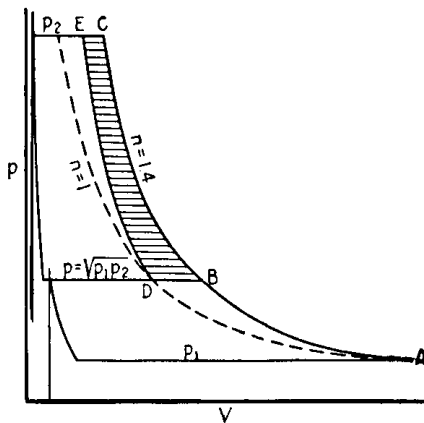
FIGURE 13.77 Two-stage compression with intercooling.

31. Power Required for Compressor.   The power required to compress and deliver to the receiver any gas is given theoretically for adiabatic compression by

$$\text{Power} = 0.01525 p_1 V_1 \left[\left(\frac{p_2}{p_1} \right)^{0.286} - 1 \right] \quad \text{hp (1 hp} = 0.75 \text{ kW)} \quad (13.47)$$

where p_1 is initial pressure and p_2 final pressure, in lb/in² abs (kPa abs), and V_1 is ft³ of gas per min (m³/min) at p_1. This is for a single stage only. For a multistage compressor with S stages and perfect intercooling, multiply 0.01525 by S and divide 0.286 by S.

For isothermal compression the equation for power becomes

$$\text{Power} = 0.01 p_1 V_1 \log_{10} \left(\frac{p_2}{p_1} \right) \quad \text{hp (1 hp} = 0.75 \text{ kW)} \quad (13.48)$$

The actual power required to drive the compressor may be obtained from either of the above by dividing by the mechanical efficiency and the proper compression efficiency as given in Sec. 30.

The theoretical power for other values of n may be obtained by the equation

$$\text{Power} = 0.00437 \left(\frac{n}{(n-1)} \right) p_1 V_1 \left[\left(\frac{p_2}{p_1} \right)^{(n-1)/n} - 1 \right] \quad \text{hp (1 hp} = 0.75 \text{ kW)}$$

$$(13.49)$$

For small pressure differences, where $(p_2 - p_1)/p_1$ is less than 0.1, such as with fans and blowers, it is sufficiently accurate to use the "hydraulic" formula for a fluid of constant density, which is power $= 0.00437 (p_2 - p_1) V_1$ hp (1 hp $= 0.75$ kW) in place of any of the preceding equations.

32. Piston Compressors.   Reciprocating piston compressors are built for pressures from 1 to 100 lb/in² or more in single stages, up to 500 lb/in² (3.4 MPa) in two stages, 1200 to 2500 lb/in² (8.3 to 17.2 MPa) in three stages, and up to 5000 lb/in² (34.5 MPa) in four stages. For moderate pressures double-acting pistons are used, but for very high pressures it is customary to use a single-acting plunger. Compressor cylinders and cylinder heads should be water-jacketed. Very small sizes may be air-cooled by metal fins.

33. Rotary Blowers.   Rotary blowers are also of the positive displacement type, and are usually limited to pressures of from 0.5 to 12 lb/in² gage (3.4 to 82.7 kPa). They are very similar to rotary pumps for liquids. For small sizes and high pressure, the slip tends to be excessive; but, under favorable conditions and for low pressures, efficiencies as high as 80 to 90 percent are claimed.

34. Centrifugal Compressors.   Centrifugal compressors are very similar to centrifugal pumps in all respects, the difference being that they handle a gas of variable density instead of a liquid of constant density. In addition to its increase in density, the temperature of the gas rises during flow through the machine, and thus introduces thermodynamic effects. Because of the increase in density, all impeller dimensions should be decreased with successive stages, but for manufacturing reasons

impellers either are made all alike if the number of stages is small, or are arranged in groups if many stages are employed. Another difference from the centrifugal pump is that the machine is usually water-jacketed.

Single-stage compressors are built for pressures of from 1 to 5 lb/in² gage (6.9 to 34.4 kPa) and even up to 15 lb/in² (103.4 kPa). Capacities range from 100 to 100,000 ft³/min (2.8 to 2831 m³/min). Multistage machines are built for pressures up to 125 lb/in² (861.8 kPa) and even up to 200 lb/in² (1379 kPa) in a few cases. Peripheral velocities up to 1000 ft/s (304.8 m/s) are employed, and efficiencies are as high as 75 percent.

35. Centrifugal Fans. The centrifugal fan is practically identical with the centrifugal pump so far as its theory is concerned, because the variation of the density of the gas within it is negligible. Fans are built to produce pressures up to 0.5 lb/in² (3.4 kPa) or 28 in (711.2 mm) of water, but in usual practice the figure is very much less than that. Capacities range as high as 400,000 ft³/min (189 m³/s).

36. Fan Pressures. Because of their small values, fan pressures are usually stated in terms of inches (mm) of water or ounces per square inch (kPa). One inch (25.4 mm) of water is equivalent to 0.577 oz/in² or 0.0362 lb/in² (0.25 kPa) or 69.3 ft (21.1 m) of air whose density is 0.0752 lb/ft³ (1.2 kg/m³). Three pressures are considered: static, velocity, and impact, which is the sum of the first two. The last is usually used for computing fan efficiency, but the static pressure may also be used. Hence it is necessary to specify which pressure is used.

If the velocity head of a gas is represented by h_v in (mm) of water, the velocity in feet per minute (m/min) is $V = (1100/\sqrt{w})\sqrt{h_v}$, where w = density in lb/ft³ (kg/m³).

37. Air Horsepower. The power in the air delivered by the fan is $5.2Qh/33,000$, where Q = ft³/min (m³/min), and h = in (mm) of water. It may be either static pressure or impact pressure.

38. Fan Efficiency. The overall efficiency of a centrifugal fan is e = ahp/bhp, where ahp is the air horsepower and bhp the brake horsepower or the power necessary to run the fan (e = akW/bkW). Its value is usually about 60 to 70 percent.

39. Fan Characteristics. The characteristics of centrifugal fans are identical with those of the centrifugal pump but instead of one curve for h there are often two, the static-pressure and the impact-pressure curves, and likewise there are two efficiency curves to correspond. The various curves may be plotted with Q as a base or, since the static head is often of greater interest and more readily measured, all values may be plotted against it.

40. Steam-Jet Blowers and Vacuum Pumps. For many purposes the steam-jet compressor or vacuum pump (Fig. 13.78) is very useful. It is simple and compact, but is inefficient unless exhaust steam is used for it. A form of this is widely used for air pumps on steam condensers. They use from 0.19 to 0.37 lb of steam per 1000 ft³ (0.09 to 0.17 kg per 28.3 m³) of air.
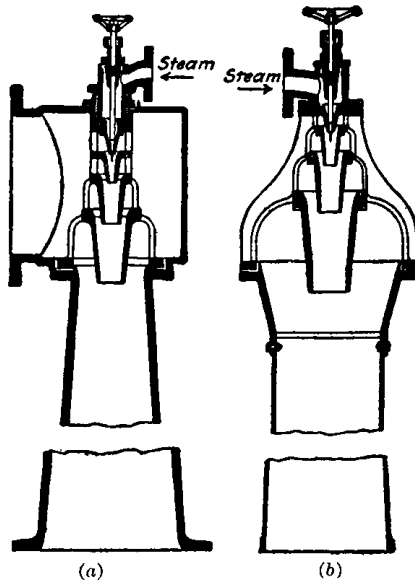
FIGURE 13.78   (*a*) Exhauster, (*b*) Steam-jet air compressor.

FUELS AND COMBUSTION

41. Fuels.   In a commercial sense, a fuel is any material of suitable cost that will combine, in part, with the oxygen of the air to liberate heat. Fuels may be classified as *solid, liquid,* and *gaseous. Natural* solid fuels are coal, lignite, peat, and wood; *prepared* solid fuels include coke, pulverized coal, and charcoal; *waste* solid fuels include sawmill refuse and bagasse (crushed sugar cane).

42. Coal.   Coal, the most important and most abundant solid fuel, is of vegetable origin and exists in a variety of forms having different chemical and physical characteristics. It contains carbon, hydrogen, oxygen, nitrogen, sulfur, and ash.

Coal Analysis.   Two kinds of analyses are commonly used to ascertain the composition of coal: the ultimate analysis and the proximate analysis.

Ultimate Analysis.   The chemical analysis of a fuel, giving the percentages of carbon, hydrogen, oxygen, nitrogen, sulfur, and ash, is known as the *ultimate analysis.* For the standard method of making these determinations, see ASME "Test Code for Solid Fuels."

Proximate Analysis.   This analysis determines the percentages by weight of moisture, volatile matter, fixed carbon, and ash in the fuel. Usually a statement of the sulfur content and the calorific value of the fuel is included in the commercial proximate analysis. *Moisture* is considered to be the loss in weight of a sample (1 g) of coal when dried at a temperature of 104 to 110°C for 1 h. *Combustible* is arbitrarily defined as that portion of the coal remaining after subtraction of ash and

moisture. *Volatile matter* is the total combustible minus the fixed carbon; it includes the carbon which is combined with hydrogen, together with free oxygen, nitrogen, and other gas-forming constituents of the dry fuel which are driven off by the heat. *Fixed carbon,* or uncombined carbon, is the combustible remaining after the volatile matter has been driven off. It is determined by subtracting from the weight of the original sample the weight of the moisture, ash, and volatile matter. *Ash* is the incombustible residue from the complete burning of the coal. For the standard method of making these determinations, see ASME "Test Code for Solid Fuels."

43. Heating Value. The *calorific* or *hearing value* of a fuel is the amount of heat recovered when the products of complete combustion of unit quantity of the fuel are cooled to the initial temperature of the air and fuel. The heating values of solid and liquid fuels are expressed on the weight basis in Btu per pound (kJ/kg). For gaseous fuels, the heating values are expressed in Btu per cubic foot (kJ/m³) of the gas; common standard conditions for the measurement of the gas volume are a temperature of 60°F (15.6°C), a total pressure of 30 in Hg (76.2 cm Hg), and "saturation" of the gas with water vapor. When a fuel contains hydrogen, there are a number of possible heating values depending upon the fractional part of water vapor formed during combustion that condenses when the products of combustion are cooled to the initial temperature. If none of the water vapor condenses, the resultant heating value is the *lower heating value;* if all the water vapor condenses, the heating value is the *higher heating value.* The numerical difference between the higher and lower heating value of any hydrogen-containing fuel is the product of the weight of water vapor formed from the complete combustion of unit quantity of the fuel and the latent heat of condensation of that water vapor. The heating value of fuels may be determined by calorimeter tests or may be estimated, with a fair degree of accuracy, from a knowledge of their chemical and physical properties. The approximate higher heating value (HHV) of coal, in Btu per pound (kJ/kg), may be calculated from the ultimate analysis by a formula of the *Dulong* type;

$$\text{HHV} = 14{,}500C + 62{,}000 \left(H - \frac{O}{8} \right) + 4000S \ (\times\ 2.3 = \text{kJ/kg}) \quad (13.50)$$

where *C*, *H*, *O*, and *S* are the fractions by weight of carbon, hydrogen, oxygen, and sulfur in the coal.

44. Liquid Fuels. Petroleum in its unrefined state, frequently called *crude oil,* is a viscous, dark brown or greenish liquid occurring in natural reservoirs in the earth's crust in many parts of the world. Regardless of their source, petroleum and petroleum products have ultimate analyses that generally fall within the following limits:

	Carbon	Hydrogen	Sulfur	O + N	Moisture and sediment
Percentage by weight	80 to 87	10 to 15	0.3 to 2.5	1 to 7	0.1 to 1.5

In the purchase and sale of petroleum and its distillates, the specific gravity is of importance. In general, the higher the specific gravity, the lower Is the content

of lighter hydrocarbons and the lower the heating value of the oil. Commercially the specific gravity is ascertained by means of a hydrometer. The reading of the hydrometer is related to the true specific gravity as indicated by the expression: Specific gravity = $141.5/(131.5 + B)$, in which B = deg. API (American Petroleum Institute) as read from the scale of the hydrometer. Table 13.3 gives the relation of API readings to density and volume. Fuel oil is generally sold by the barrel, each barrel containing 42 gal of 231 in³, at a temperature of 60°F (159 L of 3785 cm³ at 15.6°C).

The higher heating value of fuel oil, in Btu per pound (kJ/kg), is given approximately by the Sherman and Kropf empirical equations;

$$\text{HHV} = 18{,}650 + 40(B - 10) = 13{,}050 + \frac{5600}{\text{Sp. gr.}} \ (\times\ 2.3 = \text{kJ/kg}) \quad (13.51)$$

The suitability of a fuel oil for *power-plant use* depends on the cost on the basis of heating value, and freedom from grit, acids, sulfur, and other objectionable constituents; the oil should have a flash point over 150°F (65.6°C) and a suitable viscosity for pumping.

Fuel oil for use in *engines* must be clean and noncorrosive, and must have a viscosity that permits its flow through small passage. *Gasoline* is the best known and most widely used engine fuel and is obtained largely by distillation of petroleum. *Casing-head gasoline* is obtained from natural gas by compression and absorption methods; it is too volatile for commercial sale and is blended to produce satisfactory engine fuel.

The higher heating value of gasoline in Btu per pound (kJ/kg), is closely given by the following empirical equation:

$$\text{HHV} = 12{,}720 + \frac{5600}{\text{Sp. gr.}} \ (\times\ 2.3 = \text{kJ/kg}) \quad\quad (13.52a)$$

Octane (C_8H_{18}) is about the average of the mixture of several hydrocarbons that constitute gasoline.

TABLE 13.3 Relation of API Readings to Density and Volume

Deg. API	Sp. gr. 60°/60°F (15.6°C/15.6°C)	Lb per U.S. gal (kg/L)	Weight per bbl, lb (kg)
10.0	1.000	8.328 (0.998)	349.8 (158.7)
15.0	0.9659	8.044 (0.964)	337.8 (153.2)
20.0	0.9340	7.778 (0.932)	326.7 (148.2)
25.0	0.9042	7.529 (0.902)	316.2 (143.4)
30.0	0.8762	7.296 (0.874)	306.4 (138.9)
35.0	0.8498	7.076 (0.848)	297.2 (134.8)
40.0	0.8251	6.870 (0.823)	288.5 (130.9)

Kerosene is the distillation product that lies between gasoline and the distillate fuel oils; it is used in tractor and similar engines. Kerosene is composed of a number of hydrocarbons of which $C_{12}H_{26}$ is about the average. The higher heating value of kerosene, in Btu per pound (kJ/kg), is closely given by the following empirical equation:

$$\text{HHV} = 12{,}840 + \frac{5600}{\text{Sp. gr.}} \ (\times \ 2.3 = \text{kJ/kg}) \qquad (13.52b)$$

45. Gaseous Fuels. Natural gas, blast-furnace gas, producer gas, and byproduct coke-oven gas are used as fuels for the generation of steam and directly for the generation of power in large internal-combustion engines.

Natural gas is of organic origin and is usually found along with petroleum. The composition of natural gas depends upon the location of the gas field.

Coal gas results from the destructive distillation of coal and is popularly known as *city gas* or *illuminating gas.* In general, coal gas is too costly for engine use. It is, however, a highly satisfactory fuel. *Coke-oven gas* is a byproduct of coke ovens. It has a good heating value and, when freed from dust, sulfur compounds, and ammonia, makes a very good engine fuel. *Carbureted water gas* is blue water gas into which, during the manufacturing process, petroleum oil is sprayed; this process forms a composite gas that has a greater heating value per cubic foot and can be burned with a luminous flame. Much city gas is in whole or part carbureted water gas. It is a fine engine fuel, but like coal gas its cost is usually too high for use in engines. *Oil gas* is produced by the decomposition (cracking) of petroleum oils by the application of high temperatures. It is similar in composition to coal gas. Its uses are similar to those of coal gas. It is an excellent engine fuel but, like that of coal gas, its cost is ordinarily too high to allow extensive use. *Producer gas* is formed when a mixture of air and steam is blown up through a thick, hot mass of coal or coke. *Blast-furnace gas* is an important engine fuel and is generally used in large engine units. It is also burned in large quantities directly under boilers. The volume of gas produced per ton of iron made is from 130,000 to 150,000 ft^3 (3680 to 4247 m^3). When the dust has been removed, blast-furnace gas is a very good gas-engine fuel, because it is low in hydrogen and relatively free from impurities.

46. Minimum Amount of Air Required for Combustion. The air required for the complete combustion of a fuel depends upon the chemical composition of that fuel. The atomic weights of elements important in combustion are: carbon, 12.00; hydrogen, 1.00; sulfur, 32.06; nitrogen, 14.01. For combustion calculations, atmospheric air may be treated as a mechanical mixture of nitrogen and oxygen in the proportion of 79 parts N_2 to 21 parts O_2, by volume, and 77 parts N_2 to 23 parts O_2 by weight. By writing the equation of the chemical reaction, the weight of oxygen (and of air) required to burn completely each combustible element of the fuel may be determined. For example, for the complete combustion of 1 lb (0.45 kg) of carbon, $C + O_2 = CO_2$. By substitution of atomic weights, 12 lb (5.4 kg) of carbon + 32 lb (14.5 kg) of oxygen yields 44 lb (19.9 kg) of CO_2. In other words, to burn completely 1 lb (0.45 kg) lb of carbon requires 32/12(0/23) or 11.6 lb (5.3 kg) of air. The weight of air required to burn completely 1 lb (0.45 kg) of hydrogen is 34.8 lb (15.8 kg), and to burn 1 lb (0.45 kg) of sulfur, 4.35 lb (1.97 kg).

Excess Air. If the minimum amount of air required for complete combustion were supplied to an actual furnace, the fuel would not burn completely, largely because of the imperfect mixing of the oxygen and combustibles. It is necessary to supply excess air in order to prevent the loss of heat due to incompleteness of combustion. An increase in excess air causes a greater loss of heat in the exit gases, however, and there is an optimum amount of excess air for any one fuel and set of operating conditions where the total loss of heat will be minimum. For any one fuel, the CO_2 content of the products of combustion may be used as an index of the excess air. The coefficients in the chemical-reaction equations may be used to represent volumes of gaseous constituents.

INTERNAL-COMBUSTION ENGINES

47. Characteristics. Important characteristics distinguishing internal-combustion engines from steam engines and other heat engines are the combustion of the fuel with the necessary air directly in the engine cylinder and the use of the resulting products of combustion as the working substance. For all ordinary considerations, the working substances may be considered as having the properties of air. In internal-combustion engines the maximum as temperature existing in the cylinder occurs during the explosion or combustion process; 3300°R may be taken as a representative average. At full load, the temperature of the gases at the end of expansion will average about 1800°R, and the temperature of the exhaust gases will average about 1200°R. The internal-combustion engine operates over an average range of temperature of 1500°F (816°C), while the steam plant operates over an average range of about 700°F (371°C). Consequently, the theoretical efficiencies of practical internal-combustion-engine cycles are much higher than those of practical steam cycles. In general, any first-class internal-combustion engine of any given rating commercially available at the present time will operate with a much higher actual overall thermal efficiency than any first-class commercially available steam plant of the same rating.

48. Cycles. All commercial engines operate on either (1) the *Otto cycle,* (2) the *Diesel cycle,* or (3) a cycle which embodies characteristics of both the Otto and the Diesel and is variously termed the *semi-Diesel cycle, mixed cycle,* or *Sabathe cycle.*

49. Engine Classification and Performance. Engines for *stationary* uses include many sizes and types, from the small gasoline farm engines to the larger gas engines driving gas compressors, electric generators, pumps, or blowers. The development of the oil industry has made available large quantities of natural gas which has been piped to the large industrial centers and used in many gas engines.

OIL ENGINES

50. Low- and Moderate-Compression Engines. An oil engine is an internal-combustion engine that uses oil for fuel. Air, only, is drawn into the cylinder and compressed, and the fuel is metered, injected, and atomized at the proper time

in the stroke by means of a fuel pump driven from the engine. Oil engines may be arbitrarily classified as low-compression, moderate-compression, or high-compression (Diesel) engines.

Low-compression oil engines have compression pressures that are usually less than 150 lb/in^2 (1034 kPa). The temperature of the air at the end of compression is not high enough for autoignition, and some uncooled portion of the cylinder, not bulb, hot bolt, or similar device must aid in producing ignition of the fuel.

Moderate-compression oil engines employ compression pressures between 200 and 400 lb/in^2 (1379 to 2758 kPa) and operate on a mixed cycle with part of the combustion at constant volume and part of substantially constant pressure. In general, these engines now have complete cylinder cooling and will start cold under ordinary conditions, the heat of compression producing the necessary ignition temperature.

51. Diesel Engines.   Diesel or high-compression engines are used in stationary service, ships, trains, tractors, busses, trucks, and aircraft. The Diesel engine is characterized by its ability to burn low-grade and cheap fuels; by its high thermal efficiency over a wide range of loads; by the absence of all electrical ignition equipment ad carbureting devices.

Fuel injection is an important feature in engine performance and may be accomplished by *air injection* or *solid injection* of the fuel. For air injection a multistage air compressor is employed, and it delivers air at about 800 to 1500 lb/in^2 (5515 to 10,341 kPa). This compressor is usually built integrally with the engine but may be an independent unit. A fuel pump is also required to deliver the proper amount of fuel to the injection nozzle at the proper time in the stroke. This pump is driven directly by the engine.

Stationary or *heavy-duty Diesel engines* are built for two-cycle and four-cycle operation, and are of the single- and double-acting, single- and multi-cylinder, and horizontal and vertical types. The weights of these engines per brake horsepower range from 25 to 300 lb (15 to 182 kg/kW), with an average of 100 lb (61 kg).

Marine engines do not differ markedly from stationary engines, except that their piston speeds are a little higher and they are fitted with the necessary auxiliaries; fuel consumption is about the same.

High-speed Diesel engines find extensive use in applications where rotative speeds from 600 to 1200 r/min are desired. High-speed Diesel engines run at speeds up to 2500 r/min with piston speeds of 2000 ft/min (610 m/min), and they weight 10 to 15 lb/hp (6 to 9 kg/kW).

STEAM-POWER PLANT EQUIPMENT

52. Furnaces.   In steam generation, correct furnace design and operation are essential to the efficient combustion of fuels. The furnace must be adapted to the fuel; it must be constructed and operated so that the combustible gases will be thoroughly mixed with the proper amount of air to support combustion, and the gases must be maintained at a temperature above their ignition point until they have been completely burned.

Coals high in volatile matter, when burned in bulk form, require a furnace design which will cause the distillation of the volatile matter to take place at low temperatures. The resulting light hydrocarbons are more likely to burn completely without

depositing soot than the heavier compounds resulting from high-temperature distillation.

Air-cooled walls require a greater furnace volume for a given rate of energy release or output than do *water-cooled walls.* With hand or stoker firing, the cross section of the furnace is usually fixed by the area of the grate; the height of the furnace must then be great enough to give the required furnace volume. Water-cooled walls may be of the (1) bare-plate, (2) bare-tube, or (3) covered-tube type. A water-cooled wall that consists of tubes covered with protective, refractory-lined metal blocks is shown in Fig. 13.79.

Refractories.   Refractory brick for lining the furnace of a steam-power plant should possess the following properties: (1) high fusion point, (2) low thermal conductivity, (3) low thermal expansion, and (4) high resistance to abrasion. Failure in service may be due to: (1) fusion, (2) plastic deformation, (3) activity—expansion, shrinkage, and spalling—(4) slagging. Spalling means any breaking or cracking of bricks, whether due to thermal shock, pinching because of expansion, or changes in structure. Slag action results in the erosion of the brick due to fluxing by the ash of the fuel, or to the building up of layers of solid slag upon the surface.

53. Mechanical Stokers.   Mechanical stoking is always superior to hand firing when thermal efficiency and smokeless combustion are alone considered. Overall economy, however, measured in dollars, must give consideration to fixed charges and maintenance, as well as to operating expense. Automatic stokers are high in first cost and, if they are applied to too small units, the fixed charges and maintenance may offset the saving in fuel. This is very likely to be the case when no reduction in the firing force can be made and where the plant operates only fraction of the time.
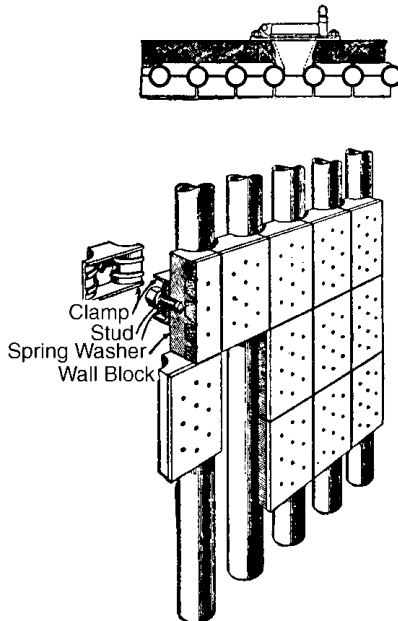
FIGURE 13.79   Bailey furnace wall.

Commercial stokers are classified as (1) *overfeed* and (2) *underfeed,* depending upon whether the coal is fed to the fuel bed above or below the level of the supply of primary air. The *overfeed* type may be divided into three classes: (a) front-feed inclined-grate stokers in which the fuel is received at the front and fed down an incline to the ash dump at the bottom, (b) double-inclined side-feed stokers with the fuel fed from both sides, and (c) traveling grates with continuous horizontal feed. The underfeed type may be *single-retort* or *multiple-retort.* Draft for mechanical stokers may be natural or forced.

54. Pulverized Coal.   Coal that has been crushed, dried, and ground to a powder has been used as a power-plant fuel since about 1918. The advantages are complete combustion with minimum excess air; high furnace efficiency and capacity; and the fact that cheaper and widely varying grades of coal may be burned, and fuel and air supply can readily be controlled to conform with variations in load. Other factors which must be given consideration are first cost, size of plant, space requirements, preparation costs, maintenance; and ash and dust collection and disposal.

Straight-shot burners (Fig. 13.80) discharge the coal and primary air in a straight line and may be horizontal, vertical, or inclined.

55. Oil Burners.   The use of oil as a fuel for steam-generating units in industrial and central-station power plants is widespread. The oil commonly used is that meeting the Bunker C specifications of the U.S. Navy. This oil has a viscosity of not over 300 sec. Saybolt Furol at 122°F (50°C) and a density of 5 to 14 API. Oil burners are classified as (1) mechanical, (2) steam- or air-atomizing, and (3) rotary-cup. Mechanical burners use mechanical means of atomizing the oil, and are supplied with oil at pressures of from 200 to 300 lb/in^2 (1379 to 2068 kPa); burners of this type are provided with registers for combustion air. Capacity is governed by changing the oil pressure or by cutting burners in and out; between 15 and 1100 gal of oil (57 to 4164 L) per hour are usually handled per burner. Most steam-atomizing burners use dry steam at pressures above 30 lb/in^2 (207 kPa) to cut
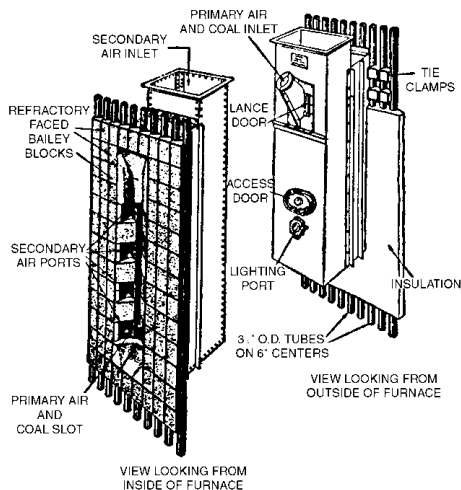
FIGURE 13.80 Calumet burner designed for firing through water-cooled furnace walls.

across the oil stream and atomize the fuel; steam required for atomization is usually between 2 and 3 percent of that generated. The capacity of a single steam-atomizing burner typically is 250 gal/h (946 L/h). In the rotary-cup burner, oil is thrown off the rim of a rotating cup and atomized; burners of this type typically burn 150 gal per hr (568 L/h). Mechanical burners are selected for the larger installations because of better economy; steam-atomizing burners are well suited to the burning of heavy sludges, and rotary-cup burners are widely used in large heating boilers. Energy-release rates of 60,000 Btu/ft^3 · h (2236 MJ/m^3 · h) of furnace volume may be efficiently attained in oil-fired furnaces.

Gas Burners.   In a gas burner, either the air, the gas, or both air and gas may be supplied under pressure. Many burners are of the venturi type in which the gas is supplied under a pressure that is adjusted to the load, while the air is supplied under atmospheric pressure; the rate of induced air flow is substantially proportional to the rate of gas flow, and the mixture ratio is maintained nearly constant. Other burners may depend upon furnace draft flow to draw in air around the gas jet.

BOILERS AND SUPERHEATERS

56. Boilers.   A complete steam-generating unit consists at least of a boiler and furnace and may include some or all of the following heat-transfer elements: superheater, economizer, air preheater, steam reheater, and auxiliaries.

Boilers are divided into two general classes: (1) fire-tube, and (2) water-tube. *Fire-tube boilers* are made in sizes up to about 6000 ft^2 (557 m^2) of heating surface. Fire-tube boilers are commonly selected for small installations where low first cost is more important than operating cost. For construction use, small, vertical, tubular boilers are built in sizes from 21 to 66 in (0.5 to 1.7 m) in diameter, with 40 to 1000 ft^2 (3.7 to 93 m^2) of surface, for pressures not exceeding 150 lb/in^2 (1034 kPa); large, vertical, tubular boilers are built in sizes ranging from 500 to 6000 ft^2 (46 to 557 m^2) for pressures up to 200 lb/in^2 (1379 kPa); stationary locomotive-type boilers are commonly built in sizes ranging from 150 to 2500 ft^2(232 m^2), and for pressures from 50 to 150 lb/in^2 (345 to 1034 kPa) the stationary Scotch marine boilers have outer shells that range from 6 to 16 ft (1.8 to 4.9 m) in diameter and have been built for pressure up to 300 lb/in^2 (2068 kPa). For power-plant use, the horizontal return-tubular boiler has been widely selected in sizes from 150 to 3750 ft^2 (14 to 348 m^2) of heating surface, with corresponding shell sizes from 36 in (diameter) × 8 ft (length) to 90 in × 20 ft (0.9 × 2.4 to 2.3 × 6.1 m); steam pressures seldom exceed 150 lb/in^2 (1034 kPa).

Water-tube boilers are made with a number of different arrangements of tubes and drums with steam-generating capacities from 10,000 to over 1,000,000 lb/h (4536 to 453,590 kg/h). They are used in central-station service and in industrial plants requiring more than about 10,000 ft^2 (929 m^2) of heating surface. The materials and construction of stationary boilers should conform to the ASME Boiler Construction Code and also to any inspection regulations to which the installation may be subject.

57. Capacity and Load.   The *capacity* of a steam-generating unit may be expressed in terms of: (1) the maximum rate of heat absorption by the unit, stated in kilo Btu (kB) per hour (W) or in mega Btu (mB) per hour (W) where mB = 1000

kB = 1,000,000 Btu; or (2) the maximum rte of steam generation, in pounds (kg) per hour, together with the pressure and temperature of the steam and the temperature of the feed water. The *load* of a steam-generating unit is the actual rate of heat absorption or steam generation at any time, expressed in the same unites as capacity. The older practice of measuring the boiler output in terms of boiler horsepower, where 1 boiler hp (0.75 W) is equivalent to a rate of heat absorption of 33,479 Btu/h (9813 W); of assigning a ''builder's rating'' of 1 boiler hp per 10 ft^2 (0.8 W/m^2) of heating surface; and of calculating the ''percent of rating'' by the quotient of the output to the ''builder's rating,'' is rapidly becoming obsolete.

58. Performance.   The performance of a steam-generating unit may be expressed in terms of its efficiency, an energy balance, or a combination of both. The *efficiency* of a steam-generating unit may be expressed as the ratio of the quantity of heat actually absorbed by the water and steam passing through the unit in a given time to the quantity of heat supplied by the fuel used in that time. In the form of an equation, the net efficiency is

$$\text{Eff.} = \frac{W_n(h_2 - h_1)}{W_f(\text{HHV})} \tag{13.53}$$

where W_n = net weight of steam delivered by the unit, in lb/h (kg/h)
             = gross weight minus auxiliary steam
       h_2 = enthalpy of steam at outlet from unit, in Btu/lb (kJ/kg)
       h_1 = enthalpy of feed water at inlet to first heating element of unit, in Btu/lb (kJ/kg)
       W_f = weight of fuel fired, in lb/h (kg/h)
     HHV = higher heating value of fuel as fired, in Btu/lb (kJ/kg)

Auxiliaries may consist of steam engines, steam turbines, steam jets, or electric motors. When the auxiliary is steam-driven, the steam consumed is subtracted directly from the gross weight delivered. If the auxiliary is not steam-driven, the equivalent weight of steam necessary to supply the auxiliary energy must be found and subtracted from the gross weight. Steam-generating units, in large sizes and with high load factors, may be expected to show monthly average net efficiencies of around 85 percent. Large industrial plants may show monthly average efficiencies of around 75 percent with lower values for the smaller plants.

59. Superheaters.   Superheaters increase the temperature of the steam without increasing its pressure. *Separately fired superheaters,* requiring separate settings, may be used for superheating steam exhausted from engines before use in low-pressure turbines, or for superheating a portion of the steam to a higher temperature than is required for the remainder. *Integral superheaters* are installed within the setting of the steam generator; they may be classified as *radiant* or *convection* superheaters, according to the predominant mode of heat transfer.

DRAFT AND DRAFT EQUIPMENT

60. Draft.   In steam-power-plant practice, draft usually means the pressure difference available to overcome the various resistances to the flow of air into and through the fuel bed and to the flow of products of combustion through the furnace. Draft

is usually measured in terms of the height, in inches (cm), of a column of water. If the water temperature is 80°F (26.7°C), 1 in (2.54 cm) of water is equivalent to a pressure of 0.036 lb/in² (0.25 kPa).

61. Natural Draft.   Draft produced by a chimney or stack is known as *natural* draft; draft produced by means of fans, blowers, or steam jets is known as *mechanical* drafts. Some small and medium-sized power plants use natural draft.

Large plants having steam-generating equipment which offers considerable resistance to the flow of gases and produces low exit-gas temperatures commonly use mechanical draft, but a chimney is then necessary to discharge the products of combustion at a suitable level.

62. Chimney.   Chimneys, or stacks, may be constructed of brick, reinforced concrete, or steel. The tallest chimneys are made of radial brick or reinforced concrete. *Self-supporting* steel stacks may be mounted on girders located near the roof level and supported by columns between the boilers. These stacks can be erected more rapidly, occupy less space, and cost less than masonry chimneys; they must be painted frequently to protect the steel from the air and from the corrosive action of the stack gases. *Guyed* steel stacks are seldom built in sizes larger than 72 in (183 cm) in diameter and 150 ft (45.7 m) high.

63. Mechanical Draft.   Large steam-power plants use both forced-draft and induced-draft fans. *Forced draft* fans force air from a closed duct into the furnace through the fuel bed. *Induced-draft* fans are placed between the boiler and the stack, and must handle gases that are at a high temperature and are frequently dust-laden. Although this duty is more severe than that of a forced-draft fan, the use of an induced-draft fan is often necessary in order to maintain a suction throughout the boiler setting. *Steam jets* are sometimes used to produce draft in small plants; forced draft may be produced when the jets are in the ashpit, and induced draft when the jets are in the flue or stack. They are low in first cost but have a large steam consumption and a limited capacity.

FEEDWATER, ACCESSORIES, AND PIPING

64. Feedwater Heaters and Economizers.   Feedwater heaters are of two general classes: (1) *open* heaters, in which the water comes into direct contact and mixes with the heating steam; (2) *closed* heaters, in which the steam and water are separated by metal surfaces through which heat is transferred. The *deaerating* heater is a special form of open heater which is designed to remove dissolved gases from engines or processes or steam extracted from turbines. The purpose of feedwater heating is to improve the thermal performance of the plant; a fuel saving of approximately 1 percent will result from each 12°F (6.7°C) rise in temperature of the feedwater if the steam used in the heater would otherwise be wasted.

Other heat exchangers may be incorporated in the feedwater system to supplement feedwater heaters. *Economizers* are closed heaters that utilize the heat in the exit gases to heat the feedwater. Economizers may be *independent* and located above or behind the boiler; or *integral,* i.e., located within the boiler setting and connected directly to the boiler circulating system.

65. Boiler-Water Conditioning.   Untreated water contains solids and dissolved gases that may cause scale formation, corrosion, caustic embrittlement, foaming, or priming.

Water conditioning may consist of sedimentation or filtration to reduce suspended solids, softening by chemical treatment to reduce formation of hard scale, evaporation of makeup water to reduce the concentration of suspended solids and dissolved gases, blowdown to reduce the concentration of solids, and deaeration to reduce the concentration of dissolved or mechanically entrained gases to minimize corrosion.

66. Boiler Accessories.   Boiler accessories are those devices which are directly connected to the boiler for the purpose of giving safe and convenient operation. Certain accessories are required by law, and others are used for improving performance. Accessories commonly provided are a steam-pressure gage, water column, safety valve, stop-and-check valve, fusible plug, feedwater regulator, blowoffs, soot blowers, and tube cleaners.

STEAM TURBINES

67. General Characteristics.   Steam turbines are built in sizes from less than 1 kW to over 200,000 kW. The small requirements of space per unit of capacity, the enormous capacities possible, the high rotative speed, and close regulation make turbines particularly suitable as prime movers for central-station generation of electrical energy. The large specific volume of steam at low pressure may be handled more satisfactorily by a turbine than by a reciprocating engine, and back pressures of under 1 in Hg (2.5 cm Hg) abs are successfully used. Highly superheated steam may be used without lubrication difficulties, and the exhaust steam is free from oil.

68. Types.   Steam turbines are divided into two general classes: (1) *impulse* turbines, in which the expansion of the steam occurs only in the stationary blades or nozzles; (2) *reaction* turbines, in which the steam expands in both the stationary and the moving blades.

In *impulse* turbines, the energy made available by the drop in pressure of the steam as it passes through the nozzles or stationary blades imparts velocity to the steam. The steam jet strikes the blades on a rotating disk and transfers its velocity energy to them. *Velocity staging* or velocity compounding is the principle illustrated in Fig. 13.81.

69. Reaction Turbine.   The *reaction turbine,* frequently called the *Parsons type,* is a multipressure-stage machine each stage consisting of a set of fixed blades or nozzles followed by a set of rotating blades, as indicated in Fig. 13.82. The pressure drop is continuous and gradual from supply to exhaust; in any one stage it is small, seldom over 3 lb/in^2 (21 kPa), and consequently the number of stages is large. The stationary blades extend around the entire circumference because partial peripheral admission is impossible. The clearances must be small to minimize steam leakage around the blades to a lower pressure region. The leakage loss diminishes toward the exhaust end.
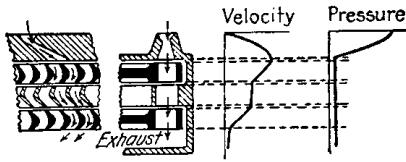
FIGURE 13.81 Pressure and velocity changes in an impulse turbine having one pressure stage and two velocity stages.
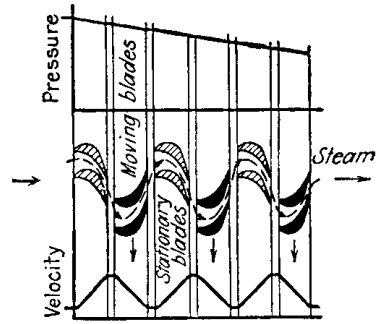
FIGURE 13.82 Pressure and velocity changes in a reaction turbine.

70. Multicylinder or Compound Turbines. The increase in capacity of steam-turbine units has led to the development of compound machines having two or three cylinders. After partial expansion of the steam in the high-pressure cylinder, the steam passes to the next cylinder, where the process is continued.

71. Back-Pressure, Low-Pressure, and Mixed-Pressure Turbines. Turbines operated noncondensing, in connection with building heating or industrial processes, are frequently called *back-pressure* turbines. Since the heat in the exhaust can be usefully applied, the power from the machine is obtained at very low cost. *Low-pressure* turbines are supplied only with exhaust steam from other apparatus and should exhaust into high vacuum. *Mixed-pressure* turbines are designed to operate on low-pressure steam, supplemented by high-pressure steam when the low-pressure supply becomes insufficient to carry the load.

72. Extraction or Bleeder Turbines. *Extraction* or *bleeder* turbines are supplied with high-pressure steam and operated condensing, but they have the distinctive feature of permitting the withdrawal of steam from the machine at one or more intermediate points between supply and exhaust, for use in building heating, feed-water heating, or manufacturing processes.

73. Turbine Performance. The best measure of comparative performance of steam turbines is the engine efficiency; brake engine efficiency is the ratio of the brake thermal efficiency of the turbine-generator to the thermal efficiency of the corresponding ideal engine (Rankine engine if there is no reheating or regenerative feedwater heating). Typical values of the brake engine efficiency of steam turbines are given in Table 13.4. These values are not exact, and manufacturers' guarantees should be obtained for new units.

CONDENSING EQUIPMENT

74. General. The object in applying condensing equipment to steam prime movers is to improve their economy and increase their capacity by lowering the back pressure against which they exhaust. For a given condition of steam supply, a

TABLE 13.4   Brake Engine Efficiencies of Steam Turbines

Brake hp	kW	Brake engine efficiency, %
100	74.6	50
500	373.0	58
1,000	746.0	63
5,000	3730.0	72
10,000	7460.0	76
20,000–100,000	14,920–74,600	80–83

reduction in back pressure increases the available energy per pound (kg) of steam. When the back pressure is reduced to 1 in Hg (2.54 cm Hg) [1 29-in (73.7-cm) vacuum referred to a 30-in (76.2 cm) barometer], the temperature of the exhaust steam is only about 80°F (26.7°C).

Condensers are of two general types: (1) *surface* and (2) *direct contact*. Both types use cooling water to change the exhaust vapor to a liquid of much smaller volume, but the surface condenser keeps the two fluids separate while the direct-contact condenser brings them together in a mixture. The selection of type of condenser depends upon: (1) water quality and abundance, (2) fixed charges on condenser and auxiliaries, (3) power cost for condenser pumps, (4) maintenance expense, and (5) space requirements.

75. Surface Condensers.   This type of condenser consists of a shell surrounding a compact nest of tubes. Cooling water is circulated through the tubes, and exhaust steam fills the shell surrounding them. Tube sizes are usually ¾ in (1.9 cm), ⅞ in (2.2 cm), or 1 in (2.5 cm), and 0.049 in (0.12 cm) (No. 18 B.W.G.) thick; ⅝-in (1.59-cm) tubes are sometimes used in small condensers. The overall coefficient of heat transfer in surface condensers depends primarily upon the water velocity, cleanliness of the tubes, and the amount of noncondensible gases in the steam. Economic water velocities are usually between 6 and 8 ft/s (1.8 and 2.4 m/s). The heat balance for a surface condenser is given as follows:

$$W_s(h_s - h_c) = W_w(t_2 - t_1) = UA\theta_m \tag{13.54}$$

where W_s = weight of steam condensed, in lb/h (kg/h)
h_s = enthalpy of exhaust steam entering the condenser, in Btu/lb (kJ/kg)
h_c = enthalpy of condensate leaving condenser, in Btu/lb (kJ/kg)
W_w = weight of cooling water flowing through condenser, in lb/h (kg/h)
t_2 = temperature of exit cooling water, in °F (°C)
t_1 = temperature of entering cooling water, in °F (°C)
U = overall coefficient of heat transfer, in Btu/h · ft² · °F (W/m² · °C) (for rough approximations, a value of 350 may be used for USCS)
A = area of tube surface, in ft² (m²)
θ_m = mean temperature difference between steam and water, °F (°C)

With single-pass condensers, the condensing surface provided for turbine installations is usually between 0.6 and 1.0 ft²/kW (0.06 and 0.09 m²/kW), while for two-pass condensers between 0.7 and 1.4 ft²/kW (0.07 and 0.13 m²/kW) is common; the larger figure in each case is for small turbines. The water circulated is commonly between 1.0 and 2.0 gal/min · ft² (0.7 and 1.4 L/s · m²).

76. Direct-Contact Condensers. This group includes (1) jet, (2) barometric, and (3) ejector condensers. In jet and barometric condensers the exhaust steam and cooling water enter near the top, the water being discharged into the steam in a fine spray. The resulting mixture of condensed steam, warmed cooling water, and air must be pumped from the condenser unless it is of the barometric type. An independent air-removal pump is necessary for high vacuum. The barometric condenser is elevated sufficiently above the hot well to cause the water to flow out by gravity against atmospheric pressure. A barometric condenser requires a circulating pump and sometimes a dry air pump. Ejector condensers operate on the same principle as a steam ejector. No removal pump is need. High velocity of the fluids is utilized to discharge the mixture of condensed steam, cooling water, and air against atmospheric pressure.

77. Recooling Condensing Water. When cooling water is deficient in quantity or quality, cooling ponds, spray cooling systems, or cooling towers may be used to cool the water after it has passed through the condenser. In this way the same cooling water may be used repeatedly with a makeup of from 2 to 8 percent to replace that lost by evaporation.

Cooling ponds are so constructed as to expose a large area to the atmosphere. The water is cooled by surface evaporation. The necessary pond area may be estimated at 8 ft^2 for each lb (1.64 m^2/kg) of steam condensed per hour. A cooling effect of about 4 Btu/h per °F (22.7 W/m^2 per °C) temperature difference between the water and the air will be obtained, per square foot of pond area, in the summer, and about 2 Btu/h · °F (11.4 W/m^2 + °C) in the winter.

Spray cooling systems accomplish the same results as cooling ponds but require much less area. The cooling water is piped to a system of nozzles through which the water is sprayed into the air. It falls back into a basin below. The amount of water surface exposed to the atmosphere is thus greatly increased. A typical spray cooling system with a capacity of 4200 gal/min at 7 lb/in^2 pressure (265 L/s at 48 kPa) discharged through 105 nozzles attached to 2-in (50.8-mm) pipes requires a pond or basin 138 ft long and 100 ft wide (42 × 30 m). The nozzles are in 21 standard clusters of 5 each; the clusters are in 3 rows; the rows are 25 ft (7.6 m) apart; each row has 7 clusters on 13-ft (3.9-m) centers. The reduction in temperature by spraying depends on the atmospheric conditions but is usually 15 to 20°F (.8.3 to 11.1°C)

A *cooling tower* is a structural assemblage of cooling surface which may conveniently be placed on the top of a building or in an open field. The water may be cooled, as an ideal limit, to the wet-bulb temperature of the air, but cooling to within 5°F (2.8°C) of this temperature represents good practice. A cooling tower is commonly guaranteed by the manufacturer to cool a prescribed weight of water per unit of time through a definite temperature range at a stated atmospheric temperature and relative humidity.

REFRIGERATION ENGINEERING

REFRIGERATION MACHINES AND PROCESSES

Refrigeration is a special aspect of heat transfer and involves the production and utilization of below-atmospheric temperatures by a number of practical process.

Substances are cooled when their heat is transferred, via a temperature drop, to solid, liquid, or gaseous media which are naturally or artificially colder, their lower temperature stemming from radiation, sensible- or latent-heat physical effects, or endothermic chemical, thermoelectric, or even magnetic effects. Effects, such as cold streams, melting ice, and sublimating solid carbon dioxide, are included.

1. Units of Refrigeration. In the United States, a *standard ton of refrigeration* corresponds to a heat absorption at a rate of 288,000 Btu/day or 200 Btu/min (3.5168 kW). The heat absorption per day is approximately the heat of fusion of 1 ton (907.19 kg) of ice at 32°F (0°C). The *standard rating* of a refrigerating machine, using a condensable vapor, is the number of standard tons of refrigeration it can produce under the following conditions: (1) liquid only enters the expansion valve and vapor only enters the compressor or the absorber of an absorption system; (2) the liquid entering the expansion valve is subcooled 9°F (5°C) and the vapor entering the compressor or absorber is superheated 9°F (5°C), these temperatures to be measured within 10 ft (3.05 m) of the compressor cylinder or absorber; (3) the pressure at the compressor or absorber inlet corresponds to a saturation temperature of 5°F (−15°C); and (4) the pressure at the compressor or absorber outlet corresponds to a saturation temperature of 86°F (30°C).

The *British unit of refrigeration* corresponds to a heat-absorption rate of 237.6 Btu/min (4.175 kW) with inlet and outlet pressures corresponding to saturation temperatures of 23°F (−5°C) and 59°F (15°C), respectively. On the European continent a unit of refrigeration capacity called the *frigorie* is used. A frigorie is approximately equivalent to 50 Btu/min (0.8786 kW), or ¼ of a standard ton of refrigeration.

PROPERTIES OF REFRIGERANTS

Refrigerants are the transport fluids which convey the heat energy from the low-temperature level to the high-temperature level, where it can, in terms of heat transfer, give up its heat. In the broad sense, gases involved in liquefaction processes or in gas-compression cycles go through low-temperature phases and hence may be termed *refrigerants,* in a way similar to the more conventional vapor-compression fluids.

Refrigerants are *designated by number.* The identifying number of the refrigerant, or the word *Refrigerant* or both, may be used in conjunction with the trade name.

OVERALL CYCLES

Figure 13.83 represents the simple closed-circuit *vapor-compression* system. The upper heat exchanger is a vapor condenser (with some superheat), and after the throttling expansion valve, the lower one is the evaporator in which the refrigerant liquid at reduced pressure and temperature evaporates with the inward refrigeration heat flow. The refrigerant vapor is elevated by the compressor to a higher pressure and condensing temperature so that it will liquefy in its transfer of heat to the atmospheric level. Figure 13.84 shows comparative temperatures for the cycle.
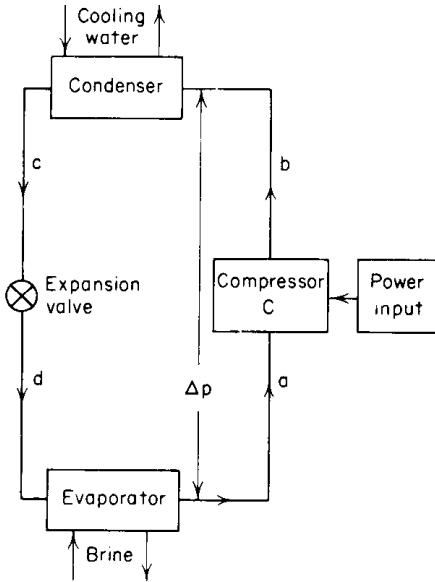
FIGURE 13.83 Simple closed-circuit vapor-compression system.
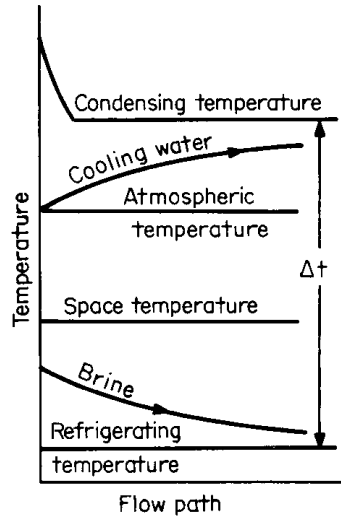
FIGURE 13.84 Comparative temperature relations for fluids of a simple system.

COMPONENTS OF COMPRESSION SYSTEMS

A refrigeration system is an energy-transport complex of assorted components. Components include vapor and gas compressors; liquid pumps; heat-transfer equipment (gas coolers, intercoolers, aftercoolers, exchangers, economizers); vapor condensers and the counterpart evaporators; liquid coolers and receivers; expanders; control valves and pressure-drop throttling devices (capillaries, refrigerant-mixture separating chambers, stream-mixing chambers); and connecting piping and insulation. Compression may be single stage (Fig. 13.83) or multiple stage.

Air-cooled compressors are used where discharge temperatures are low, such as with Refrigerant 12; *water cooling,* where discharge temperatures are high, as with ammonia. *Oil separators* return to the compressor the lubricating oil carried over by the refrigerant vapors. *Rotary compressors* of the vane, eccentric, gear and screw types are in use.

Kinetic compressors include high-speed centrifugal and axial flow machines, usually multistaged, and jet-entrainment devices. Centrifugal machines are especially adapted to high-volume flow ($>$ 500 ft³/min) (14.2 m³/min).

The entire machine-compression operation may be replaced by a secondary absorber-pump-generator system, in which the complex is known as an *absorption refrigeration system.*

Dual (or *multiple-effect*) *compression* may be used when refrigeration at two temperatures is desired. The compressor takes vapor from a lower temperature expansion coil during the first part of its intake stroke, and from a higher temperature expansion coil at or near the end of the stroke. The mixture is then compressed and condensed.

Condensers are usually shell-and-tube type, with the refrigerant passing outside the tubes. Older industrial installments still use *double-pipe condensers.* In double-pipe condensers, gas flows between the two pipes while water passes through the

inner. The outside pipe diameter may be 2 in (5 cm); the inner, 1¼ in (3.2 cm). In some instances, the pipes are exposed to the atmosphere either with or without water drip. In an *evaporative condenser* the refrigerant vapor is condensed as it passes through tubes over which water is sprayed; the water is then evaporated by air flowing over the wet tubes. In this way, cooling-water requirements are reduced to from 5 to 15 percent of the water requirements of a nonevaporative condenser with no water reuse. The evaporative condenser combines in a single unit a refrigerant condenser and the atmospheric cooling tower or spray pond which is required if the cooling water is to be reused. Air-cooled finned-tube condensers with forced ventilation are widely used on small units, and shell-and-coil or double-tube water-cooled condensers on medium units.

Evaporators may have finned or plain surfaces. Defrosting may be automatic or manual. *Flooded* evaporators are operated practically full of liquid refrigerant, the level being controlled by a float valve. *Wet-expansion* evaporators are operated with a level approaching that for flooded operation; *dry-expansion* (once-through) units operate with an indefinite amount of liquid in the evaporator.

Steam-jet refrigeration systems are used where cooling to temperatures above 32°F (0°C) is desired. Applications include industrial air conditioning and cooling of city gas to condense out tar and other objectionable impurities, of gas absorbers to increase efficiency of absorption, of reaction units where heat removal and temperature control are important during chemical transformations, and of wort and mash in the brewing and other fermentation idustries particularly in the summer months.

ABSORPTION SYSTEMS

Absorption refrigeration machines are essentially vapor-compression plants (Fig. 13.83) in which the mechanical compressor has been replaced by a thermally activated arrangement (Fig. 13.85). The basic elements are the *absorber, pump, heat exchanger, throttle valve,* and *generator.*

A typical absorption machine requires about 20 lb (9 kg) of steam for 1 ton (3.5 kW) of refrigeration effect at 45°F (7°C).

THERMOELECTRIC COOLING

Thermoelectric cooling utilizes the *Peltier effect* whereby a temperature differential will occur between two junctions of a closed loop of dissimilar metallic conductors when an electric current is imposed. It is variously called *Peltier effect cooling, thermoelectric refrigeration,* and *electronic cooling.* It is the inverse of the thermocouple operating principle whereby a temperature differential across the junctions of two dissimilar metals produces an electric current.

The capacity of a thermoelectric couple is small. System capacity can be increased by increasing the number of elements used. The temperature produced can be lowered by cascading the systems. Capacities approaching 25 tons (87.9 kW) cooling effect have been achieved, and temperatures as low as -266°F (-166°C) have been attained.
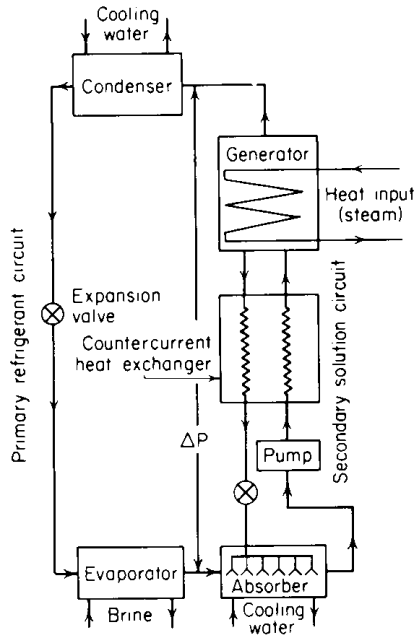
FIGURE 13.85 Elemental absorption system circuit.

METHODS OF APPLYING REFRIGERATION

In *direct expansion systems* the evaporator is placed in the space which is to be cooled; in *indirect systems* a secondary fluid (*brine*) is cooled by contact with the evaporator surface, and the cooled brine goes to the space which is to be refrigerated. Brine systems require 40 to 60 percent more surface than do direct expansion; they have an equalizing effect due to the large heat capacity of cold brine, they are safer (particularly if the refrigerating effect must be carried considerable distance or widely distributed), and they permit closer temperature regulation than is possible with direct expansion. If two temperatures are to be held, the lower may be by direct expansion, the higher by brine cooling. Development of better controls and newer piping methods has made direct expansion more attractive than previously.

Brines used for industrial refrigeration are usually aqueous solutions of calcium chloride, ethylene glycol, propylene glycol, or undiluted methylene chloride.

Brine coolers may be of three types: shell-and-tube, shell-and-coil, and double pipe. The shell-and-tube type most widely used is brine flowing through the tubes which are surrounded by the evaporating refrigerant. Tubes may be arranged for multipass operations. Effective heat-transfer surface varies from 8 to 15 ft²/ton (0.21 to 0.4 m²/kW), varying with temperature and brine velocity. A submerged coil in an open brine tank is used for ice making by the can process.

The *double-pipe cooler* is usually of 2-in (50-mm) inner or brine-flow pipe and 3-in (75-mm) outer pipe. The commercial rating is 15 to 20 ft (5 to 6 m) length of coil per ton of refrigeration.

The *shell-and-tube cooler* is used with closed heads and is erected both vertically and horizontally; brine flows through the tubes and ammonia is in the shell. It is

made in sizes from 1 to 350 tons with ratings of 8 to 15 ft^2 (0.7 to 1.4 m^2) effective surface per ton, varying with the temperature and brine velocities; tubes 1 to 2½ in (25 to 63 mm) arranged multipass. This type of cooler has largely displaced all other types in recent installations.

REFRIGERANT PIPING

It is important to the proper operation of a refrigeration system that the piping or mains interconnecting the compressors, condensers, evaporators, and receivers be properly sized. This piping must be considered in three categories, viz., liquid lines, suction lines, and discharge lines. Essentially pipeline sizing is governed by pressure drop, first cost, operating cost, noise, and oil entrainment. Excessive pressure drops penalize compressor efficiencies and may affect control-valve operation adversely. Liquid-line velocities for most refrigerants are in the order of 60 to 400 ft/min (18 to 122 m/min), suction lines from 700 to 4600 ft/min (38 to 250 m/s), and discharge lines from 1000 to 5000 ft/min (213 to 1402 m/min). Pressure drops vary approximately as the square of the velocity (or tonnage) and directly as the length of the piping. For liquid lines, when evaporator is located above condenser, the following pressure drops in pounds per square inch per foot (kilopascals per meter) of static lift should be allowed: ammonia, 0.26 (5.9); Freon 12, 0.57 (12.9); Freon 22, 0.51 (11.5); Freon 11, 0.64 (14.5).

COLD STORAGE

2. Insulation.   In recent years newer insulating materials have largely supplanted the use of plaster and corkboard in the cold-storage room construction. Insulation has generally been used in three forms for this application: board form, panelized, and poured, sprayed, or foamed in place.

Board form insulation includes corkboard, rigid polystyrene, polyisocyanurate, and polyurethane, foamed, and fibrous glass.

Panelized insulation has gained acceptance, particularly in prefabricated installations. The outer metal skin can serve as a vapor barrier, and the panels are factory insulated with polystyrene, polyisocyanurate, polyurethane, or fibrous glass. *U* values as low as 0.035 have been attained in 4-in (100-mm) thick panels.

3. Cold-Storage Temperatures.   A great deal of research and experience are required to obtain authoritative information on optimum temperatures and humidities for various products in cold storage. The safe storage period depends upon the product and the storage temperature, and operations techniques vary greatly. Modern cold-storage warehouses of the larger concerns are cooled by brine which is furnished at two different temperatures only. The higher temperature for the mild-temperature warehouses is 10 to 12°F (−12 to −11°C), and the low-temperature brine for the freezers is −10 to −12°F (−12 to −24°C). All temperatures above that of the brine are obtained by regulating the amount of brine circulated in any particular set of coils. In the low-temperature warehouses, the piping is arranged for two classes of service: (1) *sharp freezers,* where the goods which are to be frozen are kept while their temperature is brought down quickly to the holding

temperature (say, in from 6 to 10 h) after which they are stored in (2) *holding rooms* where the desired temperature is maintained.

4. Piping of Rooms. The size of pipe usually employed for piping rooms varies from 1 to 2 in (25 to 50 mm) with either brine circulation or direct expansion.

The extra cost of liberal piping allowance will often be offset by the consequent improvement in the efficiency of operation of the compressor. An expansion valve should be provided for every 500-ft (150-m) length of 1-in (25-mm) pipe, every 650 ft (200 m) of 1¼-in (30-mm) pipe, and every 1000 ft (300 m) of 2-in (50-mm) pipe when direct expansion is used.

CRYOGENICS

Cryogenics is the study, production, and utilization of low temperatures. Cryogenic temperatures have been defined so ambiguously that "upper limits" to the cryogenic range from 216 to 396°R (120 K to 220 K) may be found in the literature. In this section the cryogenic range for a given property is considered to embrace the scale between absolute zero and the temperature above which the property has the expected or normal behavior. Cryogenics thus embraces the unusual and unexpected variations which appear at low temperatures and make extrapolations of properties from ambient to low temperatures unreliable.

Progressively lower temperatures become increasingly difficult to attain in practice. As the working temperature of a refrigerator is lowered, the work required to transfer a given amount of heat increases as demonstrated by the Carnot limitation, to wit, $W = Q[(T_1 - T_2)/T_2]$, where W is the work required to extract the heat Q at a low temperature T_2 and reject it at a higher temperature T_1. The actual work is always greater than this because of the inefficiencies of mechanical equipment, thermal losses associated with finite temperature differences in heat exchangers, and heat leaks from the surroundings to the cold equipment.

INDUSTRIAL AND MANAGEMENT ENGINEERING

The American Institute of Industrial Engineers defines industrial engineering as a branch of engineering "concerned with design, improvement, and installation of the integrated systems of people, material, equipment, and energy. It draws upon the specialized knowledge and skills in the mathematical, physical and social sciences together with the principles and methods of engineering analysis and design to specify, predict, and evaluate the results to be obtained from such systems."

ACTIVITIES

Industrial engineering is similar to the other major engineering specializations (civil, mechanical, electrical, and chemical) in that it is concerned with analysis and de-

sign, and applying the laws and materials of nature to useful and constructive purposes. It is different from these other fields of engineering in that it is specifically concerned with equipment and systems in which people are in integral part. The industrial engineer must be able to use mathematics, materials, machinery, devices, chemistry, electricity, electronics, and so on, just as all of the other types of engineers. But, unlike them, the industrial engineer must also understand and be able to integrate people into his or her designs, and must know their physical, physiological, psychological, and other characteristics—singly and in groups.

The early, and still major, activities of industrial engineers include work methods analysis and improvement, work measurement and the establishment of standards, job and workplace design, plant layout, materials handling, wage rates and incentives, cost reduction, suggestion evaluation, production planning and scheduling, inventory control, maintenance scheduling, equipment evaluation, assembly line balancing, systems and procedures, overall productivity improvement, and special studies—all done, almost exclusively In the early days, in manufacturing industries. As the technology evolved, additional activities were added to this list. These include machine tool analysis; numerically controlled machine installation and programming; computer analysis, installation, and programming; linear programming; queuing and other operations research techniques; simulations; management information systems; value analysis; human factors engineering; human/machine system design; ergonomics; biomechanics; and the use of robots and automation. Moreover, it was found that the industrial engineering techniques used successfully in the factory could also be applied in the office, laboratory, classroom, hospital (including the operating room), the government, the military, and other nonindustrial areas.

The industrial engineer of today is using computers more in his or her production and test equipment, controls and systems, and for analyses and special studies. Minicomputers and microprocessors are in wide and growing use.

Computers are already being used to eliminate much of the calculation drudgery of work measurement. The power of the computer also facilitates the synthesis of various methods configurations.

Computer-aided manufacturing (CAM) makes use of the computer in an attempt to develop a total systems approach to the manufacturing process. It links together computer-aided design (CAD), automatically programmed tools, work measurement, sequencing, materials handling, and inventory control. Computer interactive graphics are also being introduced in the design and analysis functions.

Numerical control (N/C) is the term usually associated with automatically programmed equipment. An N/C machine consists of a reader, a control unit, and the machine tool. Preprogrammed machine instructions on tape are transmitted to the control unit, which interprets the instructions and causes the machine tool to execute. While N/C has been used primarily for metal removal machinery, it has also been applied successfully in material storage, assembly, and packaging.

The capability of N/C can be enhanced by direct numerical control (DNC), which uses an on-line computer to provide the instructions directly to the machine. Further extensions are in the area of adaptive control systems and robots.

The production function is dynamic by its inherent activities, and traditional industrial engineers are frequently called upon to determine the impact of worker power changes, product mix variances, and equipment additions or deletions on facilities arrangement and line balancing in the manufacturing area. This task becomes increasingly more complex today with the additional energy computation and pollution level considerations. This complicated system of interrelated activities is almost impossible to evaluate manually; however, through the use of simulation

on a computer, answers can be given to "what if" types of questions which provide significant input other decision maker.

The age-long inventory control function can now be monitored constantly by a real-time computerized system. The computer greatly facilitates the storage and retrieval of historical and current information on the status of inventory and concomitant purchase orders. Computers can be programmed to evaluate the trade-offs between the savings in inventory investment and the activity required in the planning and control phases.

Plant layout is considered a classical industrial engineering activity; however, the traditional approach to facilities design has been supplemented by the use of computers in the design and calculation process of a layout, and also by the application of mathematical modeling and optimization to those problems. It must be emphasized that the basic concepts of plant layout are still important, and the computer and operations research methods must be skillfully blended with the traditional methods in order to obtain better answers to facilities problems which become increasingly more complex—for example, by inclusion of accident risk factors.

SYSTEMS CONCEPT

The trend in industrial engineering is from micro- to macrolevel structures, such as a computer-aided manufacturing system, a total information system design of a company, or the complex materials-handling function involving many interrelated departments.

Management of industrial, business, or service activities is growing increasingly complicated as the sciences and humanities become more interdependent. These organizations need not only management personnel, but individuals capable of designing and installing new integrated systems of people, materials, and equipment which will function effectively in a society made more complex by the technological explosion. Today's industrial engineer has a foundation in the basic physical and social sciences, engineering, and computers, and a skill in systems analysis and design which crosses traditional disciplinary lines. The result is a capability to meet the demands imposed by a dynamic social system.

The logical structure of systems study may be outlined as follows:

1. Indentification of the components of a complex system. Examples include a production scheduling system, a computer software system, an educational scheduling system, a project planning and control system for underdeveloped countries, a hospital information system, and a space information system.

2. Development of the topological properties of the system, and the generation of mathematical models and analogs, or the adaptation of a heuristic approach when desirable.

3. Establishment of the functional relationships between the variables of the system, together with the necessary feedback required to control the operational system.

4. Selection and evaluation of optimization criteria. Examples of criteria are profit, costs, idle capacity, energy consumption, and information entropy.

5. Analysis and design of the nondeterministic functions to make them amenable to solution of the total system structure.

6. Integration of behavioral patterns in different environments, as required, within the physical framework of the system. Examples are the sociopolitical environment of a country for which the system is designed, the behavioral attitudes of nurses in a hospital, and those of production line workers in a fabrication shop.

7. Design of simulation models which permit a rigorous critique of the parameters of a dynamic system.

8. Economic analysis of the total system and concomitant subsystems with extensions into alternative structure.

OPERATIONS RESEARCH AND INDUSTRIAL ENGINEERING DESIGN

The real-world problems are complex systems configurations which require the use of more sophisticated methods than previously to present meaningful results to the decision maker. In the development of modern systems technology, the pacing factor is management. This is the function which directs, coordinates, and controls the many facets of a system. Technological advancement depends on optimal use of all resources and requires that new discoveries be translated into new products in minimal time. Productivity must be increased. The terms *optimal* and *minimal* are characteristics of management science and operations research.

Management science and operations research are often used synonymously since the tenor of their objectives is compatible. Operations research has been referred to as a sharper kind of industrial engineering. Even though no two definitions of operations research are exactly the same, the following four common denominators can be abstracted from most definitions: (1) formulation of objectives—seek to attain goals established by management; (2) systems perspective—this is frequently concerned with the interrelationship among many components; (3) alternatives—choices exist among decision variables; (4) optimization—attempt to make best decision relative to cited objectives, recognizing that in the real world this is not usually achieved in the strict sense.

In the current identification of engineering design and operations research, two apparently diverse disciplines, their components, and processes are found to be very similar. Industrial engineers are involved with both the design and analysis aspects of engineering in solving many of their problems. Engineering design is the process of devising a system, which includes its components and processes, to meet desired needs. It is a decision-making process (often iterative and interactive) in which the basic sciences, mathematics, and engineering sciences are applied to convert resources optimally to meet a stated objective. Among the fundamental elements of the design process are the establishment of objectives and criteria, synthesis, analysis, construction, testing, and evaluation.

The criteria established for engineering design evaluation provide a powerful impetus to utilize operations research foundations in engineering problems. It is important to note that in industrial engineering operations research is currently being used in solving the problems. The data in Table 13.5 provide insights relative to the use of operations research techniques in the production function, which is a major area of activity for industrial engineers. The techniques used most frequently are linear programming and simulation. Network models have been used extensively in project planning and control, and regression analysis in inventory and quality control.

TABLE 13.5 Application of Operations Research Techniques in Production*

Application areas	Linear programming	Dynamic programming	Network models	Simulation	Queueing theory	Game theory	Regression analysis	Others
Production scheduling	30 (41.1)	7 (9.6)	6 (8.2)	26 (35.6)	9 (12.3)	0 (0.0)	5 (6.8)	10 (13.7)
Production planning and control	19 (26.0)	3 (4.1)	7 (9.6)	18 (24.7)	4 (5.5)	0 (0.0)	3 (4.1)	3 (4.1)
Project planning and control	10 (13.7)	1 (1.4)	28 (38.4)	9 (12.3)	2 (2.7)	0 (0.0)	0 (0.0)	3 (4.1)
Inventory analysis and control	11 (15.1)	3 (4.1)	3 (4.1)	27 (37.0)	4 (5.5)	1 (1.4)	12 (16.4)	7 (9.6)
Quality control	2 (2.7)	0 (0.0)	1 (4.1)	2 (2.7)	0 (0.0)	0 (0.0)	15 (20.5)	9 (12.3)
Maintenance and repair	0 (0.0)	1 (1.4)	3 (4.1)	8 (11.0)	3 (4.1)	1 (1.4)	4 (5.5)	3 (4.1)
Plant layout	13 (17.8)	0 (0.0)	5 (6.8)	19 (26.0)	5 (6.8)	1 (1.4)	2 (2.7)	3 (4.1)
Equipment acquisition and replacement	4 (5.5)	0 (0.0)	1 (1.4)	11 (15.1)	1 (1.4)	0 (0.0)	0 (0.0)	7 (9.6)
Blending	32 (43.8)	0 (0.0)	1 (1.4)	6 (8.2)	1 (1.4)	0 (0.0)	3 (4.1)	1 (1.4)
Logistics	27 (37.0)	1 (1.4)	8 (11.0)	24 (32.9)	3 (4.1)	2 (2.7)	6 (8.2)	2 (2.7)
Plant location	32 (43.8)	2 (2.7)	8 (11.0)	23 (31.5)	1 (1.4)	0 (0.0)	5 (6.8)	4 (5.5)
Other	7 (9.6)	1 (1.4)	2 (2.7)	7 (9.6)	1 (1.4)	1 (1.14)	3 (4.1)	4 (5.5)

*All data are expressed in terms of numbers (percent). The percentages do not total 100% because many respondents indicated they used more than one technique in a given application area.

Source: From A. N. Ledbetter and J. F. Cox, ''Are OR Techniques Being Used?'' *Ind. Eng.*, p. 21, February 1977.

CHAPTER 14
CIVIL ENGINEERING AND HYDRAULIC ENGINEERING

CIVIL ENGINEERING

SURVEYING

1. Measurement of Distance

Units of Measurement. Distances are usually measured in feet and tenths, hundredths, and (for accurate work) thousandths of feet (meters, centimeters, and millimeters). For many older surveys, distances were measured in chains and links. A chain is 66 ft (20.1 m) in length and is divided into 100 links, each 7.92 in (201.2 mm) long. The metric system, in which the unit of distance is the meter and its decimal fractions or multiples, is in use in most countries and is expected to replace the English system of measurement in the United States. See Chap. 1 for conversion factors.

Four methods used for the direct measurement of distance are pacing, stadia reading, taping, and electronic distance recording.

Pacing. Pacing is a rapid means of checking more accurate measurements of distance. The precision of pacing under average conditions is from 1:100 to 1:200.

Stadia Reading. The use of stadia furnishes a rapid method of determining distances with a fair degree of accuracy. Under average conditions, a precision of from 1:300 to 1:1000 can be obtained.

Specific publications consulted during the preparation of this chapter's text include: American Association of State Highway and Transportation Officials (AASHTO) "Standard Specifications for Highway Bridges"; American Concrete Institute (ACI) "Building Code Requirements for Reinforced Concrete"; American Institute of Steel Construction (AISC) "Manual of Steel Construction," "Code of Standard Practice," and "Load and Resistance Factor Design Specifications for Structural Steel Buildings"; American Railway Engineering Association (AREA) "Manual for Railway Engineering"; American Society of Civil Engineers (ASCE) "Ground Water Management"; American Water Works Association (AWWA) "Water Quality and Treatment." In addition, the author consulted several hundred other standards, codes, and other official publications, along with well-respected texts and handbooks. Especially helpful were:

A Policy on Geometric Design of Rural Highways, AASHTO, Washington, D.C.; *A Policy on Arterial Highways in Urban Areas,* AASHTO, Washington, D.C.

Highway Capacity Manual, Transportation Research Board—National Academy of Sciences, Washington, D.C.

Specification for the Design, Fabrication and Erection of Structural Steel for Buildings, American Institute of Steel Construction, Chicago, Ill.

Concrete Reinforcing Steel Handbook, Concrete Reinforcing Steel Institute, Detroit, Mich.

General Engineering Handbook, 2d ed., by C. E. O'Rourke. Copyright © 1940. Used by permission McGraw-Hill, Inc. All rights reserved. Updated and metricated by the editors, 2002.

When using any of the formulas in this chapter that may come from an industry or regulatory code, the user is cautioned to consult the latest version of the code. Formulas may be changed from one edition of a code to the next. In a work of this magnitude it is difficult to include the latest formulas from the numerous constantly changing codes.

Measurement with Tape. The most commonly used method of determining distance is by measurement with a tape. Steel tapes, ranging in length from 50 to 300 ft (15.2 to 91.4 m), are generally used, but tapes of other materials may be used where accuracy is not essential. The precision of a tape measurement depends on the degree of refinement with which the measurement is made. The precision of taping ordinarily used in surveys is from 1:3000 to 1:5000.

For ordinary taping, a tape accurate to 0.01 ft (0.003 m) should be used. The tension of the tape should be about 15 lb (66.7 N). The temperature should be determined within 10°F (5.6°C), and the slope of the ground within 2 percent, and the proper corrections applied. The correction to be applied for temperature when using a steel tape is

$$C_t = 0.0000065s(T - T_0) \tag{14.1}$$

The correction to be made to measurements on a slope is

$$C_h = s(1 - \cos\theta) \qquad \text{exact} \tag{14.2}$$

or

$$= 0.00015s\theta^2 \qquad \text{approximate} \tag{14.2a}$$

or

$$C_h = \frac{h^s}{2s} \qquad \text{approximate} \tag{14.2b}$$

where C_t = temperature correction to measured length, ft (× 0.3048 for meters)
C_h = correction to be subtracted from slope distance, ft (m)
s = measured length, ft (m)
T = temperature at which measurements are made, °F (°C)
T_0 = temperature at which tape is standardized, °F (°C)
h = difference in elevation at ends of measured length, ft (m)
θ = slope angle, deg

In more accurate taping, using a tape standardized when fully supported throughout, corrections should also be made for tension and for support conditions. The correction for tension is

$$C_p = \frac{(P_m - P_s)s}{SE} \tag{14.3}$$

The correction for sag when not fully supported is

$$C_s = \frac{w^2 L^3}{24P_m^2} \tag{14.4}$$

where C_p = tension correction to measured length, ft (m)
C_s = sag correction to measured length for each section of unsupported tape, ft (m)
P_m = actual tension, lb (N)
P_s = tension at which tape is standardized, lb (N) [usually 10 lb (44.5 N)]
S = cross-sectional area of tape, in² (mm²)
E = modulus of elasticity of tape, lb/in² (29 million lb/in² for steel)
w = weight of tape, lb/ft (N/m)
L = unsupported length, ft (m)

Electronic Distance Measurement (*EDM*).   Electronic measuring devices, utilizing principles similar to radar, are now in general use. Distances are measured by determination of the time required for a wave (infrared, radio, or laser beam) to travel at the speed of light to and from a point. For some types, readings of distance are obtained directly from a counter, eliminating the need for calculations. Advantages include ability to take readings across bodies of water, rugged terrain, and brush much more rapidly than with tape measurement. The precision of EDM ranges up to 1:300,000.

2. Measurement of Difference in Elevation.   Difference in elevation may be measured by three methods: barometric leveling, stadia leveling, and direct leveling. Barometric methods are used for rough or preliminary work. Stadia reading is a rapid method and will give results having an error, in feet (meters), of $1.0\sqrt{\text{distance in miles}}$ (Sec. 4). Direct leveling is the most accurate and most commonly used method for determining difference in elevation. EDM instruments (Sec. 1) make it possible to determine differences in elevation by measurement of slope distances and vertical angles.

Rough Leveling.   Rough leveling is practiced on preliminary or reconnaissance surveys. Sights are permitted up to 1000 ft (304.8 m) in length, and rod readings are made to 0.1 ft (0.004 m). Precision in feet is $0.4\sqrt{\text{distance in miles}}$.

Ordinary Leveling.   Ordinary leveling is used in the construction and location of highways, railroads, and the like. Sights are permitted up to 500 ft (152 m) in length, and rod readings are made to 0.01 ft (0.003 m). Precision in feet (meters) is $0.1\sqrt{\text{distance in miles}}$.

Accurate Leveling.   Accurate leveling is used for establishing important bench marks. Sights are limited to 300 ft (91.4 m) in length, and rod readings are made to 0.001 ft (0.0003 m). Precision in feet (meters) is $0.05\sqrt{\text{distance in miles}}$.

Precise Leveling.   Precise leveling is used for establishing bench marks at widely separated locations. Sights are limited to 300 ft (91.4 m) in length, and rod readings are made to 0.001 ft (0.0003 m). Special equipment and extreme care are used, and several runs are usually made. Precision in feet (meters) is $0.02\sqrt{\text{distance in miles}}$.

3. Measurement of Angles.   Angles may be measured with either a compass or a transit. The precision of compass measurement is from 30′ to 1°. The precision of transit measurements is from 1″ to 2′, depending on the type of instrument used and the care exercised. Angle measurements should have a precision consistent with distance measurements. Surveys in which distances are measured to 0.01 ft (0.003 m) should have angles measured to 15″, with a resulting accuracy of better than 1:10,000, and surveys with distances measured to 0.1 ft (0.03 m) and angles measured to 1′ should have an accuracy of about 1:5000.

4. Stadia Surveying.   In stadia surveying, a transit having horizontal stadia cross hairs above and below the central horizontal cross hair is used. The difference in the rod readings at the stadia cross hairs is termed the rod intercept. The intercept

may be converted to the horizontal and vertical distances between the instrument and the rod by the following formulas:

$$H = Ki(\cos a)^2 + (f + c) \cos a \tag{14.5}$$

$$V = \tfrac{1}{2}Ki(\sin 2a) + (f + c) \sin a \tag{14.6}$$

where H = horizontal distance between center of transit and rod, ft (m)
     V = vertical distance between center of transit and point on rod intersected by middle horizontal cross hair, ft (m)
     K = stadia factor (usually 100)
     i = rod intercept, ft (m)
     a = vertical inclination of line of sight, measured from the horizontal, deg
$f + c$ = instrument constant, ft (m) [usually taken as 1 ft (0.3 m)]

In the use of these formulas, distances are usually calculated to feet (meters) and differences in elevation to tenths of feet (meters).

5. Circular Curves. Circular curves are the most common type of horizontal curve used to connect intersecting tangent (or straight) sections of highways or railroads. In the United States, two methods of defining circular curves are in use: the first, in general use in railroad work, defines the degree of curve as the central angle subtended by a *chord* of 100 ft (30.4 m) in length; the second, used in highway work, defines the degree of curve as the central angle subtended by an *arc* of 100 ft (30.4 m) in length. In the metric system, the degree of curve is sometimes expressed as the number of degrees subtended by an arc or chord 20 m long.

The terms and symbols generally used in reference to circular curves are listed below and shown in Figs. 14.1 and 14.2.

PC = point of curvature, beginning of curve
 PI = point of intersection of tangents
PT = point of tangency, end of curve
  R = radius of curve, ft (m)
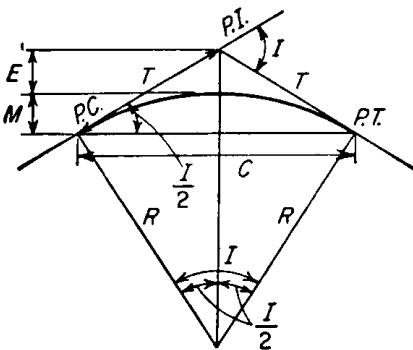
FIGURE 14.1 Circular curve.
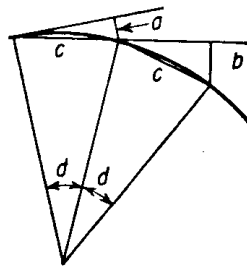
FIGURE 14.2 Offsets to circular curve.

D = degree of curve (see above)

I = deflection angle between tangents at PI, also central angle of curve

T = tangent distance, distance from PI to PC or PT, ft (m)

L = length of curve from PC to PT measured on 100-ft (30.4-m) chord for chord definition, on arc for arc definition, ft

C = length of long chord from PC to PT, ft (m)

E = external distance, distance from PI to midpoint of curve, ft (m)

M = midordinate, distance from midpoint of curve to midpoint of long chord, ft (m)

d = central angle for portion of curve ($d < D$)

l = length of curve (arc) determined by central angle d, ft (m)

c = length of curve (chord) determined by central angle d, ft (m)

a = tangent offset for chord of length c, ft (m)

b = chord offset for chord of length c, ft (m)

Equations of Circular Curves

$$R = 5729.578/D \qquad \text{exact for arc definition, approximate} \atop \text{for chord definition} \tag{14.7}$$

$$= 50/\sin \tfrac{1}{2}D \qquad \text{exact for chord definition} \tag{14.8}$$

$$T = R \tan \tfrac{1}{2}I \qquad \text{exact} \tag{14.9}$$

$$E = R \operatorname{exsec} \tfrac{1}{2}I = R(\sec \tfrac{1}{2}I - 1) \qquad \text{exact} \tag{14.10}$$

$$M = R \operatorname{vers} \tfrac{1}{2}I = R(1 - \cos \tfrac{1}{2}I) \qquad \text{exact} \tag{14.11}$$

$$C = 2R \sin \tfrac{1}{2}I \qquad \text{exact} \tag{14.12}$$

$$L = 100I/D \qquad \text{exact} \tag{14.13}$$

$$L - C = L^3/24R^2 = C^3/24R^2 \qquad \text{approximate} \tag{14.14}$$

$$d = Dl/100 \qquad \text{exact for arc definition} \tag{14.15}$$

$$= Dc/100 \qquad \text{approximate for chord definition} \tag{14.16}$$

$$\sin \tfrac{1}{2}d = c/2R \qquad \text{exact for chord definition} \tag{14.17}$$

$$a = c^2/2R \qquad \text{approximate} \tag{14.18}$$

$$b = c^2/R \qquad \text{approximate} \tag{14.19}$$

Layout of Circular Curve. The field layout of a circular curve depends on the geometric property of a circle that the angle between a tangent and a chord is one-half the included angle. The procedure is shown in Fig. 14.3, where the length of the first chord (or arc) is so chosen that point 1 is at an even 100-ft (30.4-m) station. Point 1 is located by measurement of the chord distance c from the PC and by the deflection angle $\tfrac{1}{2}d$ from the tangent. Point 2 is then located by measurement of
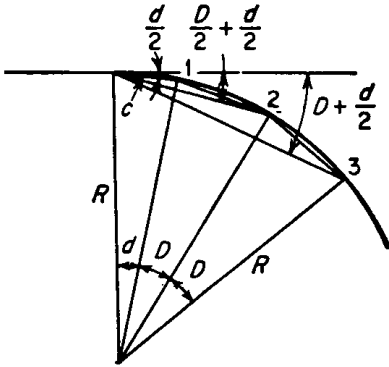
FIGURE 14.3   Layout of a circular curve.
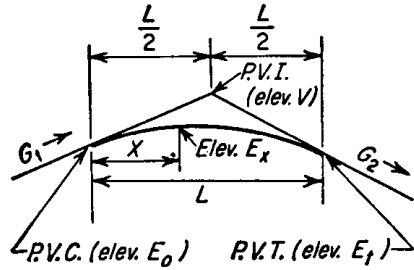
FIGURE 14.4   Vertical parabolic curve (summit curve).

the 100-ft (30.4-m) chord [or the chord corresponding to the 100-ft (30.4-m) arc] from point 1 and by the total deflection angle $(\frac{1}{2}D + \frac{1}{2}d)$ from the tangent. Succeeding points are similarly located. The entire curve can be laid out with the transit set at the PC.

6. Parabolic Curves.   Parabolic curves are used to connect sections of highways or railroads of differing gradient. The use of a parabolic curve provides a gradual change in direction along the curve. The terms and symbols generally used in reference to parabolic curves are listed below and shown in Fig. 14.4.

PVC = point of vertical curvature, beginning of curve

PVI = point of vertical intersection of grades on either side of curve

PVT = point of vertical tangency, end of curve

G_1 = grade at beginning of curve, ft/ft (m/m)

G_2 = grade at end of curve, ft/ft (m/m)

L = length of curve, ft

R = rate of change of grade, ft/ft² (m/m²)

V = elevation of PVI, ft (m)

E_0 = elevation of PVC, ft (m)

E_t = elevation of PVT, ft (m)

x = distance of any point on the curve from the PVC, ft (m)

E_x = elevation of point x distant from PVC, ft (m)

x_s = distance from PVC to lowest point on a sag curve or highest point on a summit curve, ft (m)

E_s = elevation of lowest point on a sag curve or highest point on a summit curve, ft (m)

Equations of Parabolic Curves.   In the parabolic-curve equations given below, algebraic quantities should always be used. Upward grades are positive and downward grades are negative.

$$R = (G_2 - G_1)/L \tag{14.20}$$

$$E_0 = V - \tfrac{1}{2}LG_1 \tag{14.21}$$

$$E_x = E_0 + G_1x + \tfrac{1}{2}Rx^2 \tag{14.22}$$

$$x_s = -G_1/R \tag{14.23}$$

Note. If x_s is negative or if $x_s > L$, the curve does not have a high point or a low point.

$$E_s = E_0 - G_1^2/2R \tag{14.24}$$

7. Photogrammetry. Photogrammetry is a method of obtaining measurements through use of ground or aerial photography. For large-scale projects, aerial photogrammetry techniques permit substantial savings in mapping time, but establishment of the location and elevation of control points by conventional ground survey are necessary. Through photoanalysis and interpretation, a trained interpreter can obtain reliable qualitative information concerning the type and characteristics of the soils, surface and ground waters, and manufactured features such as roads and bridges. Matched groups of overlapping photographs are viewed through a stereoscope and deductive and inductive methods used for evaluation. A selective field check is an important part of the photo-interpretative methodology.

The scale of aerial photography is given as a ratio, such as 1:6000, equivalent to 1 in = 500 ft. For a standard camera with a 6-in focal length lens and using 9-in square film, the following relationships apply:

Flight height above ground, ft (m)	Photographic scale	Coverage per single photograph, mi² (m²)	Map scale (5× enlargement of photographs), in = ft (mm = m)	Contour interval, ft (m)
1200	1:2400	0.11	1 = 40	1
(365.8)		(284,889)	(25.4 = 12.2)	(0.3)
3000	1:6000	0.72	1 = 100	2.5
(914.4)		(1,864,730)	(25.4 = 30.5)	(0.76)
4000	1:8000	1.29	1 = 133	(1)
(1219.2)		(3,340,975)	(25.4 = 40.5)	
6000	1:12,000	2.90	1 = 200	5
(1828.8)		(7,510,719)	(25.4 = 60.9)	(1.5)
8000	1:16,000	5.16	1 = 267	(2)
(2438.4)		(13,363,901)	(25.4 = 81.4)	
12,000	1:24,000	11.62	1 = 400	10
(3657.6)		(30,094,677)	(25.4 = 121.9)	(3.05)
18,000	1:36,000	26.15	1 = 600	15
(5486.4)		(67,725,973)	(25.4 = 182.9)	(4.6)
20,000	1:40,000	32.28	1 = 667	(5)
(6096.0)		(83,602,080)	(25.4 = 203.3)	

For mapping, the photographed area is usually covered in series of parallel strips, with photographs of the same strip overlapping 60 percent and photographs of adjacent strips overlapping 30 percent. Horizontal and vertical ground control is required for the preparation of maps.

SOIL MECHANICS AND FOUNDATIONS

8. Grain Size.   The grain size classification of soils used by the U.S. Department of Agriculture is given as follows:

Soil type	Particle diam, mm	Soil type	Particle diam, mm
Gravel................	>2.0	Sand, very fine	0.10–0.05
Gravel, fine	2.0–1.0	Silt	0.05–0.005
Sand, coarse	1.0–0.5	Clay	0.005–0.0002
Sand, medium	0.5–0.25	Colloids	<0.0002
Sand, fine.............	0.25–0.10		

9. Bureau of Public Roads Soil Classification.   The U.S. Bureau of Public Roads (now the Federal Highway Administration of the U.S. Department of Transportation) developed a detailed method for classifying soils for use as highway subgrades which was subsequently expanded and adopted by the American Association of State Highway and Transportation Officials (AASHTO). Soils are classified in seven major groups, A-1 through A-7 and further classified into subgroups (refer to AASHTO for detailed soil classification).

Group Index.   The group index is used as an approximate within-group evaluation of the materials of the A-2-6, A-2-7, A-4, A-5, A-6, and A-7 groups.

$$\text{Group index} = 0.2a + 0.005ac + 0.01bd$$

where a = that portion of the percentage passing the No. 200 sieve greater than 35 and not exceeding 75 percent, expressed as a positive whole number (1 to 40)
b = that portion of the percentage passing the No. 200 sieve greater than 15 and not exceeding 55 percent, expressed as a positive whole number (1 to 40)
c = that portion of the numerical liquid limit greater than 40 and not exceeding 60 percent, expressed as a positive whole number (1 to 20)
d = that portion of the numerical plasticity index greater than 10 and not exceeding 30 percent, expressed as a positive whole number (1 to 20)

Under average conditions of good drainage and thorough compaction, the supporting value of a material as a subgrade is in inverse ratio to its group index; that is, a group index of 0 indicates a good subgrade material and a group index of 20 indicates a very poor subgrade material.

10. Relationship among Soil Classifications. Other important soil classifications and measures of supporting strength include the following:

1. California Bearing Ratio, the ratio (expressed as a percentage) of the load required to cause a specified penetration in a given soil to the load required to cause the same penetration in a compacted gravel
2. Casagrande soil classification
3. Civil Aeronautics Administration soil classification
4. Resistance value R
5. Bearing value

In the Casagrande soil classification, the following symbols are used:

G = gravel, gravelly soil
S = sand, sandy soil
O = organic silt or clay
C = clay
M = silt or very fine sand
F = fine
P = poorly graded
W = well graded
L = low to medium compressibility
H = high compressibility

11. Relationship of Weights and Volumes in Soil. The unit weight of soil varies, depending on the amount of water contained in the soil. Three unit weights are in general use: the saturated unit weight γ_{sat}, the dry unit weight γ_{dry}, and the buoyant unit weight γ_b.

$$\gamma_{sat} = \frac{(G + e)\gamma_0}{1 + e} = \frac{(1 + w)G\gamma_0}{1 + e} \qquad S = 100\% \qquad (14.25)$$

$$\gamma_{dry} = \frac{G\gamma_0}{1 + e} \qquad S = 0\% \qquad (14.26)$$

$$\gamma_b = \frac{(G - 1)\gamma_0}{1 + e} \qquad S = 100\% \qquad (14.27)$$

Unit weights are generally expressed in pounds per cubic foot or grams per cubic centimeter. Representative values of unit weights for a soil with a specific gravity of 2.73 and a void ratio of 0.80 are

$$\gamma_{sat} = 122 \text{ lb/ft}^3 = 1.96 \text{ g/cm}^3$$

$$\gamma_{dry} = 95 \text{ lb/ft}^3 = 1.52 \text{ g/cm}^3$$

$$\gamma_b = 60 \text{ lb/ft}^3 = 0.96 \text{ g/cm}^3$$

The symbols used in Eqs. (14.25) to (14.27) and in Fig. 14.5 are

Gas

Water

Solids

$V = 1$ $V(\text{voids})$ $\frac{e}{1+e} =$ $V(\text{water})$ $= \frac{Se}{1+e}$ $V(\text{solids})$ $= \frac{1}{1+e}$

$W(\text{water})$ $= \frac{Se}{1+e}\gamma_0$ $W(\text{solids})$ $= \frac{1}{1+e}G\gamma_0$ $W(\text{total})(\text{for unit } V) = \gamma_t$ $= \frac{G+Se}{1+e}\gamma_0$ $= \frac{1+w}{1+e}G\gamma_0$

Total volume (solids + water + gas) = 1
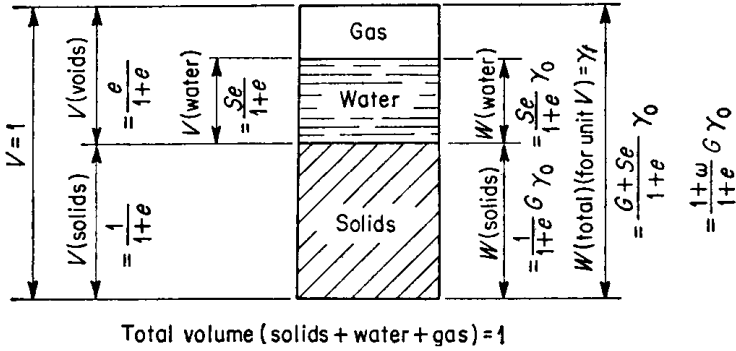
FIGURE 14.5   Relationship of weights and volumes in soil.

G = specific gravity of soil solids (specific gravity of quartz is 2.67; for majority of soils specific gravity ranges between 2.65 and 2.85; organic soils would have lower specific gravities)

γ_0 = unit weight of water (62.4 lb/ft^3 or 1.0 g/cm^3)

e = voids ratio, volume of voids in mass of soil divided by volume of solids in same mass [also equal to $n/(1 - n)$, where n is porosity—volume of voids in mass of soil divided by total volume of same mass]

S = degree of saturation, volume of water in mass of soil divided by volume of voids in same mass

w = water content, weight of water in mass of soil divided by weight of solids in same mass (also equal to Se/G)

12. Atterberg Limits. The Atterberg limits are used to define the change in the strength properties of fine-grained soils with a change in water content. The *liquid limit* w_l is the highest water content at which the soil has a small but definite shear resistance. At the liquid limit, the cohesion of the soil is practically zero. The *plastic limit* w_p is the lowest water content at which the soil is plastic. The *shrinkage limit* w_s is the lowest water content that can occur in a soil when it is completely saturated.

The *plasticity index* I_p is the liquid limit minus the plastic limit and is the range of water content throughout which the soil is plastic. When the plastic limit is equal to the liquid limit, the plasticity index is zero and the soil is entirely lacking in plasticity.

13. Permeability. The coefficient of permeability of a soil is the volume of water which would be forced through a mass of soil having a unit cross-sectional area and a unit length by a unit head of water. The permeability of sand usually ranges from 20×10^{-4} to 3000×10^{-4} cm/sec (5 to 850 ft/day). The permeability of clays is usually less than 10×10^{-4} cm/sec (2.8 ft/day).

Natural soils occurring in stratified formations have a permeability in the direction of stratification much greater than in the direction perpendicular to the stratification.

14. Internal Friction and Cohesion. The angle of *internal friction* for a soil is expressed by

$$\tan \phi = \frac{\tau}{\sigma} \tag{14.28}$$

where ϕ = angle of internal friction
$\tan \phi$ = coefficient of internal friction
σ = normal force on given plane in cohesionless soil mass
τ = shearing force on same plane when sliding on plane is impending

For medium and coarse sands, the angle of internal friction is about 30 to 35°. The angle of internal friction for clays ranges from practically 0 to 20°.

The *cohesion* of a soil is the shearing strength which the soil possesses by virtue of its intrinsic pressure. The value of the ultimate cohesive resistance of a soil is usually designated by *c*. Average values for *c* are given below:

General soil type	Cohesion c		General soil type	Cohesion c	
	lb/ft²	N/m²		lb/ft²	N/m²
Almost liquid clay	100	4,788	Medium clay	1000	47,880
Very soft clay	200	9,576	Damp, muddy sand	400	19,152
Soft clay	400	19,152			

15. Vertical Pressures in Soils. The vertical stress in a soil caused by a vertical, concentrated surface load may be determined with a fair degree of accuracy by the use of elastic theory. Two equations are in common use, the Boussinesq and the Westergaard. The Boussinesq equation applies to an elastic, isotropic, homogeneous mass which extends infinitely in all directions from a level surface. The vertical stress at a point in the mass is

$$\sigma_z = \frac{3P}{2\pi z^2 [1 + (r/z)^2]^{5/2}} \tag{14.29}$$

The Westergaard equation applies to an elastic material laterally reinforced with horizontal sheets of negligible thickness and infinite rigidity, which prevent the mass from undergoing lateral strain. The vertical stress at a point in the mass, assuming a Poisson's ratio of zero, is

$$\sigma_z = \frac{P}{\pi z^2 [1 + 2(r/z)^2]^{3/2}} \tag{14.30}$$

where σ_z = vertical stress at a point, lb/ft² (N/m²)
P = total concentrated surface load, lb (N)
z = depth of point at which σ_z acts, measured vertically downward from surface, ft (m)
r = horizontal distance from projection of surface load P to point at which σ_z acts, ft (m)

For values of r/z between 0 and 1, the Westergaard equation gives stresses appreciably lower than those given by the Boussinesq equation. For values of r/z greater than 2.2, both equations give stresses less than $P/100z^2$.

16. Lateral Pressures in Soils, Forces on Retaining Walls.   The Rankine theory of lateral earth pressures, used for estimating approximate values for lateral pressures on retaining walls, assumes that the pressure on the back of a vertical wall is the same as the pressure that would exist on a vertical plane in an infinite soil mass. Friction between the wall and the soil is neglected. The pressure on a wall consists of (1) the lateral pressure of the soil held by the wall, (2) the pressure of the water, if any, behind the wall, and (3) the lateral pressure from any surcharge on the soil behind the wall.
  Symbols used in this section are as follows:

γ = unit weight of soil, lb/ft^3 (kg/m^3) (saturated unit weight, dry unit weight, or buoyant unit weight, depending on conditions)

P = total thrust of soil, lb/linear ft of wall (kg/m)

H = total height of wall, ft (m)

ϕ = angle of internal friction of soil, deg

i = angle of inclination of ground surface behind wall with horizontal; also angle of inclination of line of action of total thrust P and pressures on wall with horizontal

K_A = coefficient of active pressure

K_P = coefficient of passive pressure

c = cohesion, lb/ft^2 (N/m^2)

Lateral Pressure of Cohesionless Soils.   For walls that retain cohesionless soils and are free to move an appreciable amount, the total thrust from the soil is

$$P = \frac{1}{2} \gamma H^2 \cos i \; \frac{\cos i - \sqrt{(\cos i)^2 - (\cos \phi)^2}}{\cos i + \sqrt{(\cos i)^2 - (\cos \phi)^2}} \qquad (14.31)$$

When the surface behind the wall is level, the thrust is

$$P = \tfrac{1}{2} \gamma H^2 K_A \qquad (14.32)$$

$$K_A = [\tan (45° - \phi/2]^2 \qquad (14.33)$$

The thrust is applied at a point $H/3$ above the bottom of the wall, and the pressure distribution is triangular, with the maximum pressure of $2P/H$ occurring at the bottom of the wall.
  For walls that retain cohesionless soils and are free to move only a slight amount, the total thrust is $1.12P$, where P is as given above. The thrust is applied at the midpoint of the wall and the pressure distribution is trapezoidal with the maximum pressure of $1.4P/H$ extending over the middle six-tenth of the height of the wall.
  For walls that retain cohesionless soils and are completely restrained (very rare), the total thrust from the soil is

$$P = \frac{1}{2}\gamma H^2 \cos i \; \frac{\cos i + \sqrt{(\cos i)^2 - (\cos \phi)^2}}{(\cos i)^2 - \sqrt{(\cos \phi)^2}} \qquad (14.34)$$

When the surface behind the wall is level, the thrust is

$$P = \frac{1}{2}\gamma H^2 K_P \qquad (14.35)$$

$$K_P = [\tan (45° + \phi/2)]^2 \qquad (14.36)$$

The thrust is applied at a point $H/3$ above the bottom of the wall, and the pressure distribution is triangular, with the maximum pressure of $2P/H$ occurring at the bottom of the wall.

Lateral Pressure of Cohesive Soils.   For walls that retain cohesive soils and are free to move a considerable amount over a long period of time, the total thrust from the soil (assuming a level surface) is

$$P = \frac{1}{2}\gamma H^2 K_A - 2cH\sqrt{K_A} \qquad (14.37)$$

or, since highly cohesive soils generally have small angles of internal friction,

$$P = \frac{1}{2}\gamma H^2 - 2cH \qquad (14.38)$$

The thrust is applied at a point somewhat below $H/3$ from the bottom of the wall, and the pressure distribution is approximately triangular.

   For walls that retain cohesive soils and are free to move only a small amount or not at all, the total thrust from the soil is

$$P = \frac{1}{2}\gamma H^2 K_P \qquad (14.39)$$

since the cohesion would be lost through plastic flow.

Water Pressure.   The total thrust from water retained behind a wall is

$$P = \frac{1}{2}\gamma_0 H^2 \qquad (14.40)$$

where H = height of water above bottom of wall, ft (m)
    γ_0 = unit weight of water, lb/ft³ (kg/m³) [62.4 lb/ft³ (999.02 kg/m³) for fresh water and 64 lb/ft³ (1024.6 kg/m³) for salt]

The thrust is applied at a point $H/3$ above the bottom of the wall, and the pressure distribution is triangular, with the maximum pressure of $2P/H$ occurring at the bottom of the wall. Regardless of the slope of the surface behind the wall, the thrust from water is always horizontal.

17. Stability of Slopes

Cohesionless Soil.   A slope in a cohesionless soil without seepage of water is stable if

$$i < \phi \qquad (14.41)$$

   With seepage of water parallel to the slope, and assuming the soil to be saturated, an infinite slope in a cohesionless soil is stable if

$$\tan i < \left(\frac{\gamma b}{\gamma_{\text{sat}}}\right) \tan \phi \tag{14.42}$$

where i = slope of ground surface
      ϕ = angle of internal friction of soil
γ_b, γ_{sat} = unit weights, lb/ft³ (kg/m³) (Sec. 11)

Cohesive Soils.   A slope in a cohesive soil is stable if

$$H < \frac{C}{\gamma N} \tag{14.43}$$

where H = height of slope, ft (m)
      C = cohesion, lb/ft² (kPa)
      γ = unit weight, lb/ft³ (kg/m³)
      N = stability number, dimensionless

For failure on the slope itself, without seepage water,

$$N = (\cos i)^2(\tan i - \tan \phi) \tag{14.44}$$

Similarly, with seepage of water,

$$N = (\cos i)^2 \left[\tan i - \left(\frac{\gamma_b}{\gamma_{\text{sat}}}\right) \tan \phi \right] \tag{14.44a}$$

where terms are as defined for Eq. (14.42).

18. Bearing Capacity of Soils.   The approximate ultimate bearing capacity under a long footing at the surface of a soil is given by Prandtl's equation as

$$q_u = \left(\frac{c}{\tan \phi} + \frac{1}{2}\, \gamma_{\text{dry}}\, b \sqrt{K_p}\right)(K_p e^{\pi \tan \phi} - 1) \tag{14.45}$$

where q_u = ultimate bearing capacity of soil, lb/ft² (kPa)
      c = cohesion, lb/ft² (kPa)
      ϕ = angle of internal friction, deg
      γ_{dry} = unit weight of dry soil, lb/ft³ (kg/m³) (Sec. 11)
      b = width of footing, ft (m)
      d = depth of footing below surface, ft (m)
      K_p = coefficient of passive pressure = $[\tan (45 + \phi/2)]^2$
      e = 2.718 . . .

For footings below the surface, the ultimate bearing capacity of the soil may be modified by the factor $1 + Cd/b$. The coefficient C is about 2 for cohesionless soils and about 0.3 for cohesive soils. The increase in bearing capacity with depth for cohesive soils is often neglected.

Typical values of the allowable bearing capacity of various soils are given in the National Building Code of the National Board of Fire Underwriters.

19. Settlement under Foundations.   The approximate relationship between loads on foundations and settlement is

$$\frac{q}{P} = C_1 \left(1 + \frac{2d}{b} \right) + \frac{C_2}{b} \tag{14.46}$$

where q = load intensity, lb/ft^2 (kPa)
  P = settlement, in (mm)
  d = depth of foundation below ground surface, ft (m)
  b = width of foundation, ft (m)
  C_1 = coefficient dependent on internal friction
  C_2 = coefficient dependent on cohesion

The coefficients C_1 and C_2 are usually determined by bearing-plate loading tests.

20. Allowable Loads on Piles.   A dynamic formula extensively used in the United States to determine the allowable static load on a pile is the *Engineering News* formula. For piles driven by a drop hammer, the allowable load is

$$P_a = \frac{2HW}{p + 1} \tag{14.47}$$

For piles driven by a single-acting hammer, the allowable load is

$$P_a = \frac{2WH}{p + 0.1} \tag{14.48}$$

For piles driven by a double-acting hammer, the allowable load is

$$P_a = \frac{2E}{p + 0.1} \tag{14.48a}$$

where P_a = allowable pile load, lb (kg)
  W = weight of hammer, lb (kg)
  H = height of drop or stroke, ft (m)
  E = actual energy delivered per blow, ft · lb (N · m)
  p = penetration of pile per blow, in (cm)

21. Types of Piles.   Foundation piles used to carry structure loads may be timber, concrete, composite (timber with concrete upper section), or steel. Piles which distribute the load throughout their length to the soil are called *friction piles.* Those which carry the load to firm substrata are *end-bearing piles.* Sheet piles used to retain soil or water may be wood planking, steel sheeting, or precast concrete sheets. Waterproofing is obtained by interlocking or overlapping of sections.

HIGHWAY AND TRAFFIC ENGINEERING

22. Highway Design Controls

Vehicle Characteristics.   Dimensions of the design vehicles recommended for use by the American Association of State Highway and Transportation Officials (AASHTO) as controls for geometric design are given in tabulations published by

the Association. Minimum turning paths for these vehicles are given in diagrams published by the Association. The vehicle which should be used in design is the largest one which represents a significant percentage of the traffic. For design of most highways accommodating truck traffic, one of the design semitrailer combinations should be used. A design check should be made for the largest vehicle expected, in order to ensure that such a vehicle can negotiate the designated turns, particularly if pavements are curbed.

Design Speed.   The design speed of a highway is the maximum safe speed that can be maintained over a specified section when conditions are favorable, so that the design features of the highway govern the speed.

Traffic.   The principal measures of traffic volume and character and the relationship between the various elements for rural highways are given in data published by the Association. Determination of the relationship between the traffic elements for urban highways usually requires special study.

Types of Arterial Highways.   A major street is an arterial highway with intersections at grade and direct access to abutting property and on which geometric design and traffic control measures are used to expedite the safe movement of through traffic. An expressway is a divided arterial highway with full or partial control of access and generally with grade separations at intersections. A freeway is an expressway with full control of access. A parkway is a type of arterial highway provided for noncommercial traffic, with full or partial control of access and usually located within a park or ribbon of parklike development.

23. Elements of Geometric Design

Stopping Sight Distance.   Design stopping sight distance is the minimum distance required for a vehicle traveling at or near the design speed to stop before reaching an object in its path. It is the sum of the distances traveled during perception and brake reaction time and the distance traveled while braking to a stop. Stopping sight distance is measured from a point 3.75 ft (1.14 m) above the road surface to a point 6 in (15.2 cm) above the road surface.

Passing Sight Distance.   Design passing sight distance is the minimum distance required to make safely a normal passing maneuver on two- and three-lane highways at passing speeds representative of nearly all drivers, commensurate with design speed. Passing sight distance is measured from a point 3.75 ft (1.14 m) above the road surface to a second point 4.5 ft (1.37 m) above the road surface. The minimum passing sight distance is given in data published by the Association.

24. Highway Capacity and Levels of Service.   The capacity of a highway is the maximum number of vehicles which has a reasonable expectation of passing over a given section of a lane or a roadway in one direction (or in both directions for a two-lane or a three-lane highway) during a given time period under prevailing roadway and traffic conditions. Capacity is equivalent to level of service *E*, defined below with other levels of service:

A = free flow, low volumes, high speeds, little or no restriction in maneuverability

B = stable flow, operating speeds somewhat restricted by traffic conditions (level B suitable for design of rural highways)

C = stable flow, operating speeds satisfactory but closely controlled by traffic conditions (level C suitable for urban design)

D = approaching unstable flow, tolerable operating speed, little freedom to maneuver

E = unstable flow, momentary stoppages (level E is capacity)

F = forced flow, congestion (volumes below capacity)

Service volumes for two-lane highways and for freeways and expressways under uninterrupted flow conditions are given in data published by the Association. These service volumes must be adjusted for roadway and traffic conditions. Adjustments for lane widths and lateral clearances are given by the Association. Adjustments for trucks may be made by converting trucks to equivalent passenger cars using the factors from the Association. The service volumes shown do not apply at intersections or in the vicinity of ramp termini.

25. Intersection Capacity and Levels of Service.   The capacity of a signalized intersection approach is the maximum number of vehicles that the approach can reasonably accommodate under the existing geometric, environmental, and traffic characteristics and controls. Capacity is equivalent to level of service E, defined below with other levels of service:

A = free operation, no vehicle waits longer than one red indication: load factor = 0.0

B = stable operation, occasional approach cycle fully utilized, many drivers somewhat restricted by traffic conditions: load factor = 0.1 or less (level B suitable for design of rural intersections)

C = stable operation, intermittent loading, most drivers somewhat restricted by traffic conditions: load factor = 0.3 or less (level C suitable for design of urban intersections)

D = approaching instability, substantial delays during short peaks within peak period: load factor = 0.7 or less

E = delays of several cycles, queues developing: load factor = 0.7 to 1.0, depending on conditions, with an average of 0.85 (level E is capacity)

F = jammed conditions, traffic flow controlled by downstream conditions, volumes unpredictable

Load factor is the proportion of green-signal intervals that are fully utilized.

Peak-hour factors of 1.00 are rarely found. Where long lines of waiting vehicles are typically present, a peak-hour factor of 0.90 or 0.95 may be used. The usual conditions in a metropolitan area are equivalent to a peak-hour factor of 0.85, but where a high rate of flow occurs over a period shorter than an hour, factors of 0.75, 0.70, or less should be used. Adjustments to service volumes for various peak hours are available from the Association.
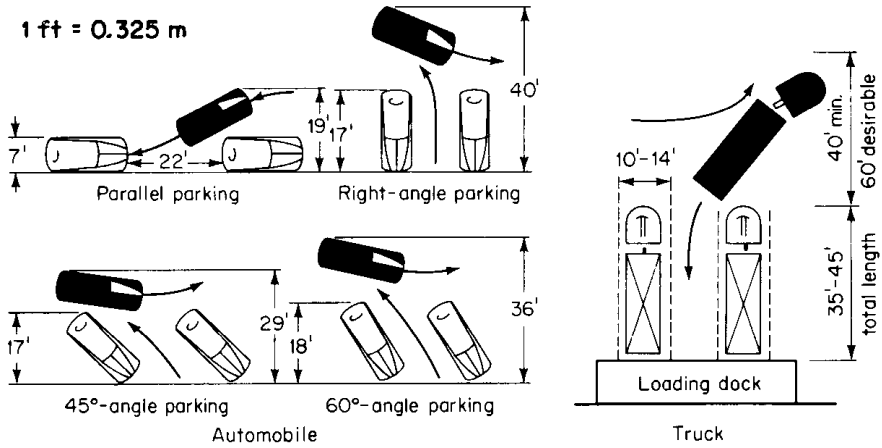
FIGURE 14.6 Storage and maneuvering space used for various parking positions.

TABLE 14.1 Street Space Used for Parking

Angle of parking at curb	Width of street used when parked		Width needed for parking plus maneuvering		Length of curb per car	
	ft	m	ft	m	ft	m
Parallel	7	2.1	19	5.8	22.0	6.7
45°	17	5.2	29	8.8	11.3	3.4
60°	18	5.5	36	10.9	9.2	2.8
90°	17	5.2	40	12.2	8.0	2.4

26. Parking Requirements

Automobile Parking.   The space used for automobile parking is shown in Fig. 14.6 and Table 14.1.

Truck Parking.   Truck loading/unloading platform heights range from 48 to 52 in (1.2 to 1.3 m) to match truck floors. Parking stall should be 10 to 14 ft (3.0 to 4.3 m) in width. Other dimensions are shown in Fig. 14.6.

RAILROADS

27. Track Gage.   Standard railroad gage in North America is 4 ft 8½ in (1435 mm) between inside of rails, measured ⅝ in (15.9 mm) below top of rail. The gage may be varied from 4 ft 8⅜ in (1432 mm) on some high-speed tangent track to 4 ft 9⅛ in (1451 mm) on curves of small radius.

28. Curvature and Superelevation. Maximum curvature for railroads is usually determined by train operating speeds and allowable superelevation. Maximum superelevation is 6 to 7 in (152 to 178 mm) but may be less on a particular railroad. The AREA (American Railway Engineering Association) formula for equilibrium superelevation is

$$e = 0.0007DV^2 \tag{14.49}$$

where D = degree of curve, deg
V = speed, mi/h (km/h)

For high-speed trains, as much as 3 in (76 mm) of unbalanced superelevation may be permitted, so that

$$e = 0.0007DV^2 - 3 \tag{14.49a}$$

Curvature may also be limited because of coupling difficulties on curves of more than 6°. Some railroads have established an absolute maximum curvature of 16° because of dimensions of the rigid wheelbase of cars and engines and permissible swing of couplers. Reverse curves should be separated by at least two car lengths.

Vertical curves should be of sufficient length to limit gradient changes to 0.05%/ 100 ft (30.5 m) for sags and 0.10%/100 ft (30.5 m) for crests on main lines; secondary lines may have vertical curves one-half the lengths required for main lines.

29. Clearances. Track clearances are established from center line of track for horizontal dimensions and from top of rail for vertical dimensions. Track spacings vary among railroads, but the following are common centerline spacings:

Main track–main track	13 ft 6 in (4.1 m)
Main track–yard or passing tracks	15 ft 0 in (4.6 m)
Yard track–yard track	13 ft 6 in (4.1 m)
Ladder track–yard track	18 ft 0 in (5.5 m)
Ladder track–ladder track	19 ft 0 in (5.8 m)

Track spacings are increased on curves at a rate of 2 in/deg (51 mm). Where adjacent tracks have different superelevation, additional clearance is provided of 3 in/in (76 mm/mm) of difference in superelevation.

Clearances from track to fixed structures vary by railroad and by state. Generally, vertical clearances of 22 ft (6.7 m) are required, except at building entrances, where 17 ft (5.2 m) may be permitted. Horizontal clearances of 8 ft (2.4 m) and 8 ft 6 in (2.6 m) are standard for tangent track, with additional allowances of 1 in/deg (25.4 mm/deg) of curve and 3 in/in (76 mm/mm) of superelevation for curved track.

30. Turnouts and Crossovers. Turnouts, used to divert trains from one track to another, consist of a switch, a frog to carry the wheel flanges over crossing rails, closure rails connecting the switch rails and the frog, and guard rails to guide the

flanges at the frog. Control points for turnouts are the actual or ½-in (3.2 mm) point of the switch [ground to a width of ⅛ in (3.2 mm)] and the actual or ½-in (12.7-mm) point of frog. Turnouts with No. 16 to No. 20 frogs are used for high-speed main-line movements, No. 10 to No. 12 for slow-speed main-line movements, and No. 8 for yards and sidings. Crossovers, used to transfer trains between parallel tracks, consist of two turnouts and connecting rails.

WATER SUPPLY, SEWERAGE, AND DRAINAGE

31. Water Supply and Treatment

Quantity of Water.   Average annual water requirements in metropolitan areas with metered systems generally range from 100 gal (378.5 L) per capita per day (gpcpd) to 200 gpcpd (757 Lpcpd), with a median value of about 150 gpcpd (567.8 Lpcpd). Unmetered supply systems have considerably higher consumption, and large water-using industries require special determination of demand.

Seasonal variations in water demand occur largely because of irrigation, lawn sprinkling, and air-conditioning loads, and maximum monthly consumption is generally about 125 percent of average annual demand but may range up to 200 percent of average annual demand. Maximum daily demands of 150 percent of average annual demand and maximum hourly demands of from 200 to 250 percent of annual average demand are commonly used for design.

Fire demand is often the determining factor in the design of mains, distribution storage tanks, and pumps, even though the total quantity of water required for fire fighting is small during a long period. For communities of less than 200,000 population, the fire demand is given by the National Board of Fire Underwriters as

$$Q = 1020\sqrt{P}(1 - 0.01\sqrt{P}) \tag{14.50}$$

where Q = fire demand, gal/min
P = population in thousands

The fire demand is added to the normal demand on the maximum day to determine the total maximum demand.

Design Period.   Pipes less than 12 in (304.8 mm) in diameter are generally designed to be adequate for the full development of the area served; pipes more than 12 in (304.8 mm) in diameter and wells, distribution systems, and filtration and treatment plants are generally designed for the flow expected 15 to 25 years in the future; and large dams and conduits are generally designed to be adequate for 25 to 50 years.

Quality of Water.   The outstanding requirement for a domestic water supply is freedom from pathogenic bacteria. In addition, there are reasonable limits for certain impurities, as listed below:

Impurity	Limit, ppm	Impurity	Limit, ppm
Turbidity	10	Iron plus manganese	0.3
Color	20	Magnesium	125
Lead	0.1	Total solids	500
Fluoride	1.5	Total hardness (calcium plus magnesium salts)	100
Copper	3.0		

Water Treatment.   Water treatment usually consists of filtration through either a slow or a rapid sand filter and disinfection with chlorine. In addition, water may be softened to remove hardness and aerated to remove iron and manganese.

The slow sand filter operates at a rate of 2 to 10 million gal/acre · day (1858 to 9291 L/m² · day) and is effective in removing tastes and odors from raw water. About 99 percent of the bacterial content is also removed.

The rapid sand filter operates at a rate of 125 to 250 million gal/acre · day (125,000 to 250,000 L/m² · day). However, preliminary treatment of the raw water is required, including chemical coagulation and sedimentation. The entire treatment process is effective in removing about 99.98 percent of the bacterial content, but removal of the color and turbidity is less dependable than for the slow sand filter and requires particular attention to the coagulation process.

Softening is accomplished by the addition of lime, or lime and soda ash, and sedimentation. The addition of lime and passage through a zeolite softener is also used. Iron and manganese may be removed by aeration and sedimentation.

Disinfection by the addition of chlorine is the final stage of any treatment process. Common practice is to add sufficient chlorine so that a small free chlorine residual is maintained.

32. Water Distribution.   Transmission mains connecting the source of supply to the distribution system must be large enough to supply at least the maximum daily demand plus fire flow. If the distribution system does not include storage, supply mains must also be adequate to deliver maximum hourly demands. Both transmission and distribution mains are usually designed by using the Hazen-Williams formula.

33. Sewage Collection and Treatment

Quantity of Sewage.   The average flow of sewage from a metropolitan area is about 100 gpcpd (gallons per capita per day) (378.5 Lpcpd). This rate may vary from 240 gpcpd (908 Lpcpd) in a maximum hour, 160 gpcpd (606 Lpcpd) on a maximum day, 70 gpcpd (265 Lpcpd) on a minimum day, and 40 gpcpd (151 Lpcpd) in a minimum hour. In addition, infiltration of ground water into sewers may be taken at about 600 gal/day · in diam. · mi.

Design Period.   Laterals and submains less than 15 in (381 mm) in diameter are generally designed to be adequate for the full development of the area served; main sewers, outfalls, and intercepter sewers are generally designed for the flow expected from 40 to 50 years in the future; and treatment works are generally designed for the flow expected from 10 to 25 years in the future.

Sewer Design.   Sewers should be at least 8 in (203 mm) in diameter and should be laid on a grade sufficient to produce a velocity of 2 ft/s (0.6 m/s) when flowing full, to prevent the deposition of suspended solids. Sewer lines should be designed with straight alignment and uniform grade between manholes, which should have a maximum spacing of 400 ft (122 m).

Quality of Sewage.   Sewage is approximately 99.92 percent water, with the remaining 0.08 percent (800 ppm by weight) composed of organic and mineral matter, as shown below:

	Organic matter, ppm	Mineral matter, ppm
Suspended solids	100	50
Colloidal solids	140	60
Dissolved solids	160	290

The "biochemical oxygen demand" of sewage, BOD, is the quantity of oxygen which must be supplied during the aerobic stabilization of sewage, and thus is a direct measure of the pollutional effect. Residential sewage has an average BOD of 0.24 lb (0.11 kg) oxygen per capita.

Sewage Treatment.   The degree of sewage treatment required should be based on the size, characteristics, and usage of the receiving body of water and upon the amount and quality of sewage to be treated. Complete sewage treatment might include preliminary treatment, such as screening to remove large suspended solids, grit removal, and grease removal; primary treatment, such as plain sedimentation or chemical precipitation; secondary treatment of a biological nature, such as the trickling filter or the activated sludge process; final treatment by chlorination; and finally disposal by dilution in a body of water. Approximate values of the BOD and suspended-solids removal of primary and secondary treatment are shown below:

Treatment	Percentage removal		Treatment process	Percentage removal	
	BOD	Suspended solids		BOD	Suspended solids
Plain sedimentation.	25–40	40–70	Trickling filter	80–95	80–90
Chemical precipitation . . .	50–75	70–90	Activated sludge	85–95	85–95

34. Sizes and Slopes of Sewers.   Sewer sizes and slopes are usually designed by using the Manning formula

$$v = \frac{1.486}{n} R^{2/3} S^{1/2} \tag{14.51}$$

where v = average velocity of flow, ft/s (m/s)
n = coefficient of roughness
R = hydraulic radius, ft (m) = A/P = $D/4$ for circular conduit flowing full
S = hydraulic gradient, ft head loss/ft length (m/m)
A = cross-sectional area of flow, ft² (m²)
P = wetted perimeter, ft (m)
D = diameter of circular conduit, ft (m)

For circular sewers flowing full, the Manning formula can be written as

$$Q = \frac{0.4632}{n} D^{2/3}S^{1/2} = \text{conveyance factor} \times S^{1/2} \qquad (14.52)$$

where Q = quantity of flow, ft³/s (m³/s).

35. Quantity of Runoff. The rational method for the determination of the quantity of storm water which appears as runoff involves the use of

$$Q = ciA \qquad (14.53)$$

where Q = runoff from rainfall, ft³/s (m³/s)
c = coefficient of runoff, dimensionless
i = rainfall intensity, expressed as a rate, in rain/h (mm/h)
A = tributary area, acres (m²)

These factors are discussed below.

Coefficient of Runoff. The coefficient of runoff for a particular area depends on the character of the surface, the type and extent of vegetation, the slope of the surface, and other less important factors. Approximate values of the coefficient of runoff c are given in Table 14.2.

Rainfall Intensity. The rainfall intensity is dependent on the recurrence interval and the time of concentration. The recurrence interval is the period of time within which, on the average, a rainfall of a given intensity will be equaled or exceeded

TABLE 14.2 Runoff Coefficients for Rational Formula

Type of area	Flat:slope <2%	Rolling:slope −10%	Hilly:slope >10%
Pavements, roofs, etc.	0.90	0.90	0.90
City business area	.80	.85	.85
Suburban residential areas	.45	.50	.55
Dense residential aeas	.60	.65	.70
Grassed areas	.25	.30	.30
Earth areas	.60	.65	.70
Cultivated land:			
Impermeable (clay, loam)	.50	.55	.60
Permeable (sand)	.25	.30	.35
Meadows and pasture lands	.25	.30	.35
Forests and wooded areas	.10	.15	.20

only once. Recurrence intervals of from 5 to 25 years are generally used, but for important structures periods of 100 years have been used.

For a particular area and a given recurrence interval, a study of rainfall records will permit the determination of an intensity-duration curve, which gives the rainfall intensity [in inches (mm) per hour as a function of the duration of rainfall. The rainfall intensity is greatest for short periods and decreases sharply as the duration of rainfall becomes greater. The intensity to use for a particular design is that for which the duration is equal to the time of concentration.

Time of Concentration.   The time of concentration for a particular inlet to a drainage system is the time required for rainfall falling on the most remote part of the tributary area drained by the inlet to reach the inlet. At this time, the entire area tributary to the inlet will be contributing to the runoff and the total runoff will be a maximum. The time for water to flow overland from the most remote part of the tributary area to the inlet may be approximated by

$$t = C \left(\frac{L}{Si^2} \right)^{1/2} \tag{14.54}$$

where t = time of overland flow, min
$\quad L$ = distance of overland flow, ft (m)
$\quad S$ = slope of land, ft/ft (m/m)
$\quad i$ = rainfall intensity, in/h (mm/h)
$\quad C$ = coefficient: 0.5 for paved areas, 1.0 for bare earth, 2.5 for turf

For any portions of the flow carried in ditches, the time of flow to the inlet may be computed by means of the Manning formula.

36. Flow in Drainage Channels.   Drainage channels are usually of such lengths that head losses other than those due to friction are negligible. Design of drainage channels is generally by the Manning formula:

$$v = \frac{1.486}{n} R^{2/3} S^{1/2} \tag{14.55}$$

where $S = h_f/l$ = hydraulic gradient, ft head loss/ft length (m/m)
$\quad n$ = coefficient of roughness, dimensionless
$\quad R$ = hydraulic radius, ft (m) = A/P
$\quad A$ = cross-sectional area of flow, ft² (m²)
$\quad P$ = wetted perimeter, ft (m)
$\quad h_f$ = head loss due to friction, ft (m)
$\quad l$ = length of channel or conduit, ft (m)
$\quad v$ = average velocity of flow, ft/s (m/s)

37. Steel Design

Working Stresses.   Structural steel has a weight of 490 lb/ft³ (7845 kg/m³), a modulus of elasticity of 29 million lb/in² (199.9 GPa), and a shearing modulus of 12 million lb/in² (82.7 GPa).

The allowable unit working stresses for structural steel are given in structural steel manuals published as codes or other guides. Somewhat lower stresses are used

for bridges, in recognition of the more severe service and the greater possibility of overloading such structures. Examples of recommended formulas for steel design are as follows.

Tension.   Tension on net section, except at pinholes:

$$F_t = 0.60F_y \tag{14.56}$$

Tension on net section at pinholes:

$$F_t = 0.45F_y \tag{14.57}$$

where F_t = allowable tensile stress, lb/in² (kPa)

The slenderness ratio Kl/r [defined following Eq. (14.63)] preferably should not exceed 240 for main members or 300 for bracing or other secondary members, other than rods.

Shear.   Shear on gross section:

$$F_v = 0.40F_y \tag{14.58}$$

where F_v = allowable shear stress, lb/in (kPa)

Compression.   Compression on the gross section of axially loaded compression members when Kl/r is less than C_c:

$$F_a = \frac{\{1 - [(Kl/r)^2/2C_c^2]\}\ F_y}{SF} \tag{14.59}$$

Compression on the gross section of axially loaded columns when Kl/r exceeds C_c:

$$F_a = \frac{149{,}000{,}000}{(Kl/r)^2} \tag{14.60}$$

Compression on the gross section of axially loaded bracing and secondary members when l/r exceeds 120:

$$F_{as} = \frac{F_a\ [\text{from Eq. (14.59) or (14.60), depending on } C_c]}{1.6 - l/200r} \tag{14.61}$$

Compression on the gross area of plate girder stiffeners:

$$F_a = 0.60F_y \tag{14.62}$$

Compression on the web of rolled shapes at the toe of the fillet:

$$F_a = 0.75F_y \tag{14.63}$$

where F_a = allowable comprehensive stress permitted in absence of bending moment, lb/in² (kPa)

F_{as} = allowable comprehensive stress permitted in absence of bending moment for bracing and other secondary members, lb/in² (kPa)

F_y = minimum yield point, lb/in² (kPa)

K = effective-length factor (suggested design values shown in Fig. 14.7)

l = actual unbraced length, in (mm)

Buckled shape of column is shown by dashed line	(a)	(b)	(c)	(d)	(e)	(f)
Theoretical K value	0.5	0.7	1.0	1.0	2.0	2.0
Recommended design value when ideal conditions are approximated	0.65	0.80	1.2	1.0	2.10	2.0
End condition code		Rotation fixed and translation fixed				
		Rotation free and translation fixed				
		Rotation fixed and translation free				
		Rotation free and translation free				

FIGURE 14.7 Effective-length factors for members subject to axial load.

r = radius of gyration corresponding to K and l, in (mm) ($= \sqrt{I/A}$)
I = moment of inertia, in^4 (mm^4)
A = gross cross-sectional area, in^2 (mm^2)
C_c = slenderness ratio separating elastic and inelastic buckling

$$C_c = \sqrt{\frac{2\pi^2 E}{F_y}} \tag{14.64}$$

SF = factor of safety

$$SF = 1.67 + \frac{3(Kl/r)}{8C_c} - \frac{(Kl/r)^3}{8C_c^3} \tag{14.65}$$

The slenderness ratio Kl/r of compression members must not exceed 200.

Bending. Tension and compression on extreme fibers of laterally supported compact shapes* having an axis of symmetry in the plane of loading:

$$F_b = 0.66 F_y \tag{14.66}$$

where F_b = allowable bending stress in absence of axial load, lb/in^2 (kPa)

---

*A compact shape has the flanges continuously connected to the web or webs; the width of un-stiffened projecting elements of the compression flange does not exceed $2050/\sqrt{F_y}$ times the flange thickness; the width of flange plates does not exceed $6000\sqrt{F_y}$ times the flange-plate thickness; and the depth of the web does not exceed $20{,}200[1 - 3.74(f_a/F_a)]/\sqrt{F_y}$ times the web thickness where (f_a/F_a) is the ratio of computed axial stress to allowable axial stress in the absence of bending moment, except that it need not be less than $8100/\sqrt{F_y}$.

Laterally supported members have transverse movement of the compression flange prevented at points of support not more than $2400b_f/\sqrt{F_y}$, or $20,000,000A_f/dF_y$ in (mm) apart, where

$$b_f = \text{compression flange width, in (mm)}$$

$$A_f = \text{cross-sectional area of compression flange, in}^2 \text{ (mm}^2\text{)}$$

$$d = \text{depth of member, in (mm)}$$

Tension and compression on extreme fibers of laterally supported unsymmetrical members (except channels) or box-type members, and tension on other rolled shapes or built-up members:

$$F_b = 0.60F_y \tag{14.67}$$

Compression on extreme fibers of other rolled shapes and built-up members (except box-type members), the larger value from Eqs. (14.68) and (14.69), but not more than $0.60F_y$:

$$F_b = \left[1.0 - \frac{(l/r)^2}{2C_c^2}\right]0.60F_y \tag{14.68}$$

$$F_b = \frac{12,000,000}{ld/A_f} \tag{14.69}$$

where l = unsupported length of compression flange, in (mm)
   r = radius of gyration of compression flange plus one-sixth web about an axis in plane of web, in (mm)
   d = depth of member, in (mm)
   A_f = cross-sectional area of compression flange, in^2 (mm^2)
   $C_c = \sqrt{2\pi^2E/F_y}$

Equation (14.68) may be further modified in certain cases by consideration of the moments at each end of the unsupported length.

Compression on extreme fibers of channels, the value from Eq. (14.69) but not more than $0.60F_y$.

Tension and compression on extreme fibers of pins:

$$F_b = 0.90F_y \tag{14.70}$$

Tension and compression on extreme fibers of rectangular bearing plates:

$$F_b = 0.75F_y \tag{14.71}$$

Bearing. Bearing on milled surfaces and pins in reamed, drilled, or bored holes:

$$F_p = 0.90F_y \tag{14.72}$$

Bearing on bolts or rivets:

$$F_p = 1.35F_y \tag{14.73}$$

where F_p = allowable bearing stress, lb/in^2 (kPa)

Columns and Tension Members.   Columns are designed on the basis of the gross area of the section used. Tension members, however, are designed on the basis of net area, with deductions made for rivet and other holes. In determining net area, net width is obtained by deducting from the gross width the sum of the diameters of all the holes in any chain of holes in any diagonal or zigzag direction and adding for each gauge space in the chain the quantity $s^2/4g$, where s is the longitudinal spacing (pitch) in inches (millimeters) of any two successive holes, and g is the transverse spacing (gage) in inches (millimeters) of the same two holes. Several chains of holes should be tried until the one giving the least net width is found. The net area of a tension member taken through a hole is limited to 85 percent of the gross area. The diameter of a rivet or bolt hole is taken as $\frac{1}{8}$ in (3.2 mm) greater than the nominal diameter of the rivet or bolt.

Beams.   The extreme fiber stress in bending for a steel beam is computed as

$$f_b = \frac{M}{S}$$ (14.74)

where f_b = maximum fiber stress, lb/in² (kPa)
        M = bending moment, lb · in (N · m)
        S = section modulus (= I/c), in³ (mm³)
        I = moment of inertia of cross-sectional area, in⁴ (mm⁴)
        c = distance from extreme fiber to neutral axis, in (mm) (c = one-half the depth for a symmetrical cross section)

The section moduli for selected standard rolled-steel sections are given in the *Manual of Steel Construction,* American Institute of Steel Construction. For beams requiring greater section moduli, built-up sections or plate girders are generally used.
    The shearing stress in a steel beam with flanges is relatively constant over the depth of the web, and may be computed as

$$f_v = \frac{V}{d_w t}$$ (14.75)

where f_v = shearing stress in web, lb/in² (kPa)
        V = total shear at section, lb (N)
        d_w = depth of web, in (mm)
        t = thickness of web, in (mm)

Built-up sections will usually require stiffeners to prevent web buckling.

38. Steel Connections

Rivets.   Rivets vary in size from $\frac{3}{8}$ to $1\frac{1}{4}$ in (9.5 to 31.8 mm) in diameter, with the $\frac{3}{4}$- and $\frac{7}{8}$-in (1.91- and 22.2-mm) sizes most commonly used. The standard sizes and cross-sectional areas of rivets are given in steel manuals. Rivet holes are considered to be $\frac{1}{8}$ in (3.2 mm) larger in diameter than the rivet.

Bolts.   Both turned and unfinished bolts are available in the same sizes as rivets. Larger sizes are also available. Cross-sectional areas are given in steel manuals.

Welds. The allowable loads on butt welds of the same size as the connected members are the same as for the members. The allowable load per inch (millimeters) of fillet weld is determined on the minimum cross section; for an equal leg weld, the minimum section at the throat is 0.707 times the dimension of the weld leg.

Working Stresses

Rivets. Allowable stresses for A502, Grade 1 hot-driven rivets are 20,000 lb/in² (137.9 MPa) in tension and 15,000 lb/in² (103.4 MPa) in shear; for A502, Grade 2 hot-driven rivets, stresses are 27,000 lb/in² (186.1 MPa) in tension and 20,000 lb/in² (137.9 MPa) in shear.

Bolts. Allowable stresses for A307 bolts are 20,000 lb/in² (137.9 MPa) in tension and 10,000 lb/in² (68.9 MPa) in shear; for other threaded parts of other steels, stresses are $0.60F_y$ in tension and $0.30F_y$ in shear.

Welds. The allowable stress for welds on A36, A242, and A441 steels is 21,000 lb/in² (144.8 MPa); except that complete-penetration groove welds with any type of loading and partial-penetration groove welds loaded in compression, bearing, or tension parallel to the axis of the weld may be stressed to the full allowable stress of the connected material.

39. Reinforced-Concrete Design

Concrete Mixes. The proportioning of concrete ingredients is by weight or volume. Weight measures are considered more reliable. Concrete mixes are designated by the proportion of each ingredient in the order: cement, sand, coarse aggregate. For example, a 1:2:3 mix is one part cement, two parts sand, and three parts stone or gravel. A bag of cement [94 lb (42.6 kg)] is equivalent to 1 ft³ (0.28 m³). Water for concrete should be free of injurious amounts of oils, acids, alkalis, salts, or organic matter.

Strength and Durability of Concrete. The most important factor affecting the strength and durability of concrete is the water-cement ratio. For concrete made from average materials, compressive strengths to be used for design are shown in *Building Code Requirements for Reinforced Concrete,* American Concrete Institute.

Design Methods. Reinforced concrete may be designed by either one of two methods: working-stress design or ultimate-strength design. Both methods are permitted under current codes, and the selection of a method is left to the designer. Refer to American Concrete Institute (ACI) Building Codes.

Reinforcing Bars. Steel bars for concrete reinforcement are available in various sizes.

Working-Stress Design

Design Loadings. In working-stress design, members should be designed to withstand actual service loads, consisting of dead loads, live loads, wind loads, and earthquake loads in any combination. Members subject to stress produced by wind or earthquake may be proportioned for stresses one-third greater than those given in applicable codes.

Working Stresses.   Allowable unit stresses for concrete and reinforcing steel are given in applicable codes.

Beams.   Concrete beams may be considered to be of three principal types: rectangular beams with tensile reinforcing only, T beams with tensile reinforcing only, and beams with tensile and compressive reinforcing.

Rectangular Beams with Tensile Reinforcing Only.   This type of beam includes slabs [for which b = 12 in (304.8 mm) when the moment and shear are expressed per foot (meter) of width]. The stresses in the concrete and steel are

$$f_c = \frac{2M}{kjbd^2} \tag{14.76}$$

$$f_s = \frac{M}{A_s jd} = \frac{M}{pjbd^2} \tag{14.77}$$

where b = width of beam [equals 12 in (304.8 mm) for slab], in (mm)
$\quad\ \ d$ = effective depth of beam, measured from compressive face of beam to centroid of tensile reinforcing (Fig. 14.8), in (mm)
$\quad\ \ M$ = bending moment, lb · in (N · m)
$\quad\ \ f_c$ = compressive stress in extreme fiber of concrete, lb/in² (kPa)
$\quad\ \ f_s$ = stress in reinforcement, lb/in² (kPa)
$\quad\ \ A_s$ = cross-sectional area of tensile reinforcing, in² (mm²)
$\quad\ \ j$ = ratio of distance between centroid of compression and centroid of tension to depth d
$\quad\ \ k$ = ratio of depth of compression area to depth d
$\quad\ \ p$ = ratio of cross-sectional area of tensile reinforcing to area of the beam ($= A_s/bd$)

For approximate design purposes, j may be assumed to be ⅞ and k ⅜. For average structures, the following guides to the depth d of a reinforced-concrete beam may be used:
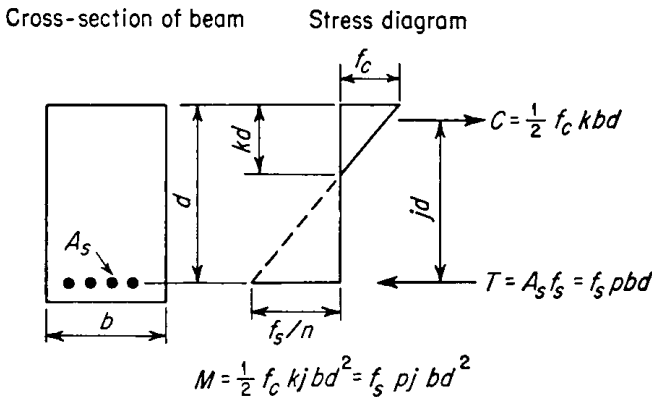
FIGURE 14.8   Rectangular concrete beam with tensile reinforcing only.

Member	d
Roof and floor slabs	$l/25$
Light beams	$l/15$
Heavy beams and girders	$l/12$–$l/10$

where l is the span of the beam or slab in inches (millimeters). The width of a beam should be at least $l/32$.

For a balanced design, one in which both the concrete and the steel are stressed to the maximum allowable stress, the following formulas may be used:

$$bd^2 = \frac{M}{K} \qquad (14.78)$$

$$K = \tfrac{1}{2}f_c kj = pf_s j \qquad (14.79)$$

Values of K, k, j, and p for commonly used stresses are given in building codes.

The reinforcing requirements may be approximately determined from

$$A_s = \frac{8M}{7f_s d} \qquad (14.80)$$

$$A_{sc} = \frac{M - M'}{nf_c d} \qquad (14.81)$$

where A_s = total cross-sectional area of tensile reinforcing, in² (mm²)
A_{sc} = cross-sectional area of compressive reinforcing, in² (mm²)
M = total bending moment, lb · in (N · m)
M' = bending moment which would be carried by beam of balanced design and same dimensions with tensile reinforcing only, lb · in (N · m)
n = ratio of modulus of elasticity of steel to that of concrete

Check of Stresses in Beam. Beams designed by the above approximate formulas should be checked to ensure that the actual stresses do not exceed the allowable, and that the reinforcing is not excessive. This can be accomplished by determining the moment of inertia of the beam. In this determination, the concrete below the neutral axis should not be considered as stressed, while the reinforcing steel should be transformed into an equivalent concrete section. For tensile reinforcing, this transformation is made by multiplying the area A_s by n, the ratio of the modulus of elasticity of steel to that of concrete. For compressive reinforcing, the area A_{sc} is multiplied by $(2n - 1)$. This factor includes allowances for the concrete in compression replaced by the compressive reinforcing and for the plastic flow of concrete. The neutral axis is then located by solving

$$\tfrac{1}{2}bc_c^2 + (2n - 1)A_{sc}c_{sc} = nA_s c_s \qquad (14.82)$$

for the unknowns c_c, c_{sc}, and c_s (Fig. 14.9). The moment of inertia of the transformed beam section is
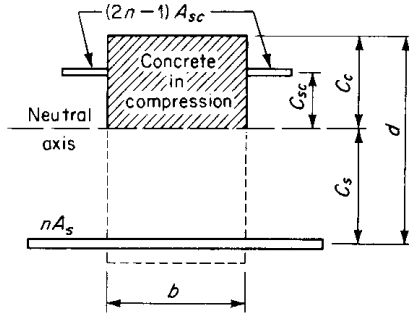
FIGURE 14.9 Transformed section of concrete beam.

$$I = \tfrac{1}{3}bc_c^3 + (2n - 1)A_{sc}c_{sc}^2 + nA_sc^2 \qquad (14.83)$$

and the stresses are

$$f_c = \frac{Mc_c}{I} \qquad (14.84)$$

$$f_{sc} = \frac{2nMc_{sc}}{I} \qquad (14.85)$$

$$f_s = \frac{nMc_s}{I} \qquad (14.86)$$

where f_c, f_{sc}, f_s = actual unit stresses in extreme fiber of concrete, in compressive reinforcing steel, and in tensile reinforcing steel, respectively, lb/in² (kPa)

  c_c, c_{sc}, c_s = distances from neutral axis to face of concrete, to compressive reinforcing steel, and to tensile reinforcing steel, respectively, in (mm)

  I = moment of inertia of transformed beam section, in⁴ (mm⁴)

  b = beam width, in (mm)

and A_s, A_{sc}, M, and n are as defined for Eqs. (14.80) and (14.81).

Shear and Diagonal Tension in Beams. The shearing unit stress, as a measure of diagonal tension, in a reinforced-concrete beam is

$$v = \frac{V}{bd} \qquad (14.87)$$

where v = shearing unit stress, lb/in² (kPa)

  V = total shear, lb (N)

  b = width of beam (for T beam use width of stem), in (mm)

  d = effective depth of beam

If the value of the shearing unit stress as computed above exceeds the allowable shearing unit stress (ν_c in building codes) web reinforcement should be provided. Such reinforcement will usually consist of stirrups. The cross-sectional area required for a stirrup placed perpendicular to the longitudinal reinforcement is

$$A_v = \frac{(V - V')s}{f_v d} \tag{14.88}$$

where A_v = cross-sectional area of web reinforcement in distance s (measured parallel to longitudinal reinforcement), in² (mm²)
  f_v = allowable unit stress in web reinforcement, lb/in² (kPa)
  V = total shear, lb (N)
  V' = shear which concrete alone could carry ($= \nu_c bd$), lb (N)
  s = spacing of stirrups in direction parallel to that of longitudinal reinforcing, in (mm)
  d = effective depth, in (mm)

Stirrups should be so spaced that every 45° line extending from the middepth of the beam to the longitudinal tension bars is crossed by at least one stirrup. If the total shearing unit stress is in excess of $3\sqrt{f_c'}$ lb/in² (kPa), every such line should be crossed by at least two stirrups. The shear stress at any section should not exceed $5\sqrt{f_c'}$ lb/in² (kPa).

Bond and Anchorage for Reinforcing Bars.   In beams in which the tensile reinforcing is parallel to the compression face, the bond stress on the bars is

$$u = \frac{V}{jd\,\Sigma_0} \tag{14.89}$$

where u = bond stress on surface of bar, lb/in² (kPa)
  V = total shear, lb (N)
  d = effective depth of beam, in (mm)
  Σ_0 = sum of perimeters of tensile reinforcing bars, in (mm)

For preliminary design, the ratio j may be assumed to be ⅞. Bond stresses may not exceed the values shown in building codes. To provide sufficient anchorage to develop the strength of reinforcing steel, tensile bars should be extended beyond the point at which they are needed to resist stress and should be terminated in a compression region, over a support, or with a hook.

Columns.   The principal columns in a structure should have a minimum diameter of 10 in (254 mm) or, for rectangular columns, a minimum thickness of 8 in (203.2 mm) and a minimum gross cross-sectional area of 96 in² (619.4 cm²).
  Short Columns, Spiral Reinforcing.   For short columns with closely spaced spiral reinforcing enclosing a circular concrete core reinforced with vertical bars, the maximum allowable load is

$$P = A_g(0.25 f_c' + f_s p_g) \tag{14.90}$$

where P = total allowable axial load, lb (N)
  A_g = gross cross-sectional area of column, in² (mm²)
  f_c' = compressive strength of concrete, lb/in² (kPa)

f_s = allowable stress in vertical concrete reinforcing, lb/in² (kPa), equal to 40 percent of the minimum yield strength, but not to exceed 30,000 lb/in² (206.8 MPa)

p_g = ratio of cross-sectional area of vertical reinforcing steel to gross area of column A_g

The ratio p_g should not be less than 0.01 nor more than 0.08. The minimum number of bars to be used is six, and the minimum size is No. 5. The spiral reinforcing to be used in a spirally reinforced column is

$$p_s = \frac{0.45(A_g/A_c - 1) f_c'}{f_y} \qquad (14.91)$$

where p_s = ratio of spiral volume to concrete-core volume (out-to-out spiral)
A_c = cross-sectional area of column core (out-to-out spiral), in² (mm²)
f_y = yield strength of spiral reinforcement, lb/in² (kPa), but not to exceed 60,000 lb/in² (413.6 MPa)

Short Columns with Ties. The maximum allowable load on short columns reinforced with longitudinal bars and separate lateral ties is 85 percent of that given in Eq. (14.90) for spirally reinforced columns. The ratio p_g for a tied column should not be less than 0.01 nor more than 0.08. The longitudinal reinforcing should consist of at least four bars, and the minimum size is No. 5.

Ties should be at least ¼ in (6.4 mm) in diameter, and should be spaced apart not over 16 bar diameters, 48 tie diameters, or the least dimension of the column.

Long Columns. Allowable column loads where compression governs design must be adjusted for column length, as follows:

1. If the ends of the column are fixed so that a point of contraflexure occurs between the ends, applied axial loads and moments should be divided by R from Eq. (14.92) (R cannot exceed 1.0):

$$R = \frac{1.32 - 0.006h}{r} \qquad (14.92)$$

2. If relative lateral displacement of the ends of the column is prevented and the member is bent in single curvature, applied axial loads and moments should be divided by R from Eq. (14.93) (R cannot exceed 1.0):

$$R = \frac{1.07 - 0.008h}{r} \qquad (14.93)$$

where h = unsupported length of column, in (mm)
r = radius of gyration of gross concrete area, in (mm)
= 0.30 times depth for rectangular column
= 0.25 times diameter for circular column
R = long-column load reduction factor

Combined Bending and Compression. The strength of a symmetrical column is controlled by compression if the equivalent axial load N has an eccentricity e in each principal direction no greater than given by Eq. (14.94) or (14.95) and by tension if e exceeds these values in either principal direction.

For spiral columns:

$$e_b = 0.43 p_g m D_s + 0.14t \tag{14.94}$$

For tied columns:

$$e_b = (0.67 p_g m + 0.17)d \tag{14.95}$$

where e = eccentricity, in (mm)
e_b = maximum permissible eccentricity, in (mm)
N = eccentric load normal to cross section of column
p_g = ratio of area of vertical reinforcement to gross concrete area
m = $f_y/0.85f_c'$
D_s = diameter of circle through centers of longitudinal reinforcement, in (mm)
t = diameter of column or overall depth of column, in (mm)
d = distance from extreme compression fiber to centroid of tension reinforcement, in (mm)
f_y = yield point of reinforcement, lb/in² (kPa)

Design of columns controlled by compression is based on Eq. (14.96), except that the allowable load N may not exceed the allowable load P [Eq. (14.90)] permitted when the column supports axial load only.

$$\frac{f_a}{F_a} + \frac{f_{bx}}{F_b} + \frac{f_{by}}{F_b} \leq 1.0 \tag{14.96}$$

where f_a = axial load divided by gross concrete area, lb/in² (kPa)
f_{bx}, f_{by} = bending moment about x and y axes, divided by section modulus of corresponding transformed uncracked section, lb/in² (kPa)
F_b = allowable bending stress permitted for bending alone, lb/in² (kPa)
F_a = $0.34(1 + p_g m)f_c'$

The allowable bending moment on columns controlled by tension varies linearly with the axial load from M_0 when the section is in pure bending to M_b when the axial load is N_b.

For spiral columns:

$$M_0 = 0.12 A_{st} f_y D_s \tag{14.97}$$

For tied columns:

$$M_0 = 0.40 A_s f_y (d - d') \tag{14.98}$$

where A_{st} = total area of longitudinal reinforcement, in² (mm²)
f_y = yield strength of reinforcement, lb/in² (kPa)
D_s = diameter of circle through centers of longitudinal reinforcement, in (mm)
A_s = area of tension reinforcement, in² (mm²)
d = distance from extreme compression fiber to centroid of tension reinforcement, in (mm)
d' = distance from extreme compression fiber to centroid of compression reinforcement, in (mm)

N_b and M_b are the axial load and moment at the balanced condition, i.e., when the eccentricity e equals e_b as determined from Eq. (14.94) or (14.95). At this condition, N_b and M_b should be determined from Eq. (14.96) so that

$$M_b = N_b e_b \tag{14.99}$$

When bending is about two axes,

$$\frac{M_x}{M_{0x}} + \frac{M_y}{M_{0y}} \le 1 \tag{14.100}$$

where M_x and M_y are bending moments about the x and y axes, and M_{0x} and M_{0y} are the values of M_0 for bending about these axes.

Uftimate-Strength Design

Loadings. In ultimate-strength design, proportioning of members is based upon design loads determined from appropriate combinations of dead loads, live loads, wind loads, and earthquake loads. The design loads are determined by multiplying the actual loads by various safety factors. Design loads are

$$U = 1.4D + 1.7L \tag{14.101}$$

or, with wind load a factor,

$$U = 0.75(1.4D + 1.7L + 1.7W) \tag{14.102}$$

or, with L absent,

$$U = 0.9D + 1.3W \tag{14.103}$$

where U = design load, lb (N) (use maximum value from above equations)
    D = dead loads, lb (N)
    L = live loads plus impact, lb (N)
    W = wind loads, lb (N)

For earthquake loading, substitute $1.1E$ for W in Eqs. (14.102) and (14.103), where E = earthquake loads, lb (N).

Assumptions in Design.   (See Fig. 14.10.) In ultimate-strength design, basic assumptions include (1) maximum strain occurs at extreme compression fiber and is 0.003 in/in (mm/mm), (2) stress in reinforcing bars below the yield strength f_y, is 29,000,000 lb/in^2 (199.9 GPa) times the steel strain, (3) strain in concrete directly proportional to the distance from neutral axis, (4) tensile strength of concrete is neglected in flexural calculations, and (5) concrete stress intensity of $0.85f'_c$ is uniformly distributed over a depth $a = kc$, where c is the distance from the extreme compression fiber to the neutral axis. For $f'_c \le 4000$ lb/in^2 (27.6 MPa), use $k = 0.85$; for each 1000 lb/in^2 (6.9 MPa) in excess of 4000 lb/in^2 (27.6 MPa), k is reduced by 0.05.

Design Values.   For conventional structures, values normally used are compressive strength of concrete $f'_c = 4000$ lb/in^2 (27.6 MPa) and yield strength of reinforcement $f_y = 60,000$ lb/in^2 (413.6 MPa).
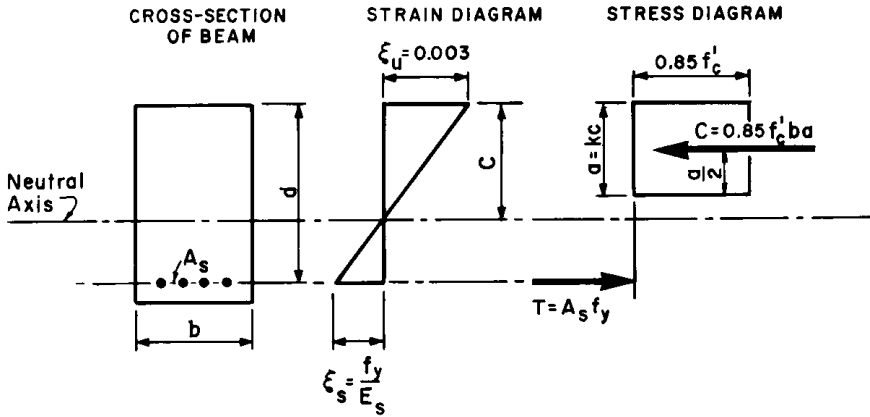
FIGURE 14.10 Assumptions in ultimate-strength design.

Rectangular Beams with Tensile Reinforcing Only. This type of beam includes slabs [for which $b = 12$ in (304.8 mm) when the moment and shear are expressed per foot of width]. The ultimate resisting moment is

$$M_u = \phi[bd^2 f_c' q(1 - 0.59q)] = \phi \left[A_s f_y \left(d - \frac{a}{2} \right) \right] \qquad (14.104)$$

where A_s = area of tensile reinforcement, in² (mm²)
$\quad a$ = depth of rectangular stress block = $A_s f_y / 0.85 f_c' b$, in (mm)
$\quad b$ = width of compressive face of flexural member, in (mm)
$\quad d$ = distance from extreme compression fiber to centroid of tension reinforcement, in (mm)
$\quad f_c'$ = compressive strength of concrete, lb/in² (kPa)
$\quad f_y$ = yield strength of reinforcement, lb/in² (kPa)
$\quad M_u$ = ultimate resisting moment, in · lb (N · m)
$\quad p$ = A_s / bd, $< 0.75 p_b$
$\quad p_b$ = reinforcement ratio producing balanced conditions*
$\quad q$ = $A_s f_y / bd f_c' = p f_y / f_c'$
$\quad \phi$ = capacity reduction factor, in concrete code
$\quad k$ = concrete stress intensity factor = 0.85 when $f_c' = \leq 4000$ lb/in² (27.6 MPa)

T Beams with Tensile Reinforcing Only. For preliminary designs, the formulas given above for rectangular beams with tensile reinforcing only can be used, since the neutral axis is usually in or near the flange. The area of tensile reinforcing will usually be critical.

---

*Balanced conditions occur when the tension reinforcement is at its yield strength f_y and the concrete in compression is at its assumed ultimate strain of 0.003; balanced conditions exist when

$$p_b = \left(\frac{0.85 k f_c'}{f_y} \right) \left(\frac{87,000}{87,000 + f_y} \right) \qquad (14.104a)$$

Beams with Tensile and Compressive Reinforcing. Beams with compression reinforcing are generally used when the size of the beam is limited, or where $p > 0.75p_b$. The ultimate resisting moment is then

$$M_u = \phi \left[(A_s - A_{sc})f_y \left(d - \frac{a}{2} \right) + A_{sc}f_y(d - d') \right] \qquad (14.105)$$

where A_s = area of tensile reinforcement, in² (mm²)
  A_{sc} = area of compressive reinforcement, in² (mm²)
  a = depth of equivalent rectangular stress block = $(A_s - A_{sc})f_y/0.85f'_c b$, in (mm)
  b = width of compressive face of flexural member, in (mm)
  d = distance from extreme compression fiber to centroid of tension reinforcement, in (mm)
  d' = distance from extreme compression fiber to centroid of compression steel, in (mm)
  f_y = yield strength of reinforcement, lb/in² (kPa)
  M_u = ultimate resisting moment, in · lb (N · m)
  p = A_s/bd
  p' = A_{sc}/bd $(p - p' < 0.75p_b)$
  p_b = reinforcement ratio producing balanced conditions, Eq. (14.104a)
  ϕ = capacity reduction factor

Shear and Diagonal Tension in Beams. The ultimate shearing unit stress, as a measure of diagonal tension, in a reinforced-concrete beam is

$$\nu_u = \frac{V_u}{bd} \qquad (14.106)$$

where ν_u = ultimate shearing unit stress, lb/in² (kPa)
  V_u = ultimate total shear, lb (N)
  b = width of beam (for T beam use width of stem), in (mm)
  d = effective depth of beam, in (mm)

For design, the maximum shear is considered to occur at a distance d from the face of the support. The shear stress carried by concrete ν_c should not exceed $2\phi\sqrt{f'_c}$. Wherever the ultimate shear stress ν_c, exceeds the shear stress ν_c, web reinforcement is mandatory. Such reinforcement will usually consist of stirrups. The cross-sectional area required for a stirrup placed perpendicular to the longitudinal reinforcement is

$$A_u = \frac{V_u s}{\phi f_y d} \qquad (14.107)$$

where A_u = total area of web reinforcement in tension within a distance measured in a direction parallel to the longitudinal reinforcements, in² (mm²)
  d = effective depth, in (mm)
  f_y = yield strength of reinforcement, lb/in² (kPa)
  s = spacing of stirrups, in (mm)
  V_u = total ultimate shear, lb (N)
  ϕ = capacity reduction factor, in design codes

The shear stress ν_u should not exceed $10\phi\sqrt{f'_c}$ and f_y should not exceed 60,000 lb/in² (413.6 MPa). Stirrups should be anchored at both ends to be considered

effective and so spaced that every 45° line extending from the middepth of the beam to the longitudinal tension bars is crossed by at least one stirrup.

Bond and Anchorage for Reinforcing Bars.   In beams in which the tensile reinforcing is parallel to the compression face, the bond stress on the bars is

$$u_u = \frac{V_u \phi}{jd \, \Sigma_0} \tag{14.108}$$

where u_u = ultimate bond stress on surface of bar, lb/in² (kPa)
  V_u = total ultimate shear, lb (N)
  jd = distance between centroid of compression and centroid of tension, in (mm)
  Σ_0 = sum of perimeters of tensile reinforcing bars, in (mm)
  ϕ = capacity reduction factor, in design codes
  d = effective depth of beam, in (mm)

Columns—General.   Columns should be designed for the axial load computed by using Eqs. (14.101), (14.102), and (14.103) and for the actual eccentricity e of the applied loading which should not be less than 0.05t for spirally reinforced columns or 0.10t for tied columns where t is the overall depth of a rectangular section or diameter of a circular section. The maximum load capacities given by Eqs. (14.109) through (14.114) apply only to short columns; adjustment factors for long columns are given by Eqs. (14.92) and (14.93).

Short Columns—Rectangular.   The ultimate strength of short rectangular columns where the reinforcement is in one or two faces, each parallel to the axis of bending and all reinforcement in any one face is located at approximately the same distance from the axis of bending, is computed by the empirical formulas

$$P_u = \phi(0.85 f'_c ba + A_{sc} f_y - A_s f_s) \tag{14.109}$$

$$P_u e' = \phi \left[0.85 f'_c ba \left(d - \frac{a}{2} \right) + A_{sc} f_y (d - d') \right] \tag{14.110}$$

where a = depth of equivalent rectangular stress block, in (mm)
  A_g = gross area of section, in² (mm²)
  A_s = area of tension reinforcement, in² (mm²)
  A_{sc} = area of compression reinforcement, in² (mm²)
  A_{st} = total area of longitudinal reinforcement, in² (mm²)
  b = width of compression face of flexural member, in (mm)
  d = distance from extreme compression fiber to centroid of tension reinforcement, in (mm)
  d' = distance from extreme compression fiber to centroid of compression reinforcement, in (mm)
  D = overall diameter of circular section, in (mm)
  D_s = diameter of the circle through centers of reinforcement arranged in a circular pattern, in (mm)
  e = eccentricity of axial load at end of member measured from plastic centroid* of the section, calculated by conventional methods of frame analysis, in (mm)

---

*Centroid of the resistance to load computed for assumptions that concrete is stressed uniformly to O.25f'_c and steel is stressed uniformly to f_y.

e' = eccentricity of axial load at end of member measured from the centroid of the tension reinforcement, calculated by conventional methods of frame analysis, in (mm)

f'_c = compressive strength of concrete, lb/in^2 (kPa)

f_s = calculated stress in reinforcement when less than the yield strength f_y, lb/in^2 (kPa)

f_y = yield strength of reinforcement, lb/in^2 (kPa)

$m = f_y/0.85f'_c$

P_u = axial load capacity under combined axial load and bending, lb (N)

$p_t = A_{st}/A_g$

t = overall depth of a rectangular section or diameter of a circular section, in (mm)

ϕ = capacity reduction factor, in design codes

Short Columns—Circular. The ultimate strength of short circular columns with reinforcing bars circularly arranged is computed by the following empirical formulas.

When tension controls:

$$P_u = \phi \left\{ 0.85 f'_c D^2 \left[\sqrt{\left(\frac{0.85e}{D} - 0.38\right)^2 + \frac{p_t m D_s}{2.5D}} - \left(\frac{0.85e}{D} - 0.38\right) \right] \right\}$$

(14.111)

When compression controls:

$$P_u = \phi \left\{ \frac{A_{st} f_y}{(3e/D_s) + 1} + \frac{A_g f'_c}{[9.6De/(0.8D + 0.67D_s)^2] + 1.18} \right\} \quad (14.112)$$

Short Columns—square. The ultimate strength of short square columns with reinforcing bars circularly arranged is computed by the following empirical formulas.

When tension controls:

$$P_u = \phi \left\{ 0.85 b t f'_c \left[\sqrt{\left(\frac{e}{t} - 0.5\right)^2 + 0.67 \frac{D_s}{t} p_t m} - \left(\frac{d}{t} - 0.5\right) \right] \right\} \quad (14.113)$$

When compression controls

$$P_u = \phi \left\{ \frac{A_{st} f_y}{(3e/D_s) + 1} + \frac{A_g f'_c}{[12te/(t + 0.67D_s)^2] + 1.18} \right\} \quad (14.114)$$

Equations (14.109) through (14.114) for the ultimate strength of columns subjected to combined axial compression and bending require assumption of a concrete cross section and reinforcement, computation of an allowable ultimate load and eccentricity, and comparison of the allowable loading with the design loading. This procedure usually requires iteration and may involve several assumptions regarding member size and amount of reinforcement. In practice, columns are usually designed by reference to extensive design tables, such as appear in *Concrete Reinforcing Steel Handbook,* which lists allowable axial loads and eccentricities for a large selection of cross sections and reinforcement arrangements.

ECONOMIC, SOCIAL, AND ENVIRONMENTAL CONSIDERATIONS

The scope of civil engineering includes many types of capital projects where the expenditure of funds must be justified by economic and financial analyses. Major civil engineering projects usually require analysis to determine their effect on community and regional social structures and on the environment.

40. Economic Analyses.   Economic analyses compare the economic benefits of a project with its economic costs, while financial analyses compare the monetary return from a project with its financial costs. Economic analyses may be summarized in terms of a benefit-cost ratio using the formula

$$B/C = \frac{\Delta t + \Delta u + \Delta a}{\Delta c + \Delta m} \tag{14.115}$$

where B/C = benefit-cost ratio
Δt = present worth of reduction in user time costs (computed on annual basis over the life of a project, discounted to present worth)
Δu = present worth of reduction in user operating costs (computed on annual basis over the life of a project, discounted to present worth)
Δa = present worth of reduction in user accident and damage costs (computed on annual basis over the life of a project, discounted to present worth)
Δc = present worth of incremental costs of initial and recurring capital investments over the life of a project
Δm = present worth of increase in maintenance and operating costs (computed on annual basis over the life of a project, discounted to present worth)

In some cases the benefits of an engineering project are determined as the net value of increased farm or industrial output attributable to the project.

In Eq. (14.115) the interest cost to be used is the opportunity cost of capital or market rate of interest; the opportunity cost of capital generally ranges from 8 to 15 percent. The life of a project is determined from consideration of the inherent durability of project components but should not exceed the period in which technological obsolescence may occur. For most construction projects, the life will range up to 30 years, but longer lives may be used for major projects. The benefit-cost ratio (B/C) indicates the economic justification of a project. A project with a benefit-cost ratio of one is marginal; values higher than one indicate greater justification.

An internal rate of return calculation is often used instead of a benefit-cost ratio. The internal rate of return is determined by setting B/C in Eq. (14.115) equal to unity and determining the interest rate used for discounting at which the numerator equals the denominator. The interest rate is then compared with the opportunity cost of capital.

Benefit-cost or internal rate of return calculations should be made on an incremental basis for comparison among alternatives. One alternative which must be included is the "do nothing" alternative. Incremental analyses among alternatives permit the determination of the incremental return obtained from successively greater capital investments.

HYDRAULIC ENGINEERING

HYDRAULIC TURBINES

1. Types. There are two types of turbine in general use, the *tangential water wheel* and the *reaction turbine.* The reaction type is further subdivided into the *mixed-flow reaction* and the *propeller* types.

2. Homologous Machines. When two machines are identical except as to size, they are said to be *homologous.* In a line of homologous turbines all of that line have the same proportions for their respective parts, and the parts are similarly located. For such a line of machines, the performance of one having been found by test, the performance of the others at any head may be found as follows: Let N = r/min, q = discharge in ft^3/s (m^3/s), P = brake hp (kW), D = diameter in in (cm), h = head in ft (m). Then

$$N = \frac{1840\phi\sqrt{h}}{D} \tag{14.116}$$

$$q = k_1 D^2 \sqrt{h} \tag{14.117}$$

$$P = k_2 D^2 h^{1.5} \tag{14.118}$$

The constants ϕ, k_1, and k_2 are typical of that line of machines, and are obtained from the test of the first machine.

3. Specific Speed. Another turbine constant of great utility is variously called characteristic speed, type characteristic, or specific speed, N_s:

$$N_s = \frac{N\sqrt{P}}{h^{1.25}} \tag{14.119}$$

The value of N_s is a criterion of type, as follows:

    Tangential water wheel, N_s up to 6
    Mixed-flow turbine, N_s = 20 to 120
    Propeller turbine, N_s = 120 to 200 or higher

Experience has shown that, for satisfactory operation of the mixed-flow turbine, the specific speed should not exceed the value

$$N_s = \frac{5050}{h + 32} + 19 \tag{14.120}$$

4. Efficiency. A well-designed tangential water wheel will operate with an efficiency of 85 to 90 percent, in the usual sizes found in power plants. The mixed-flow reaction turbine, with a specific speed between 40 and 80, may be expected to have an efficiency of 88 to 94 percent. The propeller-type machines, with specific

speeds between 100 and 150, have shown efficiencies as high as 90 percent. For any type, the higher efficiency will be associated with the larger machine operating under moderate head for its class. To estimate the efficiency of a large machine when a test on a small homologous model is available, the following empirical formula by L. F. Moody is extensively used:

$$E = 1 - (1 - e) \left(\frac{d}{D}\right)^{1/4} \left(\frac{h}{H}\right)^{1/10} \tag{14.121}$$

where E is maximum efficiency of large turbine; e is maximum efficiency of small turbine; d and D are diameters of small and large turbines, respectively; h is the head under which the model is tested; and H is the head under which the large machine will operate. Another formula sometimes used is the Cammerer formula:

$$E = 1 - (1 - e) \frac{1.4 + (1/D)}{1.4 + (1/d)} \tag{14.122}$$

In those cases where the two formulas do not agree, the Moody formula is likely to give results which are too high, while the Cammerer results will be too low.

5. Diameter of Turbine.   To determine the size of machine required in any re-action-turbine installation, Fig. 14.11 may be used. For the given head, the specific speed to be used may be determined from Eq. (14.120). Since the figure shows the diameters of the machine which will deliver 1 hp (0.75 kW) at 1-ft (0.3-m) head, the actual diameters may be found by the use of Eq. (14.118), which for this case may be rearranged as follows: $D = \sqrt{P}/\sqrt[4]{h^3}$ multiplied by the diameter given in Fig. 14.11.

For the impulse wheel an approximate diameter of the impulse circle can be found from $D = 860 \sqrt{P}/N_s\sqrt[4]{h^3}$. The impulse circle diameter is twice the perpendicular distance from the center of the shaft to the jet from the nozzle. The maximum overall diameter of the wheel will be about 15 percent greater, but varies somewhat with the design. The diameter of the jet from the nozzle is, then,
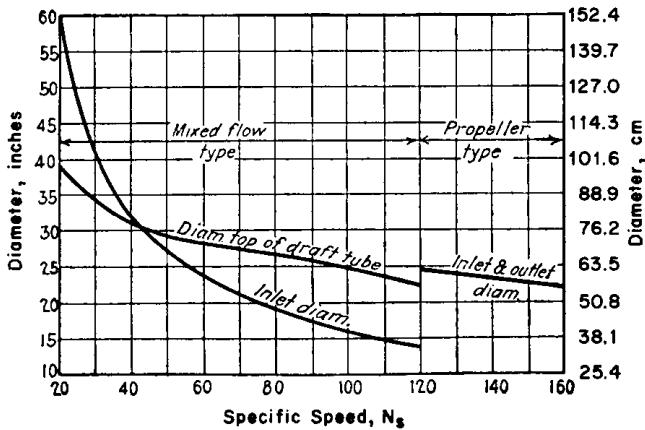
FIGURE 14.11   Turbine dimensions for 1 hp at 1-ft head (0.75 kW at 0.3 m).

$$d \text{ (in)} = \frac{DN_s}{55} \qquad [= 2.5d \text{ (cm)}]$$

In order to find the speed at which any of these machines will run, Eq. (14.119) may be used.

6. Selection of Turbines.   The specific speed, as defined by Eq. (14.119), is all that is necessary to determine the type of turbine to be used. Since both the *speed* and *power* of a unit can be varied for any given installation, it is the choice of these which will determine the type of turbine. The *power* of a unit will be selected from considerations of size of the development and the load to be served. From the standpoint of stream flow, the size of unit should be small enough so as to run at or near full load or best efficiency during the low-water season when the flow is a minimum. Having determined from the stream-flow record the minimum load which the plant is to carry, the size of unit should be but little larger than this. It should be small enough, on the other hand, so that not less than about three machines will be required altogether unless the plant is one in a large system in which breakdown of one unit will not disturb the ability to carry the load. With an isolated water-power plant, one extra machine should be provided for standby in case of damage or repairs to any one machine. Thus, for a 100,000-hp (74.6-MW) plant, install 5 units at 25,000 hp (18.7 MW), or 6 units at 20,000 hp (14.9 MW) or 11 units at 10,000 hp (7.5 MW).

The size of unit having been chosen, the *speed* is chosen either to accommodate the load to be driven or to require a type of turbine best suited for the installation. For a generator directly connected, as is the case in a great majority of hydraulic plants, the higher the speed the cheaper will be the generator for heads under 60 ft (18.3 m). In these cases the speed is normally selected for hydraulic reasons. For the mixed-flow reaction turbine the specific speed should not exceed the value given by Eq. (14.120); for the propeller-type, present practice limits specific speed to a maximum value of about 175, although higher values can be had. Where a definite choice can not be made on this basis alone, cost and efficiency must be estimated and used as a basis for a final decision. For heads above 60 ft (18.3 m), only the mixed-flow type is now being used, and installations have been made under heads as high as 850 ft (259 m). Within this range, the speed is normally selected by means of Eq. (14.120), except for the higher heads, with which lower speeds are necessary because of generator limitations. When the tangential water wheel is used, the limited economical range of specific speed fixes the revolutions per minute within rather narrow limits.

CHAPTER 15

CHEMICAL ENGINEERING, ENVIRONMENTAL ENGINEERING, AND PETROLEUM AND GAS ENGINEERING

CHEMICAL ENGINEERING*

DIFFUSIONAL OPERATIONS

1. Mass-Transfer Fundamentals. The transfer of material from one phase to another is a primary means of separating multicomponent solutions. In general, two equilibrium phases of a multicomponent mixture will have different chemical compositions, and this difference offers a means for separating a mixture into its individual components. Repetitive phase changes can provide increasingly pure solutions and in the limiting case can produce pure individual components. Analysis of these phase-change separation processes depends on three factors: thermodynamic equilibrium, mass-transfer rates, and pattern of contact between phases.

Mass-Transfer Rates. The amount of contacting required to bring two phases into equilibrium is dependent on the rate of mass transfer. The rate at which mass is transferred between phases is controlled by the *driving force* for mass transfer, the *resistance* to mass transfer, and the *interfacial area* between phases, according to

$$N_{M_i} = K_L aV(x_i - x_i^*) = K_G aV(y_i - y_i^*) \tag{15.1}$$

The overall mass-transfer coefficients are dependent on resistance to mass transfer in interfacial films (in a manner analogous to film resistances in convective heat transfer), which depend on molecular parameters, fluid turbulence near the interface, and the equilibrium relationship between phases. For equilibrium relationships which are essentially the straight lines $y = mx$, mass-transfer coefficients are given by

---

*This section is from *Engineering Manual,* 3d ed., by R. H. Perry (ed.). Copyright © 1976. Used by permission of McGraw-Hill, Inc. All rights reserved. Updating and SI units were added by the editors in the year 2002.

15.1

$$\frac{1}{K_G} = \frac{1}{k_G} + \frac{m}{k_L} \tag{15.2}$$

$$\frac{1}{K_L} = \frac{1}{k_L} + \frac{1}{mk_G} \tag{15.3}$$

The coefficients are ordinarily determined experimentally as volumetric coefficients $K_G a$ and $K_L a$, since interfacial areas are difficult to determine.

Continuous Contacting. Continuous-contact processes may be run with concurrent, crosscurrent, or countercurrent flow patterns. In terms of efficient mass transfer, countercurrent contacting, which allows the greatest amount of mass transfer between phases for a given initial composition difference between feed streams, is preferred.

The *transfer unit* is a standard degree of separation used to describe the performance of a contacting device. The greater the number of transfer units, the more thorough is the separation. The number of transfer units required to accomplish a specified degree of separation between components is dependent on the product compositions and relative flow rates of the two phases and may be based on either phase, according to

$$N_{OG} = \int_{y2}^{y1} \frac{(1 - y)_{lm}}{(1 - y)(y - y^*)} \, dy \cong \int_{y2}^{y1} \frac{dy}{y - y^*} \tag{15.4}$$

$$N_{OL} = \int_{x1}^{x2} \frac{(1 - x)_{lm}}{(1 - x)(x - x^*)} \, dx \cong \int_{x2}^{x1} \frac{dx}{x - x^*} \tag{15.5}$$

The approximation holds for dilute solutions, where transfer of material does not change the overall stream molar flow rates, and is a true equality for equimolar counter diffusion.

The *height of a transfer unit* (*HTU*) is an indication of the amount of contacting required to accomplish the standard separation of one transfer unit. It is dependent on mass-transfer coefficients, packing type, and specific surface area, flow patterns of the contacting phases, and flow rates. The heights of the overall transfer units defined by Eqs. (15.4) and (15.5) are

$$H_{OG} = \frac{G_M}{K_G a(1 - y)_{lm}} \tag{15.6}$$

$$H_{OL} = \frac{L_M}{K_L a(1 - x)_{lm}} \tag{15.7}$$

The height of a transfer unit is often much more constant that the mass-transfer coefficient for a given column and packing at various flow rates and is therefore more commonly used. HTUs are determined empirically by measuring mass transfer in a given column, calculating N_{OG} or N_{OL} by Eq. (15.4) or (15.5) and dividing into column height, that is,

$$H_{OG} = \frac{Z}{N_{OG}} \tag{15.8}$$

$$H_{OL} = \frac{Z}{N_{OL}} \tag{15.9}$$

In design applications, the number of transfer units required for a specified separation is calculated by Eq. (15.4) or (15.5). This is combined with HTU data for similar (or pilot plant) operations to obtain total column height.

Nomenclature for Eqs. (15.1) through (15.9) corresponds to that used for gas absorption. The equivalent equations for other continuous contacting operations have slightly different forms appropriate to the parameters commonly encountered. A thorough treatment of mass transfer is found in Treybal's comprehensive text.[5]

Staged Operations. In staged contacting, two phases initially not in equilibrium are held in contact for a length of time assumed sufficient to attain equilibrium. The two phases are then separated, and each is fed to an adjacent stage where it is again held in contact with a nonequilibrium mixture of the opposite phase. The two phases flow from stage to stage in opposite directions. As one phase advances through the contactor it becomes progressively more concentrated in a particular component or group of components.

2. Distillation

Definitions. *Simple distillation* is the partial vaporization of a solution of liquid components with separate recovery of vapor and liquid residue. The concentration of the more volatile components (sometimes termed the *lighter components*) is greatest in the condensed vapor, while the concentration of the less volatile components (the *heavier components*) is greatest in the liquid residue.

Rectification is continuous countercurrent contact of the vapor resulting from a simple distillation with a condensed portion of the vapor product. This countercurrent contact results in greater enrichment of the vapor (often termed *overhead*) product with the more volatile components than is possible with a single stage of simple distillation. The condensed vapor returned to accomplish this is termed *reflux.* In rectification, the feed is to the simple distillation stage (bottom) and the more important product is removed as a vapor (top).

Stripping is continuous countercurrent contacting of the liquid feed with the vapor resulting from a simple distillation. This countercurrent contacting results in a more complete removal of the more volatile components from the liquid product than could be accomplished by a single stage of simple distillation. In stripping, the liquid feed is introduced at the top of the column and the vapor is supplied by partial vaporization of the liquid stream at the bottom of the column in a *reboiler*.

Commonly, rectification is carried out on the vapor product leaving a stripping operation. This is termed *fractional distillation,* or *fractionation.* In fractional distillation, column feed is introduced at a *feed stage* within the column. Above the feed stage is the rectification section, and below it is the stripping section. Overhead product is condensed, and a portion is returned as reflux. Bottom product is partially vaporized, and the vapor is returned to the column as *boilup*.

Equilibrium Data. Vapor-liquid equilibrium data are required for the design of stills. In general, these must be observed experimentally. For binary systems, they are reported usually as tables or graphs of corresponding x and y values. Many such data are summarized by Hala et al.,[6] Smith, Block, and Hickman,[7] and by Hala et al.[8]

If equilibrium data are not available, they may be estimated for a binary system by a modification of Raoult's law:

$$y_1 = \frac{P_1}{\Pi} = \frac{\gamma_1 x_1 P_1^0}{\Pi} = \frac{\alpha_{12} x_1}{1 + (\alpha_{12} - 1)x_1} \tag{15.10}$$

For ideal mixtures, $\gamma_1 = \gamma_2 = 1.0$; as a system departs from ideality, use of Eq. (15.10) becomes less reliable. Values of γ may be estimated for a number of binaries by a method summarized in Smith, Block, and Hickman.[7]

Simple Batch Distillation (Rayleigh or Differential Distillation). In this case, a batch of material is charged to a still pot, boiling is initiated, and the vapors are continuously removed, condensed, and collected until their average composition has reached a desired value. As the distillation proceeds, the concentration of less volatile components in the vapor continually increases.

Simple distillation is analyzed by assuming that at any instant the vapors are in equilibrium with the average liquid composition in the still. Though not strictly correct, this is usually a good assumption. In this case, the composition of the liquid in the still is related to the total amount vaporized by

$$\ln \frac{S_2}{S_1} = \int_{x1}^{x2} \frac{dx}{y - x} \tag{15.11}$$

where y is related to x by the equilibrium relationship $y = y(x, \Pi, t)$. The average composition of the vapors collected during this period is related to the initial and final liquid compositions and quantities by a simple material balance:

$$x_{\text{ave}} = \frac{x_1 S_1 - x_2 S_2}{S_1 - S_2} \tag{15.12}$$

Equilibrium Flash Distillation. Liquid at an elevated temperature and pressure is throttled into a *flash chamber* maintained at a pressure below the vapor pressure of the liquid. A certain fraction of the feed vaporizes, reducing the system temperature until the vapor pressure of the remaining liquid at the new temperature matches the flash chamber pressure. Vapor and liquid fractions are led separately from the still. If vapor and liquid streams are in equilibrium, this is equivalent to a single-stage simple distillation.

Compositions and flow rates of vapor and liquid streams are obtained by simultaneous solution of material balances

$$(1 - f)x_i + fy_i = x_{Fi} \tag{15.13}$$

and energy balance

$$(t_F - t)C_F = F\gamma_m \tag{15.14}$$

together with the equilibrium relationship

$$y_i = y(x_i, \Pi, t) \tag{15.15}$$

and the relationship between temperature, vapor pressure, and total pressure

$$P_i = P_i(t) \tag{15.16}$$

$$\Pi = \Sigma P_i(t) \tag{15.17}$$

where f is the fraction vaporized, and C_F and γ_m are the heat capacity and latent heat of vaporization of the feed (assumed constant). Equations (15.13) through (15.17) are solved simultaneously by trial and error.

Continuous Binary Rectification

Plate Columns—Plate-to-Plate Calculations.   The simultaneous solution of material balance, energy balance, and equilibrium relationships between each successive two stages in a column permits the exact computation of the column behavior from one terminal stream to another.

McCabe-Thiele Graphical Method.   If the molar latent heat of vaporization is relatively independent of composition, and if heat losses are negligible, liquid and vapor molar flow rates within the column will be constant and the following graphical procedure is acceptable:

1. Plot the vapor-liquid equilibrium data as y versus x.

2. Write and plot the operating-line (material balance) equations for each section of the column which relate passing streams within the column to feed or product streams. With reference to Fig. 15.1, the equations are, for section II (enriching section),

$$y_n = x_{n+1} \frac{L_{n+1}}{V_n} + x_d \frac{D}{V_n} = \frac{R}{R+1} x_{n+1} + \frac{x_d}{R+1} \tag{15.18}$$

where $R = L_E/D = $ *reflux ratio* and for section III (stripping section),

$$y_m = x_{m+1} \frac{L_{m+1}}{V_m} - x_w \frac{W}{V_m} \tag{15.19}$$

The liquid rates in the two sections of the column are constant but generally different because of the introduction of the feed, which may be liquid, vapor, or a mixture (see below).

3. Determine the number of stages by stepping from operating line vertically to equilibrium line, then horizontally to operating line. Begin with bottom product and continue in this fashion until the feed composition is reached; then use upper operating line until upper product concentration is reached.

4. Correct the theoretical number of stages to the actual number of stages using stage efficiencies (estimated from previous experience) according to

$$N_{\text{act}} = \frac{N_{\text{theor}}}{\text{efficiency}} \tag{15.20}$$

The upper and lower operating lines intersect on a *feed line* described by Eq.
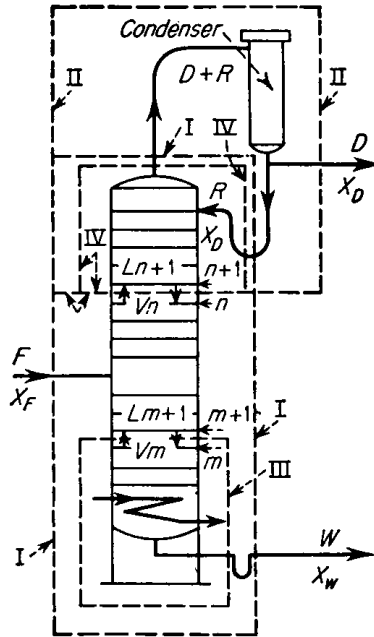
FIGURE 15.1 Schematic of continuous distillation column.

(15.13), where f is the moles of vapor introduced to the rectifying section per mole of feed introduced to the fed stage. Note that f is fractional between zero and unity if the feed is introduced as a mixture of liquid and vapor; $f = 0$ if feed is a saturated liquid; $f < 0$ if feed is a cold liquid; $f = 1$ if feed is saturated vapor; and $f > 1$ if feed is a superheated vapor. In Fig. 15.2a, five different feed lines are shown with the same upper operating line and five lower operating lines leading to the
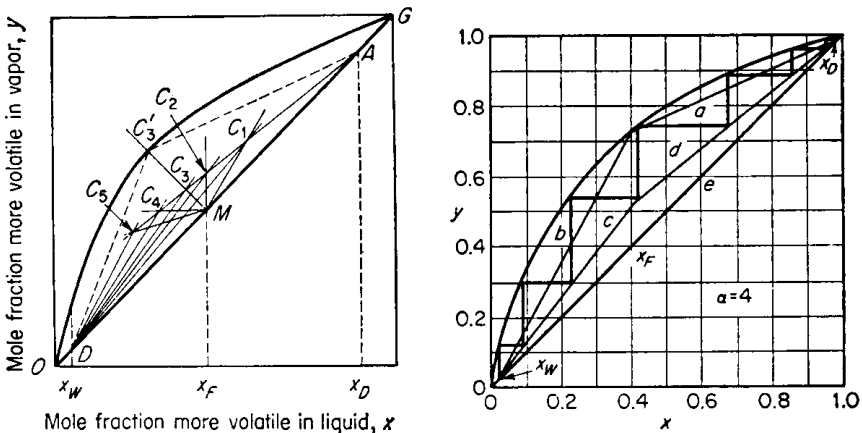
FIGURE 15.2 (*a*) Effect of thermal condition of feed on operating lines and minimum reflux ratio. (*b*) McCabe-Thiele graphical method for determining number of theoretical stages.

same bottom product composition. Intersections between operating lines for feeds of different qualities are represented by C_1 through C_5:

$$C_1(f < 0) \quad C_2(f = 0) \quad C_3((0 < f < 1) \quad C_4(f = 1) \quad \text{and} \quad C_5(f > 1)$$

Three sets of operating lines are shown on Fig. 15.2b. Lines a and b, intersecting the feedline on the equilibrium curve, represent minimum reflux and require an infinite number of theoretical stages to separate the feed with composition x_F into products with compositions x_d and x_w. Line e, where both operating lines coincide with the diagonal, represents total reflux, which requires the smallest number of stages to achieve the desired separation, but produces no product. Lines c and d represent practical operating lines between the limits of minimum and total reflux.

Analytical Method for Mixtures of Constant Relative Volatility

1. Solve for the number of theoretical plates necessary at "total" reflux (the condition when vapor and liquid rates within the column are infinitely large compared to feed, overhead, and bottom drawoff rates) using the Fenske-Underwood equation

$$N_T + 1 = \log (x_1/x_2)_d (x_2/x_1)_w / \log \alpha \tag{15.21}$$

2. Estimate the minimum reflux ratio (the ratio of liquid to distillate rates if the column were infinitely tall) by using

$$\frac{R}{R_{\min}} = \frac{x_d[1 + (\alpha - 1)x_F] - \alpha x_F}{(\alpha - 1)x_F(1 - x_F)} \tag{15.22}$$

3. By use of Fig. 15.3, estimate the number of theoretical trays necessary for the reflux ratio to be employed.

Batch Binary Rectification. In batch distillations, generally three "products" are withdrawn from the still. These are an initial product high in purity with regard to the more volatile or light-key component, an intermediate product that will usually
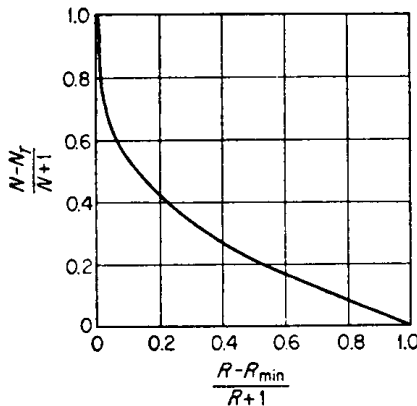
FIGURE 15.3 Correlation of Gilliland for number of theoretical stages.

be recycled for redistillation, and finally, a product high in purity with regard to the heavy or less volatile component. For most batch distillations, the following rule of thumb will hold and is useful in setting the reflux to be used:

$$\frac{L}{D}(\alpha - 1) \geq 10 \qquad (15.23)$$

If a constant-reflux ratio is maintained, the product purity of the more volatile component will drop off as the distillation proceeds. The speed at which this decline in purity occurs will be a function of the particular reflux ratio employed, the relative volatility, and the amount of volatile component originally present.

Of importance in planning for the operation of any batch-distillation apparatus are answers to the following questions:

1. What is the overhead-product composition as a function of still-pot composition?
2. How many moles of steam (as a heating medium) will be required to effect the separation?

For the constant-reflux case, the vapor requirement, and thus the steam requirement, is obtained from

$$V = (L/D + 1)D \qquad (15.24)$$

The relation between still-pot and overhead compositions is obtained by plotting on a $y - x$ diagram (as shown in Fig. 15.4) lines of constant slope equal to

$$L/V = (L/D)/(L/D + 1) \qquad (15.25)$$

and stepping off the number of theoretical plates in the column. To obtain the relation between the amount distilled and the still-pot composition, plot x_p versus $1/(x_p - x_d)$. The area under the curve is equal then to P/D, where P represents the amount of liquid in the still pot.

A second method of operating batch-distillation columns is to main product unity over a period of time by constantly increasing the reflux ratio. The relation between
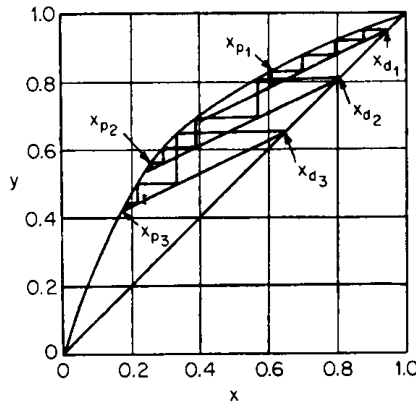
FIGURE 15.4 Diagram for batch distillation at constant reflux ratio.

amount distilled and still-pot composition is now found by a simple material balance. The steam requirement is obtained by finding the area under a curve of $[P_0(x_0 - x_p)]/[(x_p - x_d)^2(1 - L/V)]$ versus x_p. Appropriate values of x_p are read as a function of L/V from a $y - x$ plot.

Plate Efficiency.   The estimation of plate efficiencies is empirical. In a properly designed column, the value should exceed 0.60 and may exceed 0.95. Fair et al.[9] present estimation methods and typical values.

Limiting Factors.   Proper operation of a column requires gas velocities within a narrow range of values. Velocities must be high enough to give good gas-liquid contacting, yet low enough to prevent entrainment or excessive pressure drop with resultant flooding. These factors control column diameter for a given vapor rate.

3. Solvent Extraction

Definitions.   Solvent extraction consists of the transfer of a component dissolved in a liquid (called the *feed solution*) to a second liquid (called the *solvent*) to form an *extract* solution of the transferred component and to leave a *raffinate* solution relatively lean in the transferred component. Solvent extraction is used when distillation is impractical, as with close-boiling or temperature-sensitive mixtures. There is a strong analogy between extraction and distillation, solubility being the counterpart of volatility and the solvent that of heat (Fig. 15.5).

Equilibrium Data.   Phase equilibria for liquids are so specific that it is best to refer to laboratory data for the system in question. Maddox[10] cites many such data.

For some systems, the equilibrium is well approximated by the ideal-distribution law

$$K' = w/w' \tag{15.26}$$

with consequent simplification of design procedures. In solvent extraction, fewer theoretical plates and much lower plate efficiencies are encountered than in distillation or absorption, with corresponding aggravation of inaccuracies implicit in simplified methods. For this reason, shortcut approximations should be used with caution.

Countercurrent Extraction.   A feed mixture of two completely miscible components A and B is to be separated into its components by extraction with a solvent. If the solvent is partially miscible with each component of the feed, counter current extraction with reflux may be used to separate the feed components completely (in the limiting case). The method of operation in this case is presented schematically in Fig. 15.5.

Assume that the solvent is only partially miscible with each feed component, and the equipment is as shown in Fig. 15.5. Reflex is furnished to the top of the column by removing sufficient solvent from the extract phase (e.g., by distillation) to make it miscible with the raffinate phase. (Note that this does not change the ratio of A to B.) A portion of this raffinate phase is returned to the column as reflux; the remainder is purified further to remove any remaining solvent and is removed as product.
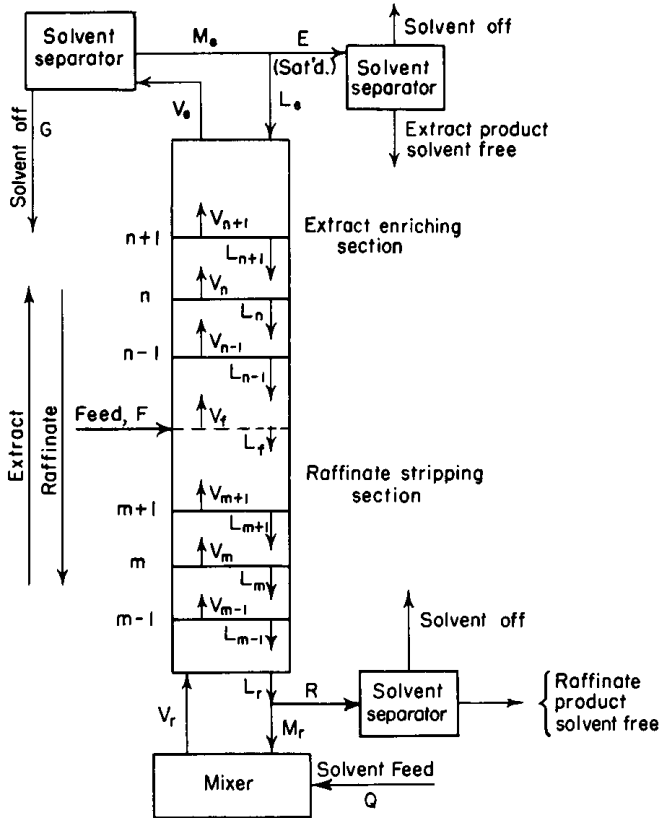
FIGURE 15.5 Flow diagram for countercurrent multistage extraction with reflux.

A simplified design procedure similar to McCabe-Thiele graphical method (see earlier subsection) may be used, assuming that extract-layer and raffinate-layer flow rates are constant throughout the column.

1. Equilibrium data are plotted as mass fraction of A in the extract layer (*ordinate*) against mass fraction of A in the raffinate layer (*abscissa*), *both fractions being on a solvent-free basis.*
2. Extract and raffinate products are located on the $Y = X$ line.
3. Operating lines through these points and of slope L_c/V_c and L_r/V_r are drawn.
4. Plates are stepped off as in the McCabe-Thiele method previously described.

A more precise design procedure is illustrated in Fig. 15.6 and is outlined below. All flow rates and concentrations are on a solvent-free basis, unless otherwise noted.

1. From known equilibrium relationships, construct the S versus Y (extract layer) and x versus X (raffinate layer) lines on a working diagram.
2. Locate the operating point K at an abscissa of X_c (extract-product composition) and an ordinate of $G/E + S_c$.
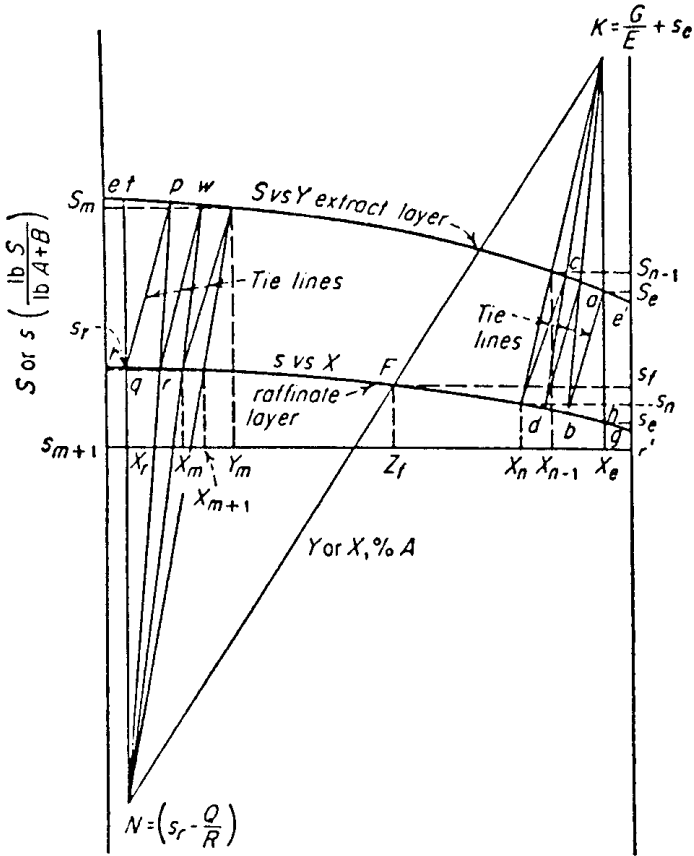
FIGURE 15.6 Graphical stepwise calculation of equilibrium stages on a solvent content-concentration diagram for operation with reflux.

3. Locate the operating point N at an abscissa of X_r (the A content of the raffinate product) and an ordinate of $(s_r - Q/R)$. A line joining K and N will intersect the s versus X line at Z_f (solvent content of the feed).

It is now possible to "walk" across the diagram to determine the number of theoretical stages necessary to effect the separation. A line is drawn from K to X_c intersecting the S versus Y curve at a. Line ab is an equilibrium tie line wherein the composition represented by b is that in equilibrium with the composition represented by a. Another line from K can than be drawn to the point b so established, and another equilibrium tie line cd is drawn to the point b so established, and another equilibrium tie line cd is drawn. The procedure is repeated until a ray from K coincides with the line joining K and N. To the left of this dividing line the same procedure is followed, using point N where point K was used before. The number of theoretical stages is then obtained by counting the total number of rays drawn from the two operating points.

Column Efficiency. For perforated-plate columns, the overall efficiency may be estimated as the fraction E_c:

$$E_c = \frac{89,500 Z_t^{0.5}}{\sigma g_c} \left(\frac{V_D}{V_c}\right)^{0.42} = \frac{0.9 Z_t'^{0.5}}{\sigma'} \left(\frac{V_D}{V_c}\right)^{0.42} \qquad (15.27)$$

where Z_t is tray spacing in feet, and Z_t' is tray spacing in inches.

For packed columns, the efficiency is expressed in the height assigned to a transfer unit (HTU) or theoretical stage (HETS). Ellis[11] shows that for *rough estimates,* the following empirical relationships are useful for towers packed with Raschig rings larger than $3/8$ in (9.5 mm):

1. Transfer of solute from continuous aqueous to dispersed organic phase:

$$\text{HETS} = \frac{94.5 \mu_C (12 d_F)^b (V_c / V_D)^{0.5}}{10^{0.0683s} \, \Delta \rho} \qquad (15.28)$$

2. Transfer of solute from dispersed organic to continuous aqueous phase:

$$\text{HETS} = \frac{69 \mu_C (12 d_F)^b}{10^{0.0535s} \Delta \rho} \qquad (15.29)$$

where $b = 2.15/10^{0.096s}$, d_F is in inches (millimeters), μ is in lb/ft · s (kg/m · s), and ρ is in lb/ft³ (kg/m³).

Extraction-Tower Diameter. Limiting flows, and hence minimum allowable diameters, for liquid-liquid extraction columns may be calculated using Fig. 15.7.

4. Gas Absorption

Definitions. *Gas absorption* consists of the transfer of a component from a gas phase to a liquid phase. The liquid phase is called the *solvent,* or *absorbent;* the transferred gas is called the *solute,* or *absorbate.* Usually, the solute is selectively
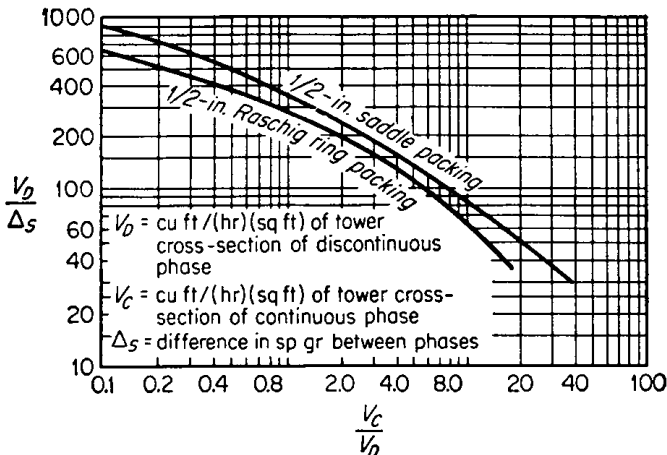
FIGURE 15.7 Colburn correlation of flooding data for packed extraction columns.

absorbed from a carrier gas. Fundamental considerations and design methods that apply to absorption are useful generally for the reverse operation of *desorption,* or *stripping.*

Equilibrium Data.   Gas solubility in a liquid is measured as a function of partial pressure or concentration of the gas in the equilibrium vapor phase. Solubilities sometimes are reported in the form of Henry's law constants. Equilibrium data for many systems may be found in standard reference sources.[2,12]

Equipment.   Gas absorption or stripping is accomplished in three principal types of equipment: *absorption columns,* packed or plate; *spray chambers* or *towers;* and *bubble-sparged tanks,* frequently agitated. Only absorption columns, by far the most important, will be treated here.

Column Height.   The height of a *packed column* is determined by the degree of separation to be achieved and by a characteristic contacting effectiveness of the packing. The former may be expressed by stream compositions or by number of transfer units [Eq. (15.4)]; the latter, by the appropriate transfer coefficient [Eq. (15.2)] or HTU [Eq. (15.6)]. The height of a transfer unit varies with application and should be determined experimentally for the gas-liquid system, column packing and column loadings employed.

The number of transfer units required may be calculated from Eq. (15.4) or, if operating and equilibrium lines are approximately straight, by Fig. 15.8, use of which requires knowledge of the slope of the equilibrium line m and stipulation of G_M/L_M. The latter should be an economic selection. For most columns, 0.7 is an acceptable value for mG_M/L_M, but something less may be used if the solute is of low economic value.

The number of ideal plates in a *plate column* may be determined by the McCabe-Thiele procedure (see earlier subsection). Once the values of K_L and L_G are estimated for the particular conditions under consideration and then combined to give a value for H_{OG} by use of Eqs. (15.6) and (15.7),

$$Z = H_{OG}N_{OG} \tag{15.30}$$

The estimation of N_{OG} may be obtained from Fig. 15.8, assuming that the operating and equilibrium lines are both straight or, at worst, only slightly curved.

5. Humidification.   The most frequently encountered gas-vapor system is that of air and water vapor. A distinct terminology has been developed for this system, but the principles and mechanisms involved apply to any gas-vapor system. *Humidification* is the process by which the moisture content of air is increased; *dehumidification* is the reverse process.

The concentration of water vapor in the air is expressed as the *absolute humidity* (pounds of water vapor per pound of dry air), the *molal humidity* (moles of water vapor per mole of dry air), *relative humidity* (ratio of the actual partial pressure of vapor to the partial pressure if saturated at the existing temperature), or *percentage humidity* (ratio of actual humidity to the humidity if saturated at the existing temperature). The moisture level in air is indicated by the *dew point* (the temperature to which an air sample must be cooled to reach saturation) and by the difference between the air temperature and the *wet-bulb temperature* (the temperature assumed by a water-wet body held in a fast-flowing stream of the air). The preceding prop-
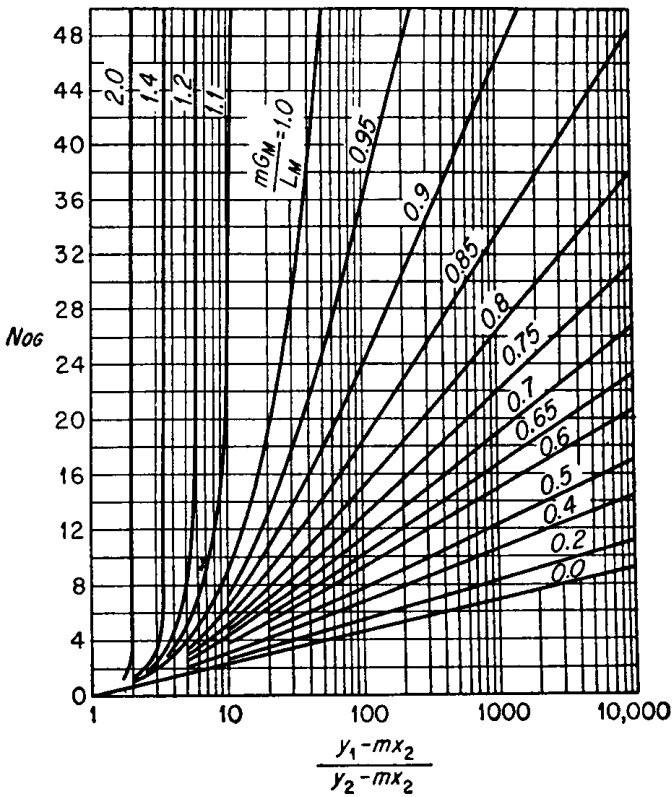
FIGURE 15.8 Number of transfer units in an absorption column. Subscripts 1 and 2 refer to the concentrated and dilute ends, respectively.

erties for the air–water vapor system are commonly presented graphically on a humidity or psychometric chart as shown in Fig. 15.9.

Humidification and dehumidification are often the unavoidable secondary results of other operations such as drying, compression, absorption, and water cooling. The evaporative cooling of water is accomplished by adiabatically contacting it with a relatively large amount of unsaturated air; the resultant vaporization of some of the water cools the remainder, which approaches wet-bulb or adiabatic saturation temperature corresponding to the condition of the air.

The contacting of air and water is accomplished with spray ponds or cooling towers. Spray ponds depend on natural forces for air movement, but power must be supplied to atomize and spray the water into the air. Spray ponds require large land areas, and water loss through entrainment is often high. Spray-nozzle performance is critical in the design, since it controls power demand, air-water contact surface, and water loss by drift.

The design of induced-draft towers is based on the allowable air velocity, the cooling load and range, the approach to wet-bulb temperature, and the transfer coefficients attained.

Norris[13] gives a design procedure, but the services of a reputable tower supplier are recommended. As a very rough estimate, 1 ft (0.3 m) of tower height will yield 1°F (0.56°C) of cooling when the water rate is 2 to 3 gpm/ft^2 (81.5 to 122.2 L/
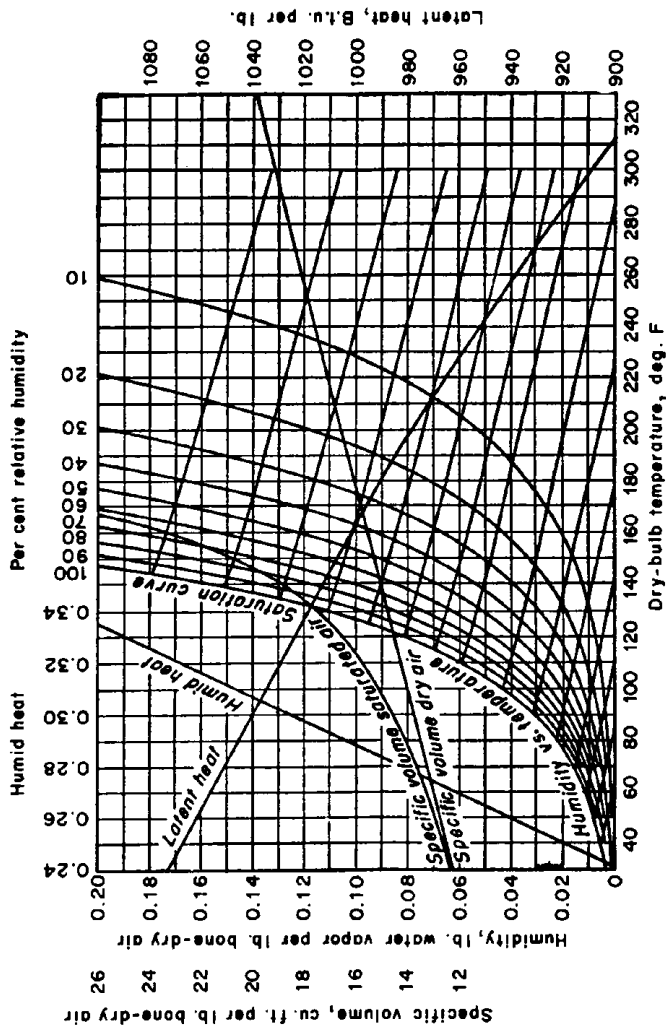
FIGURE 15.9a Humidity chart for air-water vapor mixtures, U.S. Customary units.
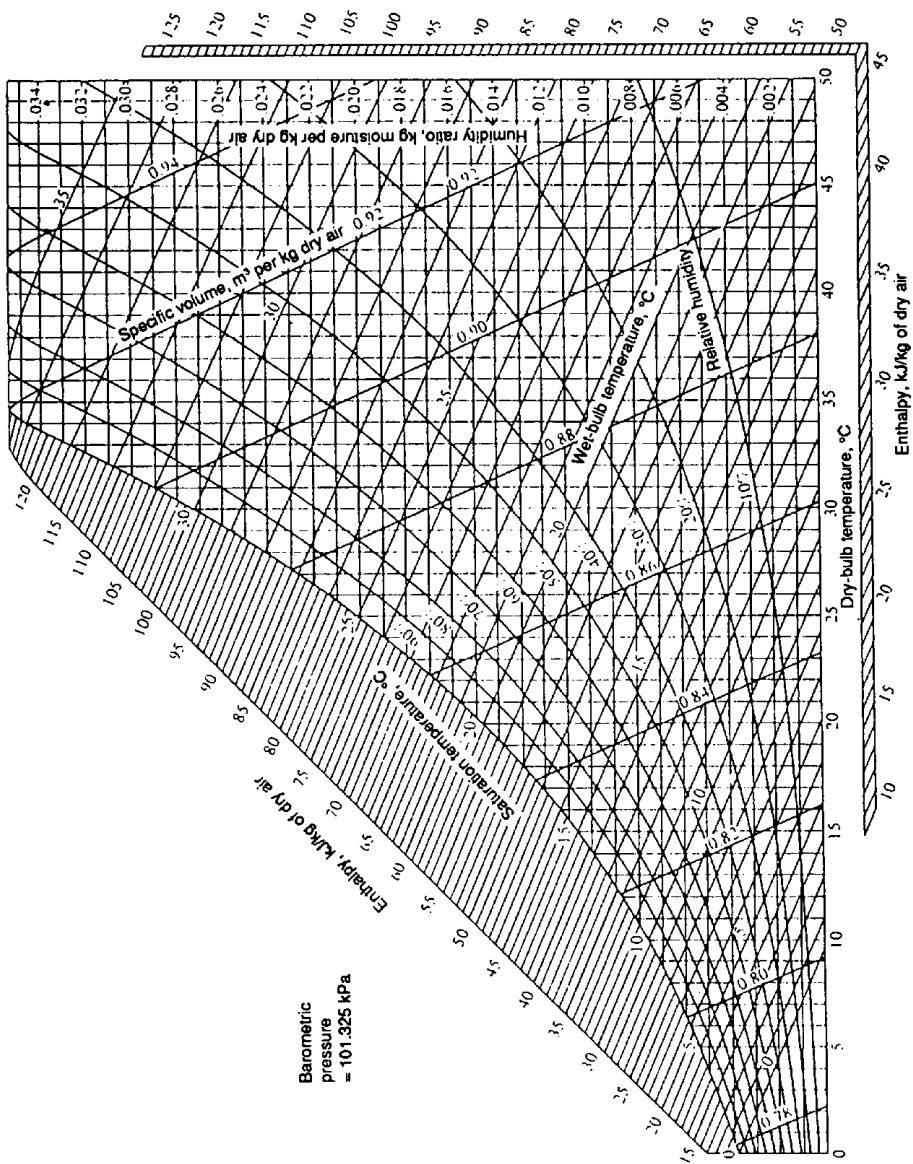
FIGURE 15.9b Humidity chart for air-water vapor mixtures, SI units. (*From W. F. Stoecker and J. W. Jones, Refrigeration and Air Conditioning, 2d ed., McGraw-Hill, New York.*)

15.16

m^2) of tower cross section, the upward air velocity is 300 ft/min (91.4 m/min), and the approach to the wet-bulb temperature is 10 to 15°F (5.6 to 8.3°C).

6. Drying. The unit operation *drying* refers to the removal by vaporization of liquid from a solid; the liquid usually constitutes a relatively small fraction of the wet solid. Beyond this, the term is applied to the removal of traces of vapor from a gas and of small amounts of water from another liquid. Only drying of a solid is discussed here, and water and air will be used as examples of wetting liquid and surrounding gas, respectively.

The equipment for drying is classified according to the means by which the necessary heat is brought to the evaporating liquid and also according to the form and disposition of the wet solid. *Direct dryers* deliver heat to the liquid by direct contact with the hot gas stream into which the liquid vaporizes; *indirect dryers* supply the heat through a wall that separates the wet solid from a heat source, such as condensing steam.

Typical of indirect dryers are drum dryers (for slurries), can dryers (for textiles), and cylinder dryers (for paper), all of which are rotating cylindrical vessels usually internally heated with steam. The material being dried is continuously applied to the outside surface, carried around for a part of a revolution, and then removed. Sufficient drying may be accomplished in one pass, as on a slow-turning drum dryer, or it may require repeated passes over a series of similar dryers, as in paper making. Generally, the drying rates can only be established experimentally; rates can range from the equivalent of 1000 to 4000 Btu/(h · ft^2) (3152 to 12,608 W/m^2) of dryer surface.

A wide variety of direct dryers are available, which differ primarily in the method of contacting the wet solid with the drying gas. The simplest and lowest cost is the batch tray drier in which the wet solid is spread on trays held in an enclosure and hot air is blown across the surface of the solid until it is sufficiently dry.

Rotary dryers offer continuous operation and are particularly adaptable to finely divided, nonsticking, and nonagglomerating solids. The wet solid is fed into the upper end of a slightly sloped, rotating cylinder through which heated air or hot combustion gas is passed. Internal flights lift the solid and shower it through the hot gas while also advancing it to the discharge end.

The steam-tube dryer is a variation of the rotary dryer in which a number of tubes are supported longitudinally within the rotating cylinder and supply the required heat. Only a small flow of air through the cylinder is needed to sweep out the vapor, and the loss of fines is minimized.

Spray dryers are particularly applicable when the feed is a solution or a very dilute suspension that can be atomized into fine droplets. The dryer consists of an atomizing device, a hot-gas source, and a drying chamber in which the droplets contact the hot gases. The atomizing device may be a pressure spray nozzle or a rotating disk.

7. Evaporation. *Evaporation* is the operation by which a volatile liquid is separated from a solution or suspension by vaporization. The separation is usually not complete, and either or both the vapor and concentrated liquor may be the desired product.

Direct-fired or steam-jacketed kettles may serve as evaporators, but the most common types use a tubular heating surface, and are steam heated. The tubes may

be horizontal or vertical, and the heating steam may be inside or surrounding the tubes. The tube bundle is placed within or external to the body of the evaporator and is so arranged that the boiling liquor can be circulated through or around the tubes. The body also serves to allow disengagement of the vapor from the liquor as it is formed. Figure 15.10 shows several types of tubular evaporators.

An evaporator is basically a heat-transfer device. Its evaporative capacity is related to the temperature difference between heating steam and the boiling solution, the area of the tubular heat-transfer surface, the heat-transfer coefficient and the latent heat of vaporization of the evaporating liquid:

$$\text{Capacity} = \frac{q}{\lambda} = \frac{(UA)(\Delta t)}{\lambda} \tag{15.31}$$

The coefficient U is chiefly a function of the velocity of the solution past the heating surface and its viscosity, the extent of fouling of the heating surface, the temperature difference between steam and liquor, and the height of the liquor level relative to the tubes. The fouling of the tubes can become significant or even the controlling factor when scale-forming materials are present in the liquor. The values of U are nearly always derived from experimentally determined rates of evaporation as a function of Δt. The temperature difference observed, however, is more often apparent than real because it is inferred from pressure measurements, which correspond to temperature values that do not reflect boiling-point rise of the liquid due to dissolved solute or temperature shifts on the steam side due to vapor superheat
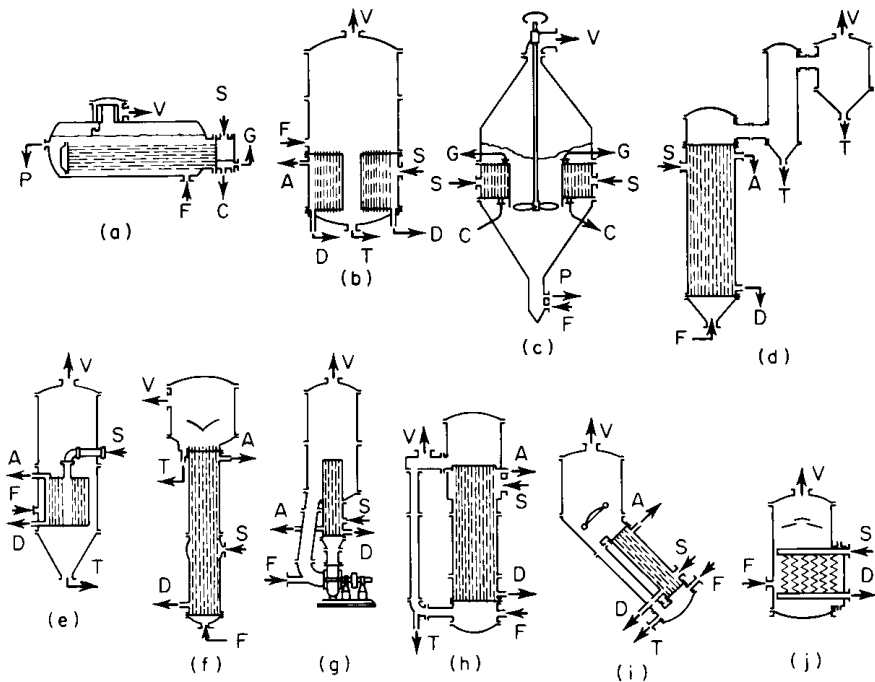
FIGURE 15.10   Typical evaporator designs: (*a*) horizontal tube, (*b*) short tube vertical, (*c*) propeller calandria, (*d*) long-tube vertical without vapor head, (*e*) basket type, (*f*) long-tube vertical, (*g*) forced circulation, (*h*) long-tube vertical with downtake, (*i*) buflovac inclined tube, (*j*) coiled tube.

or condensate subcooling. Such Δt values are known as *apparent temperature differences,* and the values of U corresponding to them are known as *apparent coefficients.* The error on the steam side usually is small. When the boiling-point rise is known and can be used to adjust the temperature difference, nearly correct values of Δt and U result; such values are known as *temperature difference* and *coefficient corrected for boiling-point rise.* When not otherwise stipulated, the values reported for evaporators usually are these.

Multiple-Effect Evaporation.   A *multiple-effect evaporator* is merely a series of similar evaporators so connected that the vapor from one body is the heating medium for the next. Passing from single to multiple effect does not alter the major features of body construction; it merely affects the interconnecting piping and the operation.

   The purpose of multiple-effect evaporation is to improve thermal efficiency. One pound (0.45 kg) of steam supplied to the first effect will evaporate approximately one pound (0.45 kg) of water in that effect. This pound (0.45 kg) of water vapor will then pass to the steam space of the second effect and, in condensing, will evaporate approximately another pound (0.45 kg) of water, and so on, so that in N effects, 1 lb (0.45 kg) of steam will evaporate approximately (but somewhat less than) N lb (kg) of water. The pressure must be progressively reduced from effect to effect in order to produce a temperature difference between the boiling liquid of that effect and the condensing vapor from the preceding effect.

Thermocompression.   The simplest though not the least expensive, means of reducing the thermal energy requirements of evaporation is to compress the vapor from a single-effect evaporator so that the vapor can be used as the heating medium in the same evaporator. The compression may be accomplished by mechanical means or by a steam jet. In order to keep the compressor cost and power requirements within reason, the evaporator must work with a fairly narrow temperature difference, usually from about 10 to 20°F (5.6 to 11.1°C). This means that a large evaporator heating surface is needed, partially offsetting the advantages of thermocompression.

8. Adsorption and Ion Exchange

Definitions.   *Adsorption* is separation of the components of a fluid phase brought about by contacting the fluid phase with a fixed solid phase. The solid phase, termed the *sorbent,* consists generally of highly porous particles with greatly extended surface area, most of which is internal. One or more components of the fluid phase, termed *sorbates,* enter the sorbent particles and bond to the surface of the solid. In *physical adsorption,* weak forces such as ionic or Van der Waals forces bond sorbate molecules to the surface. In *chemisorption,* true chemical bonds form between sorbate and sorbent surface. *Molecular sieves* possess a large number of adsorption sites of uniform type and essentially molecular dimensions and can be extremely selective toward a given component or class of components.

   Ion exchange is transfer of ions from a fluid phase to a solid phase with simultaneous transfer of other similarly charged ions in the reverse direction. The solid medium is a synthetic organic resin with functional groups having a given charge bonded to the resin matrix and hence immobile. Mobile ions of opposite charge are distributed throughout the matrix to maintain electrical neutrality. *Cation-exchange resins* have negatively charged functional groups (such as sulfonic groups)

bound to the matrix, while *anion-exchange resins* have positively charged bound groups (quaternary ammonium groups in *strongly basic resins* and other amine groups in *weakly basic resins*). If the mobile ion X is predominant, the resin is said to be in the *X form* (e.g., if Cl^- ions are the predominant mobile ion, the resin is in *chloride form*).

Once the sorbent has become saturated, the sorbate is stripped from the sorbent and recovered, returning the sorbent to its original condition for reuse. This is termed *regeneration* and may be accomplished by heating, chemical reaction, or contacting with a *regenerant.*

Design. Design for continuous operation generally depends on scale-up from an experimentally determined breakthrough curve for a particular sorbent-sorbate fluid system. It is generally found that for a given system the ratio of effluent concentration is a function of the quantity of liquid treated (expressed as volumes of liquid per volume of solid, dimensionless), residence time, and effective mass-transfer rate. For a given system at a given set of equilibrium conditions, the same functional relationship must hold:

$$X = f \left(\frac{F^a V}{D_p^b F'} \right) \left(\frac{v}{V} \right) \tag{15.32}$$

where F^a/D_p^b is a dimensional group controlling mass-transfer rate, V/F' is the residence time, and (v/V) is the total throughput in bed volumes. If mass transfer is rapid enough to ensure instantaneous equilibrium at all points, the effects of the mass-transfer parameter and residence time are not significant, and

$$X = f \left(\frac{v}{V} \right) \tag{15.33}$$

The shape of this curve is dependent on equilibrium considerations only (i.e., the feed concentration and the adsorption isotherm).

9. Membrane Processes. Several diffusional separation processes involve transfer of one or more components from one fluid phase through a porous solid medium or membrane to a second fluid phase. The porous medium, generally very thin, does not permit flow of the fluid but does allow the transfer of material between phases by molecular diffusion through pores. Membrane processes are classified according to the driving force across the membrane causing this diffusion; in *membrane permeation,* the driving force is a difference in partial pressure; in *dialysis,* the driving force is a concentration difference; in *ultrafiltration* and *reverse osmosis,* the driving force is a difference in pressure; and in *electrodialysis,* the driving force is a voltage gradient.

Membrane Permeation. For steady-state processes the rate of permeation of a component through the membrane is given by

$$N_M = \frac{P}{L} (\Delta p) \tag{15.34}$$

where Δp is the difference in partial pressure of that component across the

membrane, and P is the *permeability*. Permeability is a function of the membrane composition and its structure, diffusing component, and system temperature and pressure. Correlations are available to predict permeabilities for various permeate-membrane combinations and the effects of temperature and pressure.[3]

Dialysis and Osmosis. If two liquid solutions with different compositions are separated by a permeable membrane, the concentration gradient causes molecular diffusion through the membrane. For relatively dilute solutions, where no significant volume change occurs as the diffusion proceeds, mass-transfer rates are described by

$$N_{M_i} = U_i A \Delta C_{i,lm} \tag{15.35}$$

The overall dialysis coefficient U_i is determined by three film coefficients, analogous to heat transfer

$$\frac{1}{U_i} = \frac{1}{U_{iL1}} + \frac{1}{U_{im}} + \frac{1}{U_{iL2}} \tag{15.36}$$

where the individual coefficients are for the membrane and the two liquid films. If adequate mixing takes place in the two liquids, the membrane coefficient is controlling, $U_i = U_{im}$.

Reversed Osmosis and Ultrafiltration. Reverse osmosis and ultrafiltration processes recover the pure solvent from a solution. If a liquid solution is separated from the pure solvent by a membrane permeable to the solvent only, the osmotic pressure causes the solvent to diffuse from the pure liquid side to the solution side. A mechanical pressure applied to the solution side slows the diffusion process. If sufficient pressure is applied, the flow is reversed; i.e., solvent is transferred from the solution to the pure side, counter to the concentration gradient.

For dilute solutions, large solute molecules or high pressure differentials, Eq. (15.34) applies for reverse osmosis processes. For concentrated solutions of solutes with low molecular weight, a correction for osmotic pressure is required.

In most applications, the limiting factor is concentration polarization of the membrane. As solvent is removed from the solution, solutes greatly increase in concentration in a thin layer near the membrane. This causes an osmotic pressure that opposes the applied pressure gradient and reduces the rate of transfer. Another significant problem is buildup of deposits on the membrane surface.

Normal design tradeoffs involve membrane thickness, pressure differential, membrane surface area, and percentage of solvent recovered.

Electrodialysis. If an electric field is imposed on an electrolytic solution, negatively charged ions are attracted toward the positive pole of the field and positively charged ions are attracted to the negative pole. If the solution is partitioned by semipermeable membranes (alternately anion-permeable and cation-permeable), the flow of ions under the influence of the electric field will create alternately dilute and concentrated compartments. The feed solution is fed to the diluting compartments and a salt-recovery solution is fed to the concentrating compartments. Electrodialysis may be carried out on either a continuous or a batch basis.

MULTIPHASE CONTACTING AND PHASE DISTRIBUTION

10. Agitation. *Agitation* is motion imparted to material to promote heat or mass transfer to, from, or within the material or to distribute another phase through the material. Examples include mixing of miscible liquids, formation of uniform solids suspensions, promotion of heat transfer between a process fluid and a heat-exchange surface, promoting dissolution of a gas in a liquid. The most important process applications of agitation involve a freely fluid liquid as the primary phase. The key equipment is a rotating element (sometimes more than one) called the *agitator* or *impeller.* Turbines, paddles, propellers, and special shapes are used, the choice depending on the properties of the agitated material and the type of agitation desired.

Agitation Power. Over a wide range of operating conditions, the parameter that most completely determines the performance of an agitator is the power delivered to the fluid—the greater the power delivered, the more effective is the agitation. The power required by a rotating impeller cannot be computed directly; rather it must be measured for a geometrically similar model of the impeller and its surroundings and then be scaled up or down. The most useful scaling correlations involve the impeller Reynolds number $D_a^2 N \rho / \mu$, the power number $Pg_c / D_a^5 \rho N^3$, and the Froude number $g / N^2 D_a$. Examples of this correlation are presented in Fig. 15.11. Since agitator geometry (including impeller dimensions, impeller position, vessel dimensions, baffles etc.) affects the relationship between Reynolds number
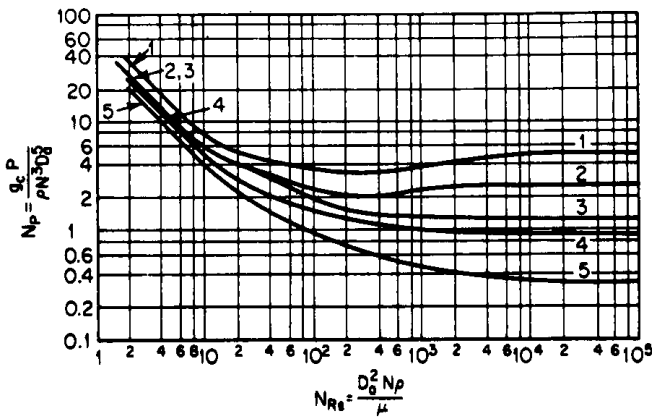
FIGURE 15.11 Impeller power correlations. (1) Six-blade turbine (curved blades), $D_a / W_i = 5$, with four baffles of height equal to $\frac{1}{12}$ agitator tank diameter. (2) Six-blade turbine (straight vertical blades), $D_a / W_i = 8$; same baffle arrangement as (1). (3) Six-blade turbine (straight blades with 45° itch), $D_a / W_t = 8$; same baffle arrangement as (1). (4) Propeller, pitch equal to $2D_a$; four baffler of height equal to $\frac{1}{10}$ agitator tank diameter. (5) Propeller (pitch equal to D_a); same baffle arrangement as (4). (*Curves 4 and 5 from Rushton, Costich, and Everett, Chemical Engineering Progress 46: 395, 467, 1950, by permission; curves 2 and 3 from Bates, Fondy, and Corpstein, Industrial Engineering and Chemical Process Design Development 2: 310, 1963, by permission of the copyright owner, the American Chemical Society.*)

and power number, it is preferable to use experimental data rather than published correlations.

Functional Performance of Agitators.   The degree of effectiveness of agitation is related to the intensity of shear, the level of turbulence, and the circulation rate produced by the impeller. For an impeller of given design, these quantities are determined by the agitator speed and power delivery. The following equations relate velocity head, power, and circulation rate for geometrically similar impellers under turbulent flow conditions:

$$Q = N_Q N D_a^3 \tag{15.37}$$

$$H = \frac{N_P N^2 D_a^2}{N_G Q_c} \tag{15.38}$$

$$P = \rho Q H = N_P \rho N^3 \frac{D_a^5}{g_c} \tag{15.39}$$

where Q is impeller discharge rate, ft^3/s (m^3/s); H is velocity head, ft · lb$_f$/lb$_m$ (N · m/N); N_P is power number; and N_Q is discharge coefficient, dimensionless.

11. Spray Generation.   A *spray* is a mechanically produced unstable suspension of liquid droplets in a gas. Sprays are generated from a continuous liquid by nozzles, of which there are three principal types: *pressure nozzles, rotating nozzles,* and *gas-atomizing nozzles* (or *two-fluid nozzles*).

Pressure Nozzles.   In pressure nozzles, classified as *hollow-cone, solid-cone, fan,* or *impact* types depending on the shape of the spray, the fluid is throttled through small openings, thereby attaining high velocities. Droplets are formed because of shear forces at the nozzle exit, by its instability at such high levels of inertia, or by impact with another jet or a solid surface.

Gas-Atomizing Nozzles.   These nozzles disintegrate a stream of liquid by contact with a high-velocity stream of gas. The liquid may be preatomized by a pressure nozzle or injected as a generally continuous sheet.

12. Gas Sparging.   A *sparger* is a distributor that disperses gas into the body of a liquid by emitting bubbles or jets of gas through an individual orifice, an array of orifices, or a porous structure.

Simple Bubblers.   Open-ended pipes or perforated tubes or plates with orifices ⅓ to ½ in (3.2 to 12.7 mm) in diameter are used as spargers. A perforated tube should be designed so that the pressure drop across the individual orifice is large compared to the pressure drop for flow through the tube.

Porous Septa.   Porous plates, tubes, or disks are made by bonding or sintering carefully sized particles or carbon, ceramic, metal, or polymer. The resulting septa may be used as spargers to produce much smaller bubbles than will result from a simple bubbler.

13. Fluidization

Definitions.   If a gas is passed upward through an unrestrained and unconsolidated bed of granular solids with ever-increasing velocity, the pressure drop across the bed due to friction will increase until it becomes equivalent to the weight of the bed plus the friction between fluid and walls. With further increase in the gas velocity, the bed tends to rise as a unit, but its unconsolidated character causes it instead to expand until the increased porosity allows the friction again just to balance the pressure drop. As the bed becomes more expanded, individual particles achieve freedom to interchange position, and the bed can circulate.

Continuous Fluidization.   One primary application for continuous fluidization is *pneumatic conveying* of a dispersed solid by means of compressed air. The solid phase is introduced into the carrier gas (e.g., by aspiration or by a screw conveyor), is immediately fluidized, and is transported to the destination where it is removed from the carrier gas by means of a cyclone or other collection device.

Batch Fluidization.   Fluidized beds sometimes are called *boiling beds.* Indeed, the expanded suspended mass of the bed does resemble a boiling liquid. This mass has a zero angle of repose, seeks its own level, and assumes the shape of the containing vessel.

Conditions for Fluidization.   The size of solid particles that can be fluidized varies greatly, from less than 1 μm to 2½ in (6.4 cm). It is generally concluded that particles distributed in size between 65 mesh and 10 μm are the best for smooth fluidization (least formation of large bubbles).

The upward velocity of the gas is usually between 0.5 and 10 ft/s (0.15 and 3.0 m/s). This velocity is based on the flow through the empty vessel and is frequently referred to as the *superficial velocity.*

Bed height is determined by a number of factors, either individually or collectively, such as: (1) space-time yield, (2) gas-contact time, (3) *L/D* ratio required to provide staging, (4) space required for internal heat exchangers, and (5) solids-retention time. Generally, bed heights are not less than 12 in (30.5 cm) or more than 50 ft (15.2 m).

Heat Transfer and Mixing in Fluidized Beds.   Heat-exchange surfaces are used to provide means of removing or adding heat to fluidized beds. Usually, these surfaces are provided in the form of vertical tubes manifolded at top and bottom. Other shapes have been used such as horizontal bayonets. In any such installations, adequately provision must be made for abrasion of the exchanger surface by the bed. Normally, the transfer rate is 5 to 25 times that for solids-free gas.

Heat transfer from solids to gas and gas to solids usually results in a coefficient of about 3 to 10 Btu/(h · ft^2 · °F) [17.0 to 56.8 W/(h · m^2 · C)]. However, the large area of the solids per cubic foot of bed [15,000 ft^2/ft^3 (49,240 m^2/m^3)] for 60-μm particles of 40 lb/ft^3 (640 kg/m^3) bulk density] results in the rapid approach of gas and solids temperatures. With a fairly good distributor, essential equalization of temperatures occurs within 1 to 3 in (2.5 to 7.6 cm) of the top of the distributor.

Bed thermal conductivities in the vertical direction have been measured in the laboratory in the range of 20,000 to 30,000 Btu/(h · ft · °F/ft) [373 to 559 kW/ (m · h · °C/m)]. Horizontal conductivities for ⅛-in (3.2-mm) particles in the range

of 1000 Btu/(h · ft² · °F/ft) [18.6 kW/(h · m² · °C/m)] have been measured in large-scale experiments.

MECHANICAL SEPARATIONS AND PHASE COLLECTION

14. Filtration.   *Filtration* is the mechanical phase separation of a fluid-solid suspension or slurry by passage of the liquid through a porous septum, or filter medium, which retains the solids. The clarified liquid product is called the *filtrate,* and the retained solids are called the *filter cake.*

Types of Filters.   Gravity filters are the simplest and are always intermittent in operation. They are usually in the form of a false-bottomed tank; the false bottom is perforated and supports the filter medium.

The *plate-and-frame filter press* is an intermittent-pressure filter. The plates are solid with channeled or ribbed faces; the filter medium (e.g., filter cloth) is laid over the faces of each plate. The frames are hollow and of the same outer dimensions as the plates but are made in a variety of thicknesses.

Leaf filters are also intermittent-pressure filters. The leaf is a hollow frame covered with a wire support screen and the filter medium.

There are several *continuous-vacuum filter* designs, but all are similar in basic principle. The filter medium is carried on a rotating drum or disk mounted on a hollow horizontal shaft. Vacuum is applied to the plenum behind the medium and the cake is formed during that part of the revolution when the medium is immersed in the slurry.

Rate of Filtration.   The rate of filtration is the rate at which the filtrate can be forced through the cake and filter medium. The flow resistance of the cake is a function of its particle-size distribution, compressibility, thickness, and the viscosity of the filtrate. The volumetric flow rate through a hard, granular, noncompressible cake can be related to the operating conditions and cake properties:

$$\frac{dV}{d\theta} = \frac{A\Delta P_c}{\mu\alpha g_c(W/A)} \tag{15.40}$$

where α is the specific cake resistance (constant). For compressible cakes, α is a function of pressure drop, often correlated by

$$\alpha = \alpha'\Delta P_c^s \tag{15.41}$$

where α' and S are empirical constants depending on the particular solids being filtered. The compressibility S varies from zero for a hard, granular, noncompressible cake to unity for a soft, readily deformed, compressible cake; for most industrial slurry solids, S ranges from 0.1 to 0.8. For compressible cakes, the filtration rate is given by

$$\frac{dV}{d\theta} = \frac{A\Delta P_c^{(1-s)}}{\mu\alpha' g_c(W/A)} \tag{15.42}$$

The sizing of a filter for a particular slurry starts with laboratory or pilot-plant-

scale tests followed by scale-up to the desired capacity. The choice of the type of filter involves a variety of factors and is best made with the advice of filter man-ufacturers.

15. Settling and Sedimentation.   The separation of suspended solids from a fluid by gravitational forces is termed *settling* or *sedimentation;* differentiation be-tween the two terms is not precise. *Settling* usually refers to very dilute suspensions of discrete rigid particles or discrete liquid droplets. *Sedimentation* is applied to higher concentrations of solids and to solids that *flocculate* (i.e., agglomerate to form rigid lattices).

In the free or unhindered settling of discrete particles, the particle will accelerate until drag forces offered by the liquid exactly balance the net gravitational forces (i.e., particle weight minus buoyant forces) and thereafter will settle at a constant *terminal velocity* expressed by

$$v_t = \sqrt{\frac{4gD_p(\rho_p - \rho_f)}{3\rho_f C_D}} \tag{15.43}$$

where the drag coefficient C_D is a function of particle shape and the Reynolds number based on particle diameter, $N_{Re} = D_p v_t \rho_f / \mu$. At very low N_{Re} (<0.3), Stokes' law is valid for spherical or near-spherical rigid particles, and $C_D = 24/N_{Re}$; at higher N_{Re}, the coefficient decreases with increasing N_{Re}, but becomes less dependent on N_{Re}, becoming essentially constant; as the sphericity of a particle decreases, the drag coefficient increases. Nonrigid particles exhibit similar behavior, but droplet deformation and internal circulation affect the relationship between C_D and N_{Re}.

16. Centrifugation.   Centrifugal force can be applied to enhance the separation of liquid-solid or liquid-liquid suspensions by either settling or filtration. Centrifugal force is exerted on a mass that is following a curved path; the extent of the force depends on the radius of curvature and the angular velocity, increasing as velocity is increased or curvature is decreased. Thus a centrifugal force field thousands of times stronger than normal gravity can be generated in properly designed devices called *centrifuges.* A centrifugal force field can produce the same effects as a grav-itational field but at rates proportional to the relative strengths of the fields.

Centrifuges may be classified by the operation carried out, namely, sedimentation or filtration. *Sedimentation centrifuges* are usually continuous-operation, solid-bowl machines yielding a dense phase of thickened sludge and a light phase of clear liquid.

Filtration centrifuges usually have a perforated-wall cylindrical basket lined with a suitable filter medium, such as fine woven cloth or wire screen, which will retain the solids and allow the liquid to pass through. The centrifugal force acts directly on the liquid to push it through a cake of solids during buildup of the cake and also to drive it out of the pores of the cake during the dry spin.

17. Screening.   *Screening* is the mechanical separation of a mixture of particles into two or more fractions by means of a surface with multiple uniform perforations, termed a *screen*. Material retained on the first screen in a series (i.e., largest open-ings) is termed *oversize;* that passing through the last in a series (i.e., smallest

openings) is termed *undersize,* or *fines.* A screen may consist of cloth woven from various fibers, a perforated plate, or uniformly spaced parallel bars. Screens are specified by the number of openings per linear inch (*mesh count*) or by the dimension of the openings, measured between and perpendicular to adjacent wires or bars (*aperture* or *clear opening*).

18. Wet Classification. *Wet classification* is the separation of a mixture of particles into two or more fractions according to particle size or particle density by contact with a fluidizing medium, often water. It is used as a unit operation in the chemical process industry primarily for raw materials treatment (i.e., ore beneficiation, coal washing, etc.). Most types of wet classification utilize the different settling velocities of the different particles to remove coarse or dense particles (termed *sand*) from fine particles.

Jigging. Jigging is the separation of materials of different specific gravities by the pulsation of a stream of liquid flowing through a bed of the materials. The liquid pulsates, or ''jigs,'' up and down, causing the heavy material to work down to the bottom of the bed and the lighter material to rise to the top. Each product is then drawn off separately.

Tabling. *Tabling* is the classification of particulate solids by means of an inclined, riffled, shaking surface (called a *table*) across which water or air is flowed. The particles are classified principally on the basis of density difference.

Wet tables require finer feed (dense ore, 6 to 150 mesh; light material, such as coal, <1 in [25.4 mm]) than air tables [which handle ore up to $\frac{1}{4}$ in (6.4 mm) and coal up to 3 in (76.2 mm)].

Froth Flotation. *Froth flotation* is the fractionation of particulate solids based on differences in interfacial tensions between the solids, water, and air. It has been an important process in the beneficiation of ore. The ore particles are suspended in a liquid at a pulp density of 15 to 35 percent solids by mechanical or air agitation. The slurry is treated with chemicals, called *promoters,* which render the surfaces of specific minerals air-avid and water-repellent.

19. Crystallization. *Crystallization* is the production and recovery of solid material from a solution brought about by reduction of solubility due to temperature change or by evaporation of solvent. A saturated solution is fed to a vessel, often an agitated vessel, and heat is added or removed so that the concentration in the solution is above the solubility at the operating conditions, i.e., supersaturated.

CHEMICAL KINETICS AND REACTOR DESIGN

20. Introduction. The principles governing the mechanisms of chemical reactions and the rates at which they proceed comprise the field of *chemical kinetics.* Although the theory of chemical kinetics is imperfect, it is a useful guide for analyzing the results of experimental investigations and is thus the basis for design of practical industrial reactor systems.

The rate of a chemical reaction is best described in terms of number of moles of reactant converted (or moles of product produced) per unit of reactor volume per unit time. Thus for the reaction with the stoichiometric equation

$$aA + bB + cC \rightleftharpoons pP + qQ \tag{15.44}$$

the rate of reaction may be described by

$$rV = \frac{-1}{a}\frac{dN_A}{d\theta} = \frac{-1}{b}\frac{dN_B}{d\theta} = \frac{-1}{c}\frac{dN_C}{d\theta} = \frac{1}{p}\frac{dN_P}{d\theta} = \frac{1}{q}\frac{dN_Q}{d\theta} \tag{15.45}$$

The rate of appearance (or disappearance) of any product (or reactant) can be obtained in terms of the rate of change of any other participant by multiplication of Eq. 45 by the appropriate stoichiometric coefficient. If the reaction volume is constant, the concentration of reactants may be introduced:

$$r = \frac{1}{a}\frac{d(N_A/V)}{d\theta} = -\frac{1}{a}\frac{dC_A}{d\theta} \tag{15.46}$$

This is approximately true for most liquid-phase reactions taking place in a tank where material is neither added nor removed (*batch reaction*) or where no significant change in liquid level occurs in a flow reaction. It is exact for gas-phase reactions (flow or batch) confined in a rigid vessel.

For steady-state flow reactions with no longitudinal mixing, the composition at any point is constant with time, and the reaction rate may be defined in terms of change in composition with position as

$$r = \frac{F}{a}\frac{dX_A}{dV} \tag{15.47}$$

where F is volumetric feed rate, X_A is moles of A converted per unit volume of total feed, and V is the reactor volume upstream of that point in the reactor.

21. Homogeneous Reactions.   Reactions in which both reactants and products are in the same phase throughout the reaction are termed *homogeneous reactions.* Some reactions involving phase change may be treated as homogeneous even if there is a phase change between the initial reactant state and final product state provided the phase change occurs rapidly enough so as not to affect the overall reaction rate. (A reaction which results in a precipitated product, for example, may be treated as a homogeneous reaction during which the product concentration is constant at its solubility level.)

Reaction Order.   In general, the rates of reactions whose mechanisms are simple have been found to be proportional to integral powers of the concentrations of some or all of the reacting components. That is, for a reaction involving reactants A, B, and C,

$$r = kC_A^\alpha C_B^\beta C_C^\gamma \tag{15.48}$$

The exponents are experimentally determined and are not necessarily equal to the stoichiometric coefficients of the reaction equation. In general, they have a value

between 0 and 3. For the special case of single-step reactions involving a small number of reacting molecules, the exponents α, β, and γ will be integers.

Rate laws for more complex reactions having several steps, involving catalyzed reactions, chain reactions, etc., will be much more complicated than Eq. (15.48). Such reactions may sometimes be described by Eq. (15.48) in order to obtain a correlation between rates and reactant concentrations. If this is done, the exponents α, β, and γ will often have noninteger values. When the reaction mechanism is unknown, this is often the only available procedure. It should be used with caution, however, and usually the values of α, β, and γ which are experimentally determined apply only to the conditions (reactant concentration, catalyst concentration, temperature, etc.) for which the data were taken.

The reaction kinetically described by Eq. (15.48) is said to be αth order with respect to A, βth order with respect to B, and γth order with respect to C; as a whole, its order is $(\alpha + \beta + \gamma)$th. Theory suggests that order may be related to molecularity of the reaction mechanism; if so, the order of simple homogeneous reactions with respect to each component should be finite and represented by an integer.

Equation (15.48) gives the rate of a reaction proceeding irreversibly among three components. If the reaction of interest were, instead, a reversible one (as, strictly speaking, all reactions are), such as

$$A + B + C \rightleftharpoons P + Q \tag{15.49}$$

Eq. (15.49) would describe only the rate of the forward half-reaction. The reverse might be expected to exhibit a rate proportional to simple powers of the concentrations of the products P and Q:

$$r' = k'C_P^\rho C_Q^\sigma \tag{15.50}$$

Thus the reverse half-reaction would be ρth order with respect to P, σth order with respect to Q, and $(\rho + \sigma)$th order overall.

It should be noted that the net rate of a reversible reaction is the algebraic sum of its forward and reverse rates. For the reaction described by Eqs. (15.48) and (15.50),

$$r_{\text{net}} = r - r' \tag{15.51}$$

The coefficients k and k' of Eqs. (15.49) and (15.50) are known as specific rate constants, peculiar to a particular reaction and temperature but independent of concentrations of reactants.

Integrated Rate Equations.   If one substitutes the appropriate rate law [e.g., Eq. (15.48)] into Eq. (15.46) and relates the concentrations by means of a material balance, it is possible to integrate Eq. (15.46) to give reactant concentrations as a function of time. For irreversible reactions of integral order, these integrated equations have a simple form. Table 15.1 presents integrated equations for simple irreversible, constant-volume reactions.

Reaction rate constants [k in Eq. (15.49)] may be determined from experimental data by plotting concentration versus time in a manner determined by the form of the integrated rate law [e.g., for a first-order reaction plot $\ln(C_{A0}/C_A)$ versus time; the reaction rate constant is the slope of the line].

Reversible reactions, consecutive reactions ($A \rightarrow B \rightarrow D$), and parallel reactions

TABLE 15.1 Rate Equations for Reactions of Simple Order

Order	Differential equation	Constant-volume process
Zero	$-\dfrac{dN_A}{Vd\theta} = k$	$k(\theta - \theta_0) = C_A^0 - C_A$
One-half	$-\dfrac{dN_A}{Vd\theta} = kC_a^{1/2}$	$k(\theta - \theta_0) = 2(C_A^{01/2} - C_A^{1/2})$
First	$-\dfrac{dN_A}{Vd\theta} = kC_A$	$k(\theta - \theta_0) = \ln\dfrac{C_A^0}{C_A}$
Second	$-\dfrac{dN_A}{Vd\theta} = kC_A^2$	$k(\theta - \theta_0) = \dfrac{1}{C_A} - \dfrac{1}{C_Z^0}$
	$-\dfrac{dN_A}{Vd\theta} = kC_A C_B$	$k(\theta - \theta_0) = \dfrac{1}{C_B^0 - C_A^0}\ln\dfrac{C_A C_A^0 + C_A^0 C_B^0 - C_A^{03}}{C_A C_B^0} \qquad C_A^0 \neq C_B^{0*}$
Third	$-\dfrac{dN_A}{Vd\theta} = kC_A^2$	$2k(\theta - \theta_0) = \dfrac{1}{C_A^2} - \dfrac{1}{C_A^{02}}$
	$-\dfrac{dN_A}{Vd\theta} = kC_A C_B C_c$	$k(\theta - \theta_0) = \dfrac{1}{(C_B^0 - C_A^0)(C_C^0 - C_A^0)}\ln\dfrac{C_A^0}{C_A} + \dfrac{1}{(C_B^0 - C_C^0)(C_B^0 - C_A^0)}\ln\left(\dfrac{C_B^0}{C_A + C_B^0 - C_A^0}\right)$ $+ \dfrac{1}{(C_C^0 - C_B^0)(C_C^0 - C_A^0)}\ln\left(\dfrac{C_C^0}{C_A + C_C^0 - C_A^0}\right)C_B^0 \neq C_C^0 \neq C_A^{0\dagger}$

NOTE: C^0 and θ_0 are initial conditions for time and concentration, respectively.
*If $C_A^0 = C_B^0$, use expression for $-dN_A/Vd\theta = kC_A^2$.
†If $C_A^0 = C_B^0 = C_C^0$, use expression for $-dN_A/Vd\theta = kC_A^3$.

15.30

$$\left(A \lessgtr {B \atop C}\right)$$
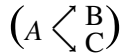
require much more complicated rate equations for their description.

Equilibrium and Kinetics.   Inasmuch as all chemical reactions are limited by a chemical equilibrium, reaction kinetics really describes the rate of approach to that equilibrium rather than to a stoichiometric completeness of the reaction. At equilibrium, the net rate of reaction Eq. (15.51) is zero, whence it follows that for a reaction whose rate law is described by the molecularity (i.e., stoichiometric equation), the equilibrium constant for the reaction is related to the forward and reverse specific rate constants. Thus,

$$\frac{C_A^a C_B^l C_C^c}{C_P^p C_Q^q} = K_c = \frac{k}{k'} \tag{15.52}$$

It is clear that the larger the value of K_c, the larger is the magnitude of the forward rate constant relative to the reverse and the closer to stoichiometric completeness is the equilibrium conversion.

Effect of Temperature.   Homogeneous reactions are strongly temperature-dependent. Their specific reaction rate always increases with increasing temperature. The effect of temperature is described by the semitheoretical relation of Arrhenius:

$$k = Ae^{-EIRT} \tag{15.53}$$

The coefficient A (called the *frequency factor*) and the exponent E (called the *energy of activation*) have theoretical interpretations, but they are best regarded by the process designer as empirical quantities peculiar to a particular chemical reaction and valuable from experimental rate data. Thus a plot of k against $1/T$ should be linear and should have a slope of $-E/R$ and an intercept of $\ln A$, provided A and E are independent of temperature. In fact, both the frequency factor and the energy of activation vary slightly with temperature, but over the temperature ranges normally encountered, they may be assigned constant average values without serious error.

The failure of rate data to fit Eq. (15.53) may be accepted as evidence that

1. A reversible reaction has been treated as if it were irreversible, and the effect of temperature on the equilibrium is significant.
2. An otherwise incorrect mechanism has been assigned to the reaction.
3. The specific rate constant has been evaluated from the experimental data incorrectly.
4. The reaction is heterogeneous, and its rate is influenced by adsorption or by some other physical process.

Energies of activation range from less than 1000 to greater than 100,000 cal/(g · mol). For most reactions, the value will be between 10,000 and 70,000 cal/(g · mol).

Effect of Concentration. At constant temperature, the specific rate constant is assumed to be independent of the concentration of reactants and products, so that equations like (15.48) and (15.50) show explicitly the effect of concentration on the progress of the reaction.

Homogeneous Catalyzed Reactions. A *catalyst* is a substance that affects the rate of a chemical reaction without entering the reaction in any stoichiometric sense. The catalyst may undergo net physical or chemical change in the course of the reaction, but often it does neither.

22. Heterogeneous Reactions.

A chemical reaction is said to be *heterogeneous* if more than one phase is an active participant and if transfer of materials to phase boundaries has an effect on the rate of reaction. Heterogeneous reactions commonly involve fluid-solid mass transfer (e.g., catalysis of a fluid reaction mixture on the surface of solid catalyst pellets, combustion of a solid fuel in air, acid leaching of metals from ores) or mass transfer between two fluid phases (e.g., absorption of gaseous sulfur dioxide by weak aqueous sodium hydroxide, nitration of toluene by nitric acid).

Uncatalyzed Heterogeneous Reactions. In an uncatalyzed heterogeneous reaction, chemical action occurs among components that are simultaneously being transferred physically from phase to phase. The apparent rate of the reaction is in fact the rate of a more complicated process. It will be influenced not only by factors affecting chemical kinetics, but also by those affecting the rate of interphase mass transfer.

Catalyzed Heterogeneous Reactions. Although heterogeneous reactions responsive to catalysis may involve any combination of phases, the examples most common and industrially most important are solid-fluid systems in which the catalyst is the solid phase. The reactants and products may be gaseous, liquid, or both. The solid catalyst may be a container wall, a metal gauze or a granular mass.

Figure 15.12 is a typical plot of effectiveness factor against a modulus
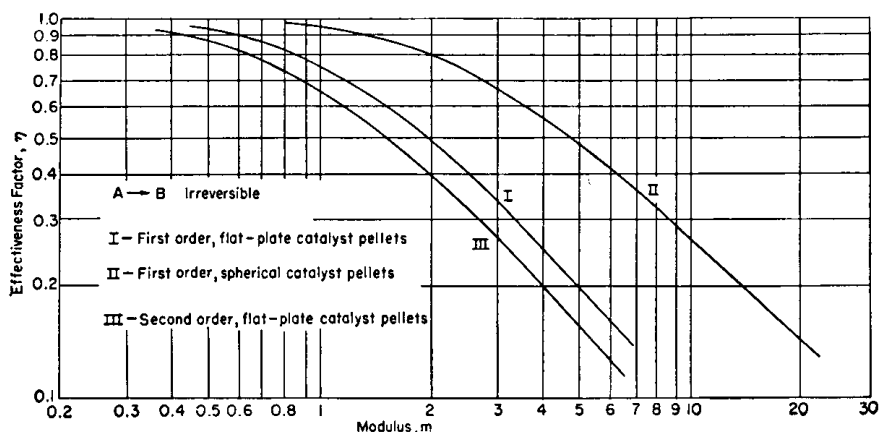
FIGURE 15.12 Effectiveness factor for equations of simple order.

$$m = 1 \sqrt{2k(C_i^0)^{n-1}(\bar{r}D_i)} \tag{15.54}$$

in which C^0 is the concentration of reactant at the external catalyst surface, and $\bar{r}$, the average pore radius, is calculated as $2V_g/S_g$.

3. If absorption-desorption or chemical reaction at the catalyst surface is the controlling process, a mechanism must be found that will identify which of the possibilities are rate-controlling and which are equilibrium steps. For example, for the stoichiometric reaction

$$A \rightleftharpoons R \tag{15.55}$$

the following mechanistic steps involving the participants A and R and a catalyst site s may be postulated:

$$A + s \rightleftharpoons As \tag{15.55a}$$

$$As \rightleftharpoons Rs \tag{15.55b}$$

$$Rs \rightleftharpoons R + s \tag{15.55c}$$

Assumption of each of these in order as the controlling step results in a different rate equation, which may be validated against experimental kinetic data. If none meets the test, a new mechanism must be tried. Rate equations for a number of simple examples are summarized in Table 15.2.

23. Interpretation of Kinetic Data.   The design of a chemical reactor should be based on properly collected and analyzed laboratory data. These data will, in general, be one of three types:

1. Measurements of composition as a function of time in a batch reactor of constant volume operated at various temperatures and pressures
2. Measurements of outlet composition as a function of feed rate to a flow reactor of constant volume operated at various pressure and temperature levels
3. Measurements of composition as a function of time in a variable-volume batch reactor operated at constant temperature and substantially constant pressure

For simple reactions involving more than one reactant, the order with respect to any one component can be determined by holding all other reactant concentrations constant. This may be accomplished by using large excesses of these other reactants. For the specific reaction $A + B \rightarrow C$, with reactant B present in great excess, assumed to be substantially irreversible, several analysis techniques are presented below as examples of the general approach.

1. *Differential rate equation:* A logarithmic plot of $dC_A/d\theta$ (or, by approximation, $\Delta C_A/\Delta\theta$) against C_A yields a straight line, of which the slope is n, the order of the reaction (Fig. 15.13).

2. *Integrated rate equation, first order:* A plot of $-\log(C_A/C_{AO})$ against θ yields a straight line through the origin (Fig. 15.14).

3. *Integrated rate equation, other than first order:* When $n > 1$, the general rate equation

TABLE 15.2 Mechanisms and Their Corresponding Rate Equations

Chemical equation	Catalytic steps	Rate equation*
$A \leftrightharpoons R$	$A + s \leftrightharpoons As$	$r = \dfrac{k(C_A - C_R/K)}{1 + K_R C_R}$
	$As \leftrightharpoons Rs$	$r = \dfrac{k(C_A - C_R/K)}{1 + K_A C_A + K_R C_R}$
	$Rs \leftrightharpoons R + s$	$r = \dfrac{k(C_A - C_R/K)}{1 + K_A C_A}$
$A \leftrightharpoons R$	$2A + s \leftrightharpoons Ast$	$r = \dfrac{k(C_A^2 - C_R^2/K^2)}{1 + K_R C_R + K_B C_B^2}$
	$A_{st} + s \leftrightharpoons 2As$	$r = \dfrac{k(C_A^2 - C_R^2/K^2)}{(1 + K_R C_R + K_A C_A^2)^2}$
	$As \leftrightharpoons Rs$	$r = \dfrac{k(C_A - C_R/K)}{1 + K_A C_A^2 + K_A' C_A + K_R C_R}$
	$Rs \leftrightharpoons R + s$	$r = \dfrac{k(C_A - C_R/K)}{1 + K_A C_A^2 + K_A' C_A}$
$A \leftrightharpoons R$	$A + 2s \leftrightharpoons 2A_{1/2}s$	$r = \dfrac{k(C_A - C_R/K)}{(1 + \sqrt{K_R C_R} + K_R' C_R)^2}$
	$2A_{1/2}s \leftrightharpoons Rs + s$	$r = \dfrac{k(C_A - C_R/K)}{(1 + \sqrt{K_A C_A} + K_R C_R)^2}$
	$Rs \leftrightharpoons R + S$	$r = \dfrac{k(C_A - C_R/K)}{1 + \sqrt{K_A C_A} + K_A' C_A}$
$A \leftrightharpoons R + S$	$A + s \leftrightharpoons As$	$r = \dfrac{k(C_A - C_R C_S/K)}{1 + K_{RS} C_R C_S + K_R C_R + K_S C_S}$
	$As + s \leftrightharpoons Rs + Ss$	$r = \dfrac{k(C_A - C_R C_S/K)}{(1 + K_A C_A + K_R C_R + K_S C_S)^2}$
	$\left.\begin{array}{l} Rs \leftrightharpoons R + s \\ Ss \leftrightharpoons S + s \end{array}\right\}$	$r = \dfrac{k(C_A - C_R C_S/K)}{C_S(1 + K_A C_A + (K_{AS} C_A/C_s) + K_S C_S)}$
$A \leftrightharpoons R + S$	$A + s \leftrightharpoons As$	$r = \dfrac{k(C_A - C_R C_S/K)}{1 + K_R C_R + K_{RS} C_R C_S}$
	$As \leftrightharpoons Rs + s$	$r = \dfrac{k(C_A - C_R C_S/K)}{1 + K_A C_A + K_R C_R}$
	$Rs \leftrightharpoons R + s$	$r = \dfrac{k(C_A - C_R C_S/K)}{C_S(1 + K_A C_A + K_{AS} C_A/C_S)}$

TABLE 15.2 Mechanisms and Their Corresponding Rate Equations (*Continued*)

Chemical equation	Catalytic steps	Rate equation*
$A + B \rightleftharpoons R$	$A + s \rightleftharpoons As$	$r = \dfrac{k(C_A - C_R/KC_B)}{1 + (K_{RB}C_R/C_B) + K_BC_B + K_RC_R}$
	$B + s \rightleftharpoons Bs$	$r = \dfrac{k(C_B - C_R/KC_A)}{1 + K_AC_A + (K_{RA}C_R/C_A) + K_RC_R}$
	$As + Bs \rightleftharpoons Rs + s$	$r = \dfrac{k(C_AC_B - C_R/K)}{(1 + K_AC_A + K_BC_B + K_RC_R)^2}$
	$Rs \rightleftharpoons R + s$	$r = \dfrac{k(C_AC_B - C_R/K)}{1 + K_AC_A + K_BC_B + K_{AB}C_AC_B}$
$A + B \rightleftharpoons R + S$	$A + s \rightleftharpoons As$	$r = \dfrac{k(C_A - C_RC_S/KC_B)}{1 + (K_{RS}C_RC_S/C_B) + K_RC_B + K_RC_R + K_SC_S}$
	$B + s \rightleftharpoons Bs$	$r = \dfrac{k(C_B - C_RC_S/KC_A)}{(1 + K_{RS}C_RC_S/C_A^{1/2} + K_AC_A + K_RC_R + K_SC_S)}$
	$As + Bs \rightleftharpoons Rs + Ss$	$r = \dfrac{k(C_AC_B - C_RC_S/K)}{(1 + K_AC_A + K_BC_B + K_RC_R + K_SC_S)^2}$
	$\left.\begin{array}{c} Rs \rightleftharpoons R + s \\ Ss \rightleftharpoons S + s \end{array}\right\}$	$r = \dfrac{k[(C_AC_B/C_S) - C_R/K]}{1 + K_AC_A + K_BC_B + K_SC_S + K_{AB}C_AC_B/C_s}$
$A + B \rightleftharpoons R + S$	$A + 2s \rightleftharpoons 2A^{1/2}s$	$r = \dfrac{k(C_A - C_RC_S/KC_B)}{[1 + K_{RS}C_RC_S/C_B + K_BC_B + K_RC_R + K_SC_S]^2}$
	$B + s \rightleftharpoons Bs$	$r = \dfrac{k(C_B - C_RC_S/KC_A)}{1 + \sqrt{K_AC_A} + (K_{RS}C_RC_S/C_A) + K_RC_R + K_SC_S}$
	$2A_{1/2}s + Bs \rightleftharpoons Rs$ $+ Ss + s$	$r = \dfrac{k(C_AC_B - C_RC_S/K)}{(1 + \sqrt{K_AC_A} + K_BC_B + K_RC_R + K_SC_S)^2}$
	$Rs \rightleftharpoons R + s$	$r = \dfrac{k(C_AC_B/C_S - C_R/K)}{1 + K_A\sqrt{C_A} + K_BC_B + (K_{AB}C_AC_B/C_S) + K_SC_s}$
	$Ss \rightleftharpoons S + s$	$r = \dfrac{k(C_AC_B/C_R - C_s/K)}{1 + \sqrt{K_AC_A} + K_BC_B + K_RC_R + K_{AB}C_AC_B/C_R}$
$A + B \rightleftharpoons R + S$	$B + s \rightleftharpoons Bs$	$r = \dfrac{k(C_B - C_BC_R/KC_A)}{1 + K_RC_R + K_{RS}C_RC_S/C_A}$
	$A + Bs \rightleftharpoons Rs + S$	$r = \dfrac{k(C_AC_B - C_RC_S/K)}{1 + K_RC_R + K_BC_B}$
	$Rs \rightleftharpoons R + s$	$r = \dfrac{k[(C_AC_B/C_S) - C_R/K]}{1 + (K_{AB}C_AC_B/C_S) + K_BC_B}$

NOTE: K_{AB} . . . = combined equilibrium constants; K = over-all equilibrium constant for the chemical equation; k = constant.

*The rate equation is opposite the catalytic step assumed to be rate-controlling.
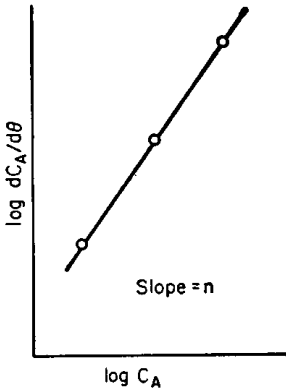
FIGURE 15.13 Determining reaction order: differentiation. (*From Walas,*[19] *Fig. 2.1.*)
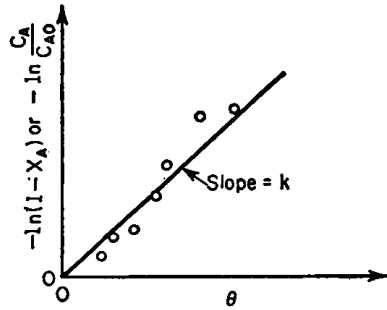
FIGURE 15.14 Test for first-order reaction. (*From Levenspiel,*[18] *p. 48.*)

$$-\frac{dC_A}{d\theta} = kC_A^n \tag{15.56}$$

is integrated between the limits C_{AO} and C_A to give

$$\left(\frac{1}{C_A}\right)^{n-1} - \left(\frac{1}{C_{AO}}\right)^{n-1} = (n-1)k\theta \tag{15.57}$$

From Eq. (15.57), a plot of $(1/C_A)^{n-1}$ against θ yields a straight line with the intercept $(1/C_{AO})^{n-1}$ and the slope $k(n-1)$ (Fig. 15.15).

4. *Integrated rate equation: Method of half-life:* The *half-life,* or time for 50 percent conversion, is a useful criterion for order. Integration of Eq. (15.56) between the limits C_{AO} and $0.5C_{AO}$ yields the following values for half-life $\theta_{1/2}$:
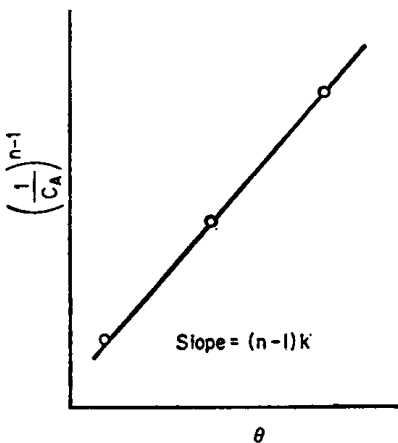
FIGURE 15.15 Determining order of action: integrated equation. (*From Walas,*[19] *Fig. 2.2.*)
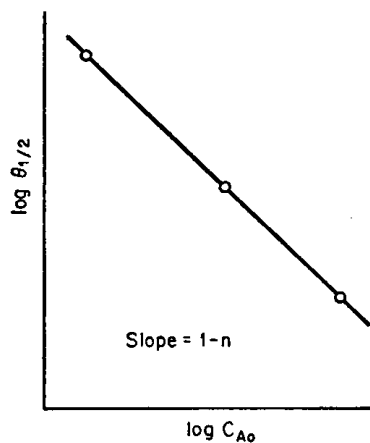
FIGURE 15.16 Determining order of reaction: half-lives. (*From Walas,*[19] *Fig. 2.3.*)

$$\theta_{1/2} = \begin{cases} \dfrac{1}{2k} C_{AO} & n = 0 \\[3mm] \dfrac{0.69}{k} & n = 1 \\[3mm] \dfrac{1}{kC_{AO}} & n = 2 \\[3mm] \dfrac{2^{n-1} - 1}{k(n - 1)(C_{AO})^{n-1}} & n = 3 \end{cases}$$

Hence, if experiments are run at different initial concentrations and log $\theta_{1/2}$ is plotted against log C_{AO}, a straight line of slope $1 - n$ results (Fig. 15.16).

5. *Integrated rate equation: Method of reference curves:* Inspection of the integrated rate equations indicates that the ratio of times required for any two degrees of conversion is dependent only on those conversion fractions and on the reaction order. Walas[19] has calculated such ratios, taking 90 percent conversion as the arbitrary convenient reference.

24. Reactor Design. Once a suitable rate equation that fits the experimental kinetic data has been discovered, design of the plant reactor can proceed. Five steps are involved:

1. Selection of the type of reactor
2. Selection of the shape or proportions of reactor
3. Sizing the reactor
4. Selection of materials of construction
5. Design of reactor auxiliaries

Sizing the Reactor. The volume of a reactor is calculated directly from the rate equation, such as Eq. (15.49), which may be rewritten as

$$\int_0^\theta V d\theta = \int_{N_{A_2}}^{N_{A_1}} \frac{dN_A}{r_A} \tag{15.58}$$

the correct-order expression being inserted for r_A. For flow reactors, a more useful form of Eq. (15.58) is

$$\int_0^V \frac{dV}{F} = \int_0^{X_A} \frac{dX_A}{r_A} \tag{15.59}$$

written in terms of the molal flow rate F and degree of conversion X_A, moles of A converted per mole of F. For steady-state continuous stirred tanks, the rate expression is simply

$$r_A = (F/V)(C_{A1} - C_{A2}) \tag{15.60}$$

where the subscripts 1 and 2 refer to the concentration of reactant A in the entering and leaving streams, respectively.

NOMENCLATURE

Basic Dimensions.   Basic dimensions of parameters are given in terms of the following symbols and may be found with various compatible combinations of the units shown.

L = length, ft, cm

θ = times, s, h

T = temperature, °R or K

M = mass, lb mass, grams, tons

moles = mass/molecular weight, gram-moles, lb-moles

F = force, lb force, newtons

conc. = concentration, moles per unit volume

E = energy, cal, Btu, joules

Nomenclature

A = filter area, L^2

= collector plate area, L^2

= heat-transfer surface area, L^2

= membrane area, L^2

= chemical reactant species

= frequency factor, $\text{moles}/(\theta \cdot \text{conc.}^{n-1})$

a = surface area of packing per unit volume of contactor, L^2/L^3

= stoichiometric coefficient for reactant A

B = chemical reactant species

b = stoichiometric coefficient for reactant B

C = chemical reactant species

= concentration, when subscripted moles/L^3

C_D = drag coefficient

C_F = heat capacity of feed, $E/(M \cdot T)$

c = stoichiometric coefficient of reactant C

= controlled variable

D = distillate product rate, moles/θ

= chemical reactant species

D_a = impeller diameter, L

D_P = particle diameter, L

= diameter of ion-exchange resin bead, L

D_i = fluid-phase diffusion coefficient, L^2/θ

d = indicates differentiation

d_F = packing characteristic length, L

E = energy of activation, E

     = extract flow rate, M/θ

E_c = column plate efficiency

exp = exponential function (natural antilogarithm)

F = feed flow rate, moles$/\theta$, M/θ, L^3/θ

f = amount of vapor introduced to flash still or feed tray in rectifying column, moles/mole

G = gas volumetric flow rate, L^3/θ

     = rate of solute-free solvent removed in solvent separator, M/θ

G_M = gas molar flow rate, moles$/\theta$

g = gravitational acceleration, L/θ^2

g_c = gravitational constant, $LM/(\theta^2 F)$

H = velocity head of agitated fluid, E/M

H_{OG} = height of transfer unit (HTU) based on overall (gas-phase) driving force, L

H_{OL} = height of transfer unit (HTU) based on overall (liquid-phase) driving force, L

K' = distribution coefficient for liquid-liquid equilibrium

K_c = equilibrium constant

K_G = overall gas-phase mass-transfer coefficient, moles$/(L^2\theta$ conc.)

K_L = overall liquid-phase mass-transfer coefficient, moles$/(\theta L^2$ conc.)

k = reaction rate constant, forward reaction

k' = reaction rate constant, reverse reaction

k_G = gas-phase mass-transfer coefficient, moles$/(\theta L^2$ conc.)

k_L = liquid-phase mass-transfer coefficient, moles$/(\theta L^2$ conc.)

L = membrane thickness, L

     = liquid molar flow rate in distillation column, moles$/\theta$

L_M = liquid molar flow rate, moles$/\theta$

L_e = extract-layer flow rate from last stage, M/θ

L_r = raffinate-layer flow rate from last stage, M/θ

l = nominal pore length in catalyst particle (radius of sphere or cylinder, half-thickness of slab), L

m = slope of vapor-liquid equilibrium line

N = rotation speed of impeller, rpm

     = number of equilibrium stages

     = number of moles of subscripted species

N_{OG} = number of transfer units, based on over-all (gas-phase) driving force

N_{OL} = number of transfer units, based on over-all (liquid-phase) driving force

N_M = molar mass-transfer rate, moles$/\theta$

N_p = impeller power number

N_Q = impeller discharge coefficient

N_T = number of theoretical stages at total reflux

N_{act} = actual number of stages in column

N_{theor} = number of equilibrium stages (theoretical stages) in column

n = order of chemical reaction

p = permeability of membrane

P = pressure of liquid, F/L^2

= agitator shaft power, E/θ

= chemical reactant species

= vapor pressure of subscripted component, F/L^2

= number of moles in still-pot (batch distillation)

p = partial pressure of subscripted component, F/L^2

= stoichiometric coefficient of reactant P

Q = impeller discharge rate, L^3/θ

= rate of feed of solute-free solvent to raffinate end of extractor, M/θ

= chemical reactant species

q = rate of heat transfer, E/θ

= stoichiometric coefficient of reactant Q

R = ideal gas constant, $E/(T \cdot \text{moles})$

= raffinate-layer flow rate, M/θ

= reflux ratio

r = specific reaction rate, or specific reaction rate of subscripted species (forward reaction), moles$/(\theta L^3)$

r' = specific reaction rate (reverse reaction), moles$/(\theta L^3)$

$\bar{r}$ = average pore radius, L

τ_{net} = net forward rate of a reversible reaction, moles$/(\theta L^3)$

S = amount of material distilled in Rayleigh distillation, moles

= solvent content of extract layer, $M_{solvent}/M_{dissolved\,material}$

s = empirical constant for filter-cake compressibility

= average of mutual solubilities of solute-free contacted liquids

= solvent content of raffinate layer, $M_{solvent}/M_{dissolved\,solids}$

T = absolute temperature, °R or K

t = temperature, °C or °F

U = overall heat transfer coefficient, $E/(L^2\theta T)$

U_i = dialysis coefficient in subscripted phase, moles$/(L^2\theta)$ conc.

V = total volume upstream of a given point in a plug flow reactor, L^3

= volume of contactor or reactor, L^3

= molar vapor rate in column, M/θ

V_C = superficial velocity of continuous phase, $L^3/(L^2\theta)$

V_D = superficial velocity of dispersed phase, $L^3/(L^2\theta)$

V_e = extract-layer flow rate from last stage, M/θ

V_r = raffinate-layer flow rate from last stage, M/θ

v = total volume of feed to ion-exchange process, L^3

v_t = terminal settling velocity, L/θ

W = mass of solids retained in filter cake, M

    = bottom product rate, moles$/\theta$

w = equilibrium composition of raffinate phase, $M_{solute}/M_{solute\text{-}free\,phase}$

w' = equilibrium composition of extract phase, $M_{solution}/M_{solute\text{-}free\,phase}$

X = mass fraction solute in raffinate, solvent-free basis

    = moles of subscripted component converted per mole of feed to continuous reactor

x = mole fraction (of subscripted component) in liquid phase

x^* = mole fraction in liquid phase which would be in equilibrium with actual gas phase at that point in contactor

Y = mass fraction of solute in extract, solvent-free basis

y = mole fraction (of subscripted component) in vapor or gas phase

y^* = mole fraction in vapor phase which would be in equilibrium with actual liquid phase at that point in contactor

Z = total column height, L

Z_t = tray spacing, L

Z_f = solvent content of feed, $M_{solvent}/M_{solvent\text{-}free\,stream}$

α, α_{12} = relative volatility

α = reaction order with respect to reactant A

    = specific cake resistance, L/M

    = empirical constant for wet scrubber efficiencies

α' = empirical constant for filter cake resistance

β = reaction order with respect to reactant B

γ = activity coefficient

    = empirical constant for wet scrubber efficiencies

    reaction order with respect to reactant C

Δ = difference or change

ΔP_c = pressure drop across filter cake, F/L^2

θ = time

$\theta_{1/2}$ = reaction half-life

λ_m = molar latent heat of vaporization, $E/moles$

λ = latent heat of vaporization, E/M

μ = viscosity $M/(L\theta)$

Π = total pressure, F/L^2

ρ = density, M/L^3

    = reaction order with respect to reactant P

ρ_f = density of fluid, M/L^3

ρ_p = density of particle, M/L^3

σ = reaction order with respect to reactant Q

σ = interfacial tension, lb mass/hr^2

σ' = interfacial tension, dynes/cm

Subscripts

A = reactant A

B = reactant B

C = reactant C

   = continuous phase

D = disperse phase

d = distillate product

F = feed stream

i = component i

   = at interface

L_1 = liquid phase 1

L_2 = liquid phase 2

lm = logarithmic mean

M = molar property

m = molar property

O = initial value

p = in still pot

P = reactant P

Q = reactant Q

w = bottom product

in = entering

out = leaving

1 = component 1

   = at position 1 (upstream) or initial time

2 = component 2

   = at position 2 (downstream) or final time

ENVIRONMENTAL ENGINEERING*

INTRODUCTION

The concept of environmental engineering is relatively new. Historically, governments have concerned themselves with only the most flagrant forms of pollution, such as a public water supply contaminated with sewage. In general, the world's atmosphere and bodies of water were considered a limitless garbage dump, and natural purification processes were sufficient to handle human wastes for thousands of years. But now the combined impacts of continued industrial development and population growth make environmental engineering a necessity.

The field spans an enormous range of activities and embraces all the traditional engineering disciplines. This section will concentrate on the legal and technological aspects of wastewater treatment and control of air pollution from stationary sources.

1. Clean Air Act of 1970.   The Clean Air Act of 1970 was the first law to require national ambient air quality standards based on geographic regions. Air quality is regulated by two sets of standards determined by EPA. The primary standards are intended to be the minimum level of air quality that is necessary to preserve human health, while the secondary standards are aimed at preventing damage to animals, plant life, and property.

State governments retain the authority to determine how ambient air quality standards are to be met within their borders. However, the state implementation plans are subject to approval by EPA, which can revise any state plan it considers unsatisfactory.

2. Water Act of 1972.   The U.S. Congress tackled water pollution with the Water Pollution Control Act Amendments of 1972. This bill is considerably more stringent than any previous water legislation and provides penalties up to $25,000 per day for willful or negligent violations. All discharges to navigable waters must have been treated by best practicable technology by July 1, 1977 and best available technology by July 1, 1983. The EPA holds responsibility for determining exactly what is required under the two levels of treatment technology.

During 1973 and 1974, EPA set effluent guidelines for 28 industry categories. These serve as the basis for pollutant limitations in the discharge permits. The guidelines sparked much controversy (and a number of legal suits). Further action by the Congress may be required. The objections to the guidelines involve complex legal questions, but in general they are being challenged on the grounds that EPA exceeded its statutory authority under the 1972 law.

3. Information Sources.   The engineer who needs to determine the specific pollution-control requirements for a particular operation in a certain location has many sources of information available—too numerous to be listed here. In general, the starting point for legal information should be at the lowest applicable level of

---

government. Usually this is a state body, although in metropolitan areas a county or city agency may have jurisdiction.

WASTEWATER TREATMENT

4. Preliminary Treatment

Equalization.   Wastewater flow rates are erratic. Municipal flow rates peak during the day and reach a minimum during the night; within these periods, the flow rises and falls with the time of day and with the weather (rainfall).

Neutralization.   Biological wastewater treatment proceeds optimally at pH values near 7 (neutrality). Slight deviations from this value will reduce efficiency, while large differences may result in total inactivation of the bacteria.
    While municipal sewage normally needs no pH adjustment, industrial waste quite often does. Using pH recorder-controllers, some incorporating analog computers, acids and bases can be added in optimum fashion.

Oil Removal.   Generally, oil concentrations greater than 50 mg/L inhibit biological action. Thus gravity settling basins are usually installed to separate oil and water when they are combined in a nonemulsified form.
    In general, oil-water separators are rectangular, multichannel structures that produce low flow velocity and a minimum of turbulence while allowing time for the oil globules to float to the surface and be removed.

Metallic-Ion Removal.   Ions of the heavy metals (Cu, Zn, Al, Fe, Cr, etc.) in concentrations greater than 1 to 10 mg/L, depending on their nature, must be removed because they inhibit subsequent treatment processes. Adding lime, raising the pH to about 10.5, normally results in a reasonable reduction of metallic-ion concentration, because many of the metals form insoluble hydroxides at this pH level.

5. Primary Treatment.   Screens, comminutors, and/or grit chambers usually precede primary sedimentation in municipal plants. Screens, normally composed or iron bars or grates with ¾- to 3-in (1.9- to 7.6-cm) openings, remove the largest suspended particles, while comminutors grind large particles into smaller ones. Grit chambers are designed to remove the heavier solids (i.e., inorganic matter such as sand whose specific gravity is greater than 2) through a slight reduction in flow velocity.
    Simplistic design of sedimentation systems is accomplished by choosing either a residence detention time (generally 90 to 150 min) or an overflow rate [the normal range is 600 to 1200 gal/ft^2 per day (24,466 to 48,891 L/m^2 per day)]. From these data, the tank depth can easily be selected.

Chemical Addition.   Chemicals are being used increasingly in the initial stages of wastewater treatment for (1) phosphorus removal, (2) increased efficiency of BOD removal in the primaries, and (3) proper conditioning of the wastewater for filtration, carbon adsorption, and reverse osmosis. Hence the proper selection of chem-

icals that assist in the aggregation of colloidal particles (0.001 to 10 μm) is crucial to the process operation.

6. Bioxidation.   Simple sedimentation removes a major fraction of the suspended solids but none of the dissolved contaminants. Microorganisms can utilize dissolved organic material as nutrients for their growth and metabolism. The process, shown mechanistically in Fig. 15.17, could be termed *flameless oxidation* in the liquid state because bacteria convert carbon to carbon dioxide via metabolic pathways, therein consuming oxygen and yielding energy.

There are two different biological oxidation, or bioxidation, processes, with the distinction between them based on microbial mobility:

1. Suspended—the microorganisms are suspended in the liquid; the system operates as a continuous biological reactor, called an *activated sludge plant* or *oxidation pond.*

2. Fixed bed—the microorganisms are embedded in a gelatinous mass attached to a solid support medium. This is referred to as a *trickling filter* or *bioxidation tower.*

Three types of reactions take place in the bioxidation process: (1) assimilative respiration, (2) synthesis of cells, and (3) endogenous respiration. Assimilative respiration provides the necessary energy for the life processes of cells and can be represented by the following equation:

$$C_xH_yO_z + (x + \tfrac{1}{4}y - \tfrac{1}{2}z)O_2 \rightarrow CO_2 + \tfrac{1}{2}yH_2O + \text{energy} \qquad (15.61)$$
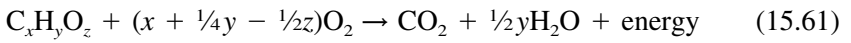
Besides the energy released, the cells derive energy for the oxidation process to use in building new cells:
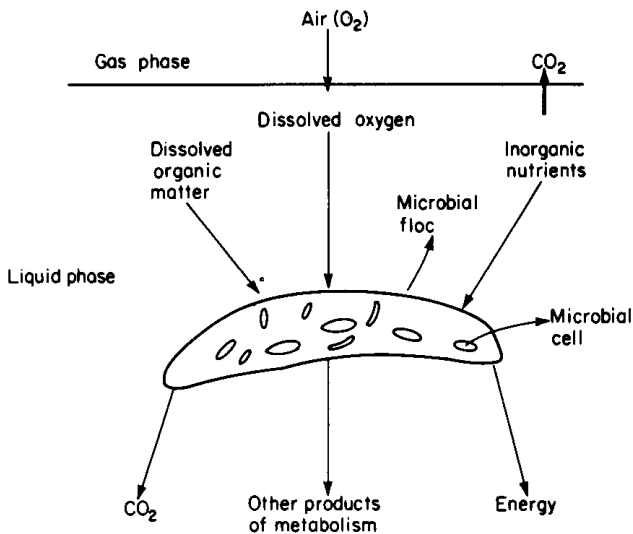
FIGURE 15.17   Schematic for bioxidation of dissolved organic matter.

$$n(C_xH_yO_z) + nNH_3 + n(x - \tfrac{1}{4}y - \tfrac{1}{2}z - 5)O_2 \rightarrow (C_5H_7NO_2)n$$
$$+ n(x - 5)CO_2 + \tfrac{1}{2}n(y - 4)H_2O \quad (15.62)$$
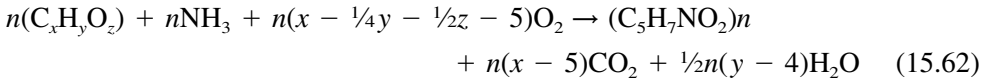
This process is termed *endogenous respiration;* it is actually autoxidation or self-destruction of cellular material and can be represented by

$$(C_5H_7NO_2)n + 5nO_2 \rightarrow 5nCO_2 + 2nH_2O + nNH_3 + \text{energy} \quad (15.63)$$
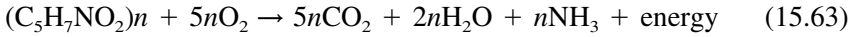
The formula representative of bacterial cells is $C_5H_7NO_2$. The need for nitrogen in order to grow cells is obvious by its inclusion in the formula.

Activated Sludge System.   The conventional activated sludge process is analogous to a continuous biological fermentor because nutrients in the untreated wastewater are continually added, oxygen is supplied through an aeration system, and cells are produced and removed in the effluent stream (Fig. 15.18). The product is a combination of treated wastewater and suspended cells, which are removed by sedimentation and partially recycled to the reactor. The treated effluent overflows to discharge or reuse systems.

The activated sludge process provides a smooth-running, highly efficient operation, with excellent removal of organic matter; BOD removals of 90 percent are easily achieved.

Tapered Aeration.   As the organic contaminants are consumed by the bacteria in a long reactor, the F/M ratio and oxygen demand decrease with length. The process modification from conventional activated sludge is a variation of the airflow rate according to oxygen needs achieved by putting in decreasing numbers of diffuser tubes with increasing distance from the inlet. The obvious advantage is lower power consumption than uniform (and excess) aeration would require.
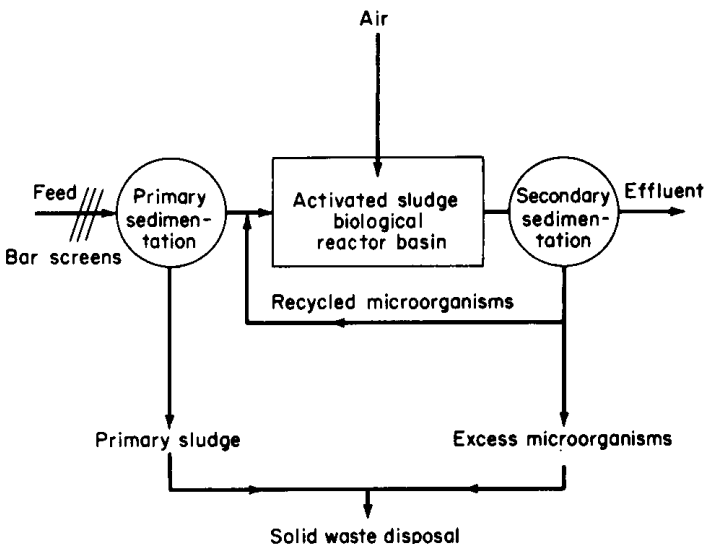
FIGURE 15.18   Simplified flowsheet for activated sludge plant.

Step Aeration.   Another way to balance air requirements is to add the wastewater at different points along the aeration tank. This modification keeps the *F/M* ratio constant as a function of tank length and leads to uniform oxygen demand by the microorganisms. With these changes, BOD loading may be increased and detention time and consumption rate may be lowered.

Contact Stabilization.   The adsorptive properties of activated sludge particles are advantageous when a large part of the BOD is present in suspended form. Sludge that has been conditioned by reaeration is brought into short (15 to 30 min) but unaerated contact with the wastewater.

Extended Aeration.   Extended aeration systems operate essentially without any sludge production because they include neither a primary sedimentation tank nor any allowance for continuous sludge withdrawal from the bioxidation system.

Aeration tanks usually are sized to provide 24- to 30-h retention with air added mechanically or by submerged diffusers. A final settling basin where the detention is approximately 4 h removes most solids for total recycle—although normally provision is made for intermittent sludge withdrawal to holding tanks for later off-site disposal. In general, the process functions best at BOD loadings of less than 15 lb (6.8 kg) per day per 1000 ft^3 (28.3 m^3) of aeration tank and less than 0.1 lb BOD/lb (6.05 kg BOD/kg) of cells.

Oxygen Demand and Supply.   Generally, 2 mg/L of dissolved oxygen should be maintained at all points of the activated sludge basin to keep the cells working at their maximum rate. Hence the mixed liquor in the reactor is continuously aerated by either a diffused air system or by mechanical surface aerators. The aeration process serves three functions: (1) it mixes the sludge and the sewage, (2) it keeps the sludge in suspension, and (3) it supplies the oxygen needed for bioxidation.

Lagoons.   The term *lagoon* or *stabilization pond* generally is applied to all bodies of water artificially created with the intention of retaining sewage or organically contaminated wastewaters until the wastes are rendered stable or unobjectionable through biological decomposition and the waters are suitable for disposition, either by discharge into receiving waters or by way of ground seepage and evaporation.

Lagooning is a common and cheap means of eliminating dissolved or suspended organics. The feed enters at one end of the lagoon. Sedimentation occurs first, with settable solids being deposited near the entrance.

Trickling Filters.   *Trickling,* or *biological, filter* is the name given to the process whereby microorganisms attached to stones (or plastic media) strip wastewater of its organic components as it flows (trickles) down through a packed tower. The attached growths adsorb, absorb, and oxidize suspended and colloidal matter in the wastewater. Crushed stone [1½ to 4 in (3.8 to 10.2 cm)] normally is used as the packing medium for the bed, although recently plastic packing has found substantial application. The wastewater is applied to the filter through a central rotating distributor, which intermittently applies water to the various segments of the bed.

Design Equations.   The rate of removal of BOD in a filter is a function of the BOD concentration of the wastewater and the adsorptive capacity of the biological growth. The rate of stabilization controls the adsorption capacity of the biological growth.

Eckenfelder[21] has proposed equations for design of rock filters with and without recirculation:

No recirculation:
$$\frac{L_c}{L_0} = \frac{100}{1 + 2.5(D^{0.67}/Q^{0.50})} \tag{15.64}$$

Recirculation:
$$\frac{L_c}{L_0} = \frac{1}{(1 + N)[1 + 2.5(D^{0.67}/Q^{0.5})] - N} \tag{15.65}$$

when
$$L_0 = \frac{L_a + NL_c}{N + 1}$$

where L_c = BOD of effluent, mg/L
L_a = influent BOD, mg/L
L_0 = BOD of mixed influent plus recirculated streams, mg/L
N = recirculation ratio
D = filter depth, ft (m)
Q = hydraulic application rate, million gal/(acre · day) [million L/(m² · day)]

Temperature Effect.   The effect of temperature on treatment efficiency is given by the formula

$$E = E_{20}1.035^{(T-20)} \tag{15.66}$$

when E is the efficiency of BOD removal at any temperature $T(°C)$, and E_{20} is the efficiency of removal at 20°C.

Plastic Packing.   Plastic media have found wide acceptance in municipal and industrial systems. The medium consists of corrugated plastic sheets running parallel to each other and bonded together to form a bundle or a pack. The uniform but open pattern of the plastic grid largely eliminates clogging problems that exist with rock and allows the grid to operate at higher organic loadings.

The basic design formula for plastic-media trickling filters was developed by Germain:[22]

$$\frac{L_c}{L_0} = e^{-kD/Q^{1/2}} \tag{15.67}$$

where D = the depth of the filter, ft (m)
Q = the hydraulic dosage rate of primary effluent, gal/(min · ft²) [L/(min · m²)]
k = the reaction rate coefficient (treatability factor) in consistent units to render the exponent dimensionless

The treatability factor is important because its magnitude determines rate of BOD removal. For municipal sewage, Germain determined a value of 0.088.

7. Sludge Treatment.   *Sludge,* a suspension of solids in water, must be processed further to reduce its high water content and to reduce the concentration of organic matter. If untreated and used as landfill, it will decay and produce an offensive odor. Concentration and stabilization are the primary functions of the sludge-

processing system that can constitute as much as 25 to 50 percent of the operating and capital cost of a wastewater treatment plant.

Conditioning.   Sludge is most often conditioned chemically. Coalescence can be enhanced, in theory, by chemicals that neutralize charges on suspended particles, causing them to agglomerate and simultaneously lessening the particle's tendency to bind water.

Synthetic polymers have taken on a role of increasing importance in flocculation, but inorganic compounds such as alum and ferric chloride are still widely used. Other conditioning agents include fly ash and diatomaceous earth.

Gravity Thickening.   Gravity thickening basins are usually circular and are provided with a raking mechanism to convey the solids to the point of discharge at the center of the tank. For municipal sewage sludges, the following loading rates, given in lb/ft^2 per day (kg/m^2 per day), are recommended for thickeners: (1) 22 (107.4) for primary sludge, (2) 15 (73.2) for primary plus trickling-filter biofloc, (3) 8 to 12 (39.1 to 58.6) for a blend of primary and waste-activated sludge; and (4) 4 (19.5) for waste-activated sludge alone.

Flotation.   Flotation is attractive for sludge thickening because of a faster dewatering rate that is not affected adversely by decomposing solids, which evolve gases and cause sludge bulking. In flotation units, air is dissolved under pressure in the liquid being treated. When the pressurized air-wastewater stream is discharged into an aeration tank, maintained at atmospheric pressure, the supersaturated air nucleates as small bubbles in the 10- to 100-μm range.

Aerobic Digestion.   In *aerobic digestion,* waste sludge is simultaneously the food and the oxidation system; i.e., in autoxidation or endogenous systems the living cells utilize nutrients released when other cells die and dissolve. Thus the microbial population and the amount of biodegradable organic matter are continually decreasing. The process is carried out until the volatile suspended solids are reduced to about 50 percent, a level at which the sludge product will not cause a nuisance. The residence time should be about 15 days to produce a good sludge that forms a filter cake of 20 to 25 percent solids.

Anaerobic Digestion.   *Anaerobic digestion* is the most popular method of sludge stabilization. The term applies to the process in which organic material is decomposed biologically in an environment devoid of oxygen. This decomposition results from the action of two major groups of bacteria: acid formers, consisting of facultative bacteria that convert carbohydrates, fats, and proteins to organic acids and alcohols; and methane bacteria, which are strict anaerobes that convert the organic acids and alcohols produced by the acid formers into carbon dioxide and methane. Small amounts of hydrogen and hydrogen sulfide are also formed.

Thermal Conditioning.   Heat treatment of sludge under pressure for about 30 min produces a sludge that is easily dewatered without the use of chemicals. Essentially, the process consists of grinding the sludge to eliminate large particles, and sparging the sludge with live steam in a reactor for 30 to 45 min at 250 lb/in$_g^2$ (1.7 MPa$_g$).

Sludge is withdrawn to a decanter where the solid material settles rapidly to about one-third of its original volume. Vacuum filters handling heat-treated sludge average 12 lb/ft^3 per day (58.6 kg/m^3 per day). The filter cake, which prior to heat treatment averaged 20 to 25 percent solids, is increased to 40 to 45 percent.

Dewatering. Used for many years, applicable to a wide variety of waste, and relatively cheap, vacuum filtration is a popular method of mechanically dewatering sludge. A typical vacuum filter consists of a rotating cylindrical drum partially submerged in sludge. A vacuum of 10 to 20 inHg (254 to 508 mmHg) is applied to the center of the drum, drawing the water through the filter medium but leaving the solids deposited on the periphery. Continual application of vacuum as the drum rotates out of the liquid removes more water. Just before the drum reenters the sludge, the cake is removed.

Pressure Filtration. This essentially batch operation is relatively labor-intensive, but the product is so dry (50 to 60 percent solids) that in many cases the advantages in disposal of the resultant solid waste outweigh the cost of labor. Besides a well-dewatered product, the filtrate generally contains a low concentration (20 mg/L or less) of suspended solids.

Centrifugation. In recent years, centrifuges have gained increasing acceptance in the handling of municipal and industrial sludges. Among the reasons are the development of centrifuges requiring less maintenance than in the past while operating at higher feed rates and greater sludge recovery. Disadvantages still include the maintenance problems and noise that are inherent with high-speed equipment.

Sand Drying Beds. Although mechanical methods of sludge dewatering have received the greatest research attention recently, operators, especially those of small plants, still use sand beds effectively. Drying beds are constructed by placing 4 to 6 in (102 to 152 mm) of sand over graded layers of gravel or crushed stone underlaid by tile drains. The sides of the beds are made of concrete or wood; internally the beds are subdivided into cells, usually less than 20 ft (6.1 m) wide and 100 ft (30.5 m) long.

8. Sludge Disposal. Waste treatment does not destroy pollution, but merely converts the pollutant to another form. Hence a large fraction of wastewater contaminants, both suspended and dissolved, are converted to a solid waste whose ultimate disposal location is on the land, in the atmosphere, or in the ocean. Economics, location, and environmental impact are the three factors normally controlling process selection.

Incineration. With suitable sanitary landfill areas rapidly disappearing and ocean dumping being closely scrutinized, incineration is becoming an increasingly popular method of sludge disposal. Burning accomplishes a significant reduction in sludge volume and produces a sterile ash. The fundamental concern is the supplementary heat required (if any) to sustain combustion. Dried solids contain considerable heat value. Based on the ultimate analysis, the heat value is

$$Q = 14,600 \text{ C} + 62,000 \text{ (H} - \text{O/8)} \tag{15.68}$$

where Q = Btu/lb (J/kg) dried sludge
     C = percent carbon
     H = percent hydrogen
     O = percent oxygen

A typical sewage sludge will have a heat value of 7000 Btu/lb (16.3 MJ/kg) dry

solids. If inorganic coagulants are used in the thickening process, this value is reduced.

Land Spreading. There are two principal methods of disposing of sludge on land, a process that has the advantage of recovering the fertilizer value of the sludge. Drying and sale of the material for fertilizer have had limited success. The city of Milwaukee for many years has produced Milorganite, a fertilizer that gained prominence largely due to marketing techniques. However, synthetic fertilizers have provided severe competition, because they deliver more nitrogen and phosphorous at lower cost.

The sludge also can be spread wet, thus avoiding the cost of drying. Many farmers seek this material if available locally, although there may be problems with odor and possible accumulation of heavy metals in the soil.

Ocean Disposal. For cities near the ocean, barging or pipeline transport of sludge to deep water has been popular because ocean disposal is cheap and simple. Although all traces of sludge disappear within ½ h of the discharge and studies by large cities using the method claim no detectable degradation of the sea, there are serious concerns over the environmental impact of this method. Thus ocean dumping of sludge will probably not last.

9. Tertiary Treatment. *Tertiary treatment* is defined as any process that follows secondary biological systems. For the purpose of this section, it has been divided into four areas: removal of nutrients, fine solids, organic material, and dissolved solids. None of the steps to remove these pollutants is cheap in view of the incremental cost. Microscreening, for example, can reduce BOD by a further 5 percent at a cost of 1.5 cents per 1000 gal (1.5 cents per 3785 L)—compared to a 90 percent reduction in secondary treatment for 11 cents. For dissolved solids removal, the cost increases dramatically.

Nitrogen. Although conventional sewage plants remove some nitrogen, there are still varying amounts released in diverse forms, such as ammonia, organic nitrogen, and nitrates. A process for nitrogen removal receiving much attention at the present time is nitrification-denitrification. Nitrogen is first oxidized to its highest oxidation state; the nitrates are then reduced to nitrogen, which is air-stripped. Nitrification occurs in two steps. The bacteria capable of oxidizing ammonia to nitrate are believed to be primarily autotrophic nitrifiers, among which *Nitrosomonas* are most common:

$$NH_4^+ + 1.50_2 \rightarrow 2H^+ + H_2O + NO_2^- \tag{15.69}$$
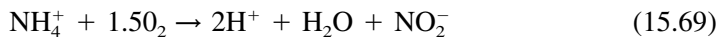
The second phase of the reaction, nitrite to nitrate, is carried out by *Nitrobacter:*

$$NO_2^- + 0.50_2 \rightarrow NO_2^- \tag{15.70}$$

One factor of great importance in nitrification is the large amount of oxygen required per unit of ammonia—theoretically 4.56 mg O_2/mg NH_3.

Denitrification is also a biological process in which bacteria, in the absence of free molecular oxygen, reduce the nitrates to nitrogen. In order to achieve a reasonable efficiency, a supplementary source of carbon, commonly methanol, is added to accomplish the following:

$$5CH_3OH + 6NO_3^- \rightarrow 5CO_2 + 7H_2O + 3N_2 + 6OH^- \qquad (15.71)$$
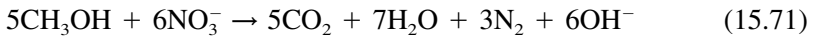
The air-stripping process for ammonia removal consists in raising the pH to 10.5 to 11.5 and degassing the ammonia through air-water contact. Cooling towers have been used for this step.

Ammonia removal proceeds according to the reaction

$$3Cl_2 + 2NH_3 \rightarrow N_2 + 6HCl \qquad (15.72)$$

The *breakpoint* is defined as the point where the nitrogen in the form of ammonia is eliminated and free available chlorine is detected.

Removal of Trace Suspended Solids. The characterization of trace suspended solids is difficult. Particle sized, concentration, and physical and chemical properties (specific gravity, toxicity, stickiness, etc.) are all highly variable and dependent on the pollution source. Sewage plant effluents, especially those from biological plants, contain small amounts of suspended solids that can have a serious impact on the receiving body of water.

Microstraining. Essentially, a *microstrainer* consists of a rotating drum with a fine screen around its periphery. Feedwater enters the drum through an open end and passes radially through the screen while solids are deposited on the inner surface of the screen.

The effectiveness of removal depends on the screen pore size. The screens used in microstraining are made of a variety of plastics and stainless steel and have extremely small openings. They have high porosity to effectuate high flow rate at low pressure drops [6 in (152 mm) due to the fabric and 12 to 15 in (305 to 381 mm) overall].

Sand Filtration. The most common filter medium is a graded bed of silica sand. Developed in Great Britain and used for water clarification, these filters were operated at rates of 0.04 to 0.12 gpm/ft² [0.027 to 0.081 L/(s · m²)]. In the United States, precoagulation has increased flow rates to 1 to 4 gpm/ft² [0.68 to 2.7 L/(s · m²)]. The two processes have thus been called *slow* and *rapid sand filtration.*

Mixed-Media Filter. The mixed-media filter uses combinations of anthracite coal, sand, garnet, and ilmenite, with average specific gravities of 1.5, 2.6, 4.2 and 4.8, respectively. Properly formed after washing, the bed has the largest particles on top and the smallest on the bottom (if a downflow filter).

Organic Removal by Activated Carbon. Activated carbon selectively adsorbs dissolved organics (phenol, xylenes, etc.) and has a high capacity because of its large surface area, typically 500 to 1400 m²/g.

Activated carbon is used to produce high-quality effluents. Two processes involve tertiary treatment: passing biologically treated wastewater through carbon columns or adding powdered activated carbon to the activated sludge basin. Another method of using activated carbon is part of physical-chemical treatment that eliminates biological oxidation entirely; this process is described in the next subsection.

Reverse Osmosis. Filtration through semipermeable membranes under pressure removes biological and colloidal matter, as well as most dissolved organics that affect color, odor, and taste.

The municipal application for reverse osmosis is the desalting of brackish water having total dissolved solids between 1000 and 15,000 mg/L. Industrially, there are a wide variety of uses. The important design parameter is the rate of transport of water across the membrane, called the *flux J,* which is

$$J = C(\Delta P_g - \Delta P_o) \qquad (15.73)$$

where ΔP_g = physical (gage) pressure (feedside-productside)
$\quad \Delta P_o$ = osmotic pressure (feed-product)

Ion Exchange.   *Ion exchange* is a process in which ions, held by electrostatic force to charged functional groups on the surface of a solid, are exchanged for ions of a similar charge in a solution in which the solid is immersed. Because the charged functional groups at which exchange occurs are on the surface of the solid, and because the exchanging ions must undergo a phase transfer from liquid phase to residence on a solid phase, ion exchange is a sorption process.

10. Physical-Chemical Processes.   Physical-chemical treatment (involving coagulation, carbon adsorption, and filtration) is an alternative to biological wastewater treatment methods. It replaces microbial stabilization of the dissolved organic matter with a combination of physical treatment processes that do not rely on bacterial action. A typical flow diagram for physical-chemical treatment is shown in Fig. 15.19.

11. Disinfection.   Wastewater treatment processes remove some but not nearly all of the pathogens. Hence the discharge from biological treatment plants, as well as natural waters contaminated by human sources, can carry diseases caused by bacteria and viruses. Thus, disinfection of sewage plant effluents is widely practiced.

Chlorination.   Of all the chemical disinfectants, chlorine is perhaps most commonly used. It is reasonably economical, toxic to microorganisms that carry waterborne diseases, tasteless and nonpoisonous to humans at low concentrations, and can be detected in residual amounts some time after application.
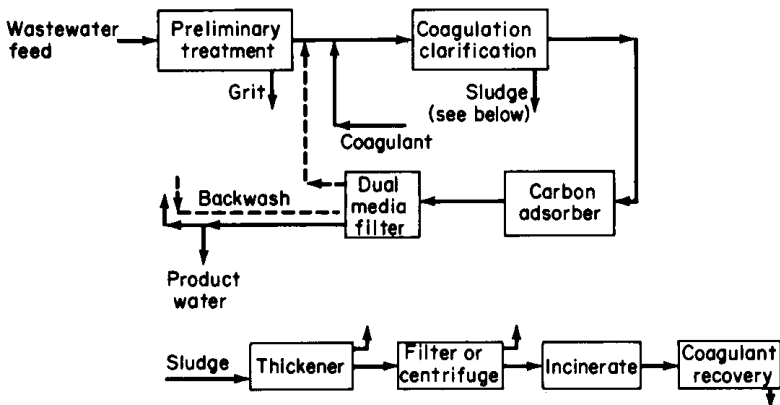
FIGURE 15.19   Principal operations in physical-chemical treatment.

Chlorine is quite soluble in water—7160 mg/L at 20°C and at atmospheric pressure, forming hypochlorite according to the following equation:

$$Cl_2 + H_2O \rightleftharpoons H\ O\ Cl + H^+ + Cl^- \qquad (15.74)$$
$$\uparrow \quad \downarrow$$
$$H^+ \quad OCl^-$$
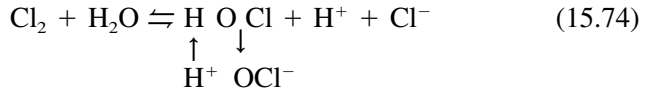
When ammonia is present, it reacts to form chloramines, which are slow-acting disinfectants:

$$NH_3 + HOCL \rightarrow NH_2Cl + H_2O \qquad (15.75)$$

$$NH_3 + 2HOCl \rightarrow NHCl_2 + 2H_2O \qquad (15.76)$$

$$NH_3 + 3HOCl \rightarrow NCl_3 + 3H_2O \qquad (15.77)$$
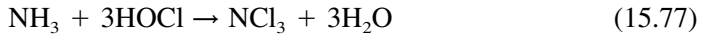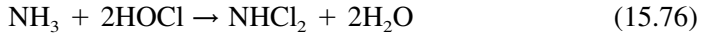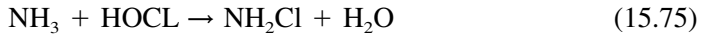
Chlorine, present as NH_2Cl or $NHCl_2$, is called *combined available chlorine.* When all reducing agents have been satisfied, then free chlorine is detected. This is called the *breakpoint,* which occurs only after the chloramines have been oxidized to N_2O and N_2.

Ozonization.   A relatively new disinfection method utilizes ozone, a molecule that consists of three atoms of elemental oxygen (O_3). In water, these atoms break down rapidly, so they are free to oxidize organic matter and inactivate bacteria and viruses. Additionally, ozonization adds oxygen to the system but not dissolved solids. Further, ozone will deodorize gases in a few seconds of contact time at low levels of concentration.

AIR POLLUTION CONTROL

12. Definitions.   *Air pollution* is defined as the "presence in the atmosphere of one or more contaminants of such quantity and duration as may be injurious to human, plant, or animal life, or property, or which may unreasonably interfere with comfortable enjoyment of life, property, or conduct of business." Air pollution results from a two-part phenomenon, a time-concentration relationship.

Classes of Pollutants.   Atmospheric pollutants may exist as gases or particulates. Frequently, a third category, odors, is recognized. Chemically, odors fall into one of the previous categories. They are frequently separated because of their objectionable olfactory sensation at concentrations where they are neither toxic nor hazardous.

Effects. SO_2 and other acidic pollutants corrode architecture and sculpture (both metals and masonry, especially limestone and marble). Deterioration of historic works has increased significantly in the last 50 years. Carbonaceous deposits discolor buildings. Particulates and condensation nuclei reduce atmospheric visibility, hindering aircraft and ground transportation, block off distant scenic views, and decrease the quantity of solar radiation reaching the earth. Pollutants damage vegetation and reduce crop yields.

13. Transport and Meteorological Effects.   Pollutants are transported through the atmosphere by wind currents from their point of release to downwind receptors. They are dispersed and diluted so that an emission, toxic at its release point, may be harmless at ground level downwind. The higher the release point above the surroundings and the more buoyant is the plume, the greater is the dilution.

Best conditions for dispersion are high wind speeds and a highly unstable temperature gradient (superadiabatic). As the atmosphere becomes progressively more stable, wind speeds tend to decrease, and plume dispersion becomes progressively poorer. An indication of atmospheric stability can be gained from smoke plume behavior (Fig. 15.20).

Dispersion Equations.   Standard gaussian distribution can predict downwind concentrations with fair accuracy. Models currently most in use are described by Turner.[26] Figure 15.21 illustrates the model. The plume is discharged from a stack of height Kh_s. A buoyant plume continues to rise some distance above the stack, Δh. The overall effective stack height H is the sum of h_s and Δh. The plume is assumed to be released from a point source upwind from the stack such that it
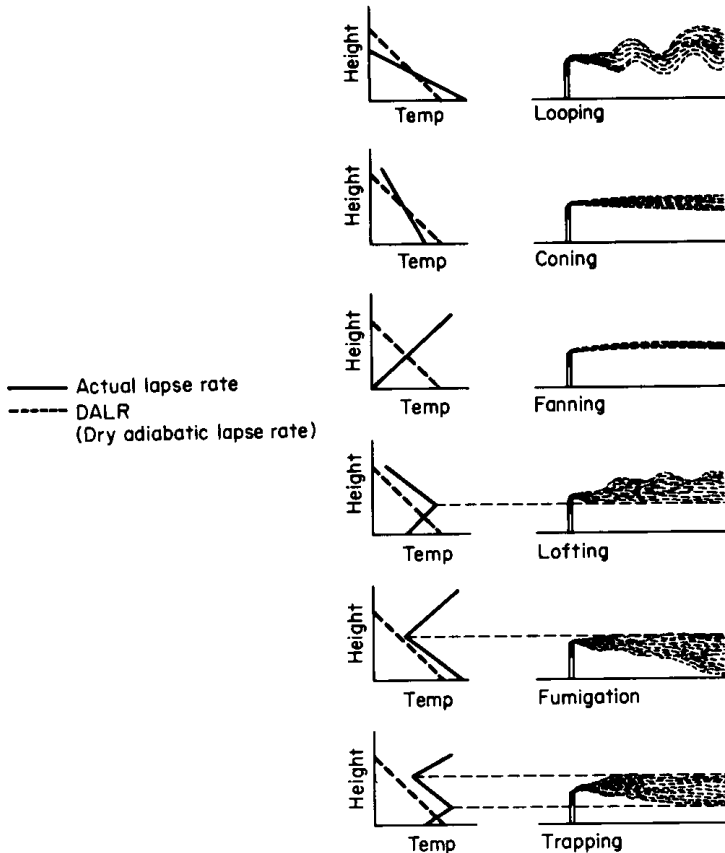
FIGURE 15.20   Smoke plume behavior as a function of temperature. (*From H. L. Perkins, Air Pollution, McGraw-Hill, New York, 1974, p. 169.*)
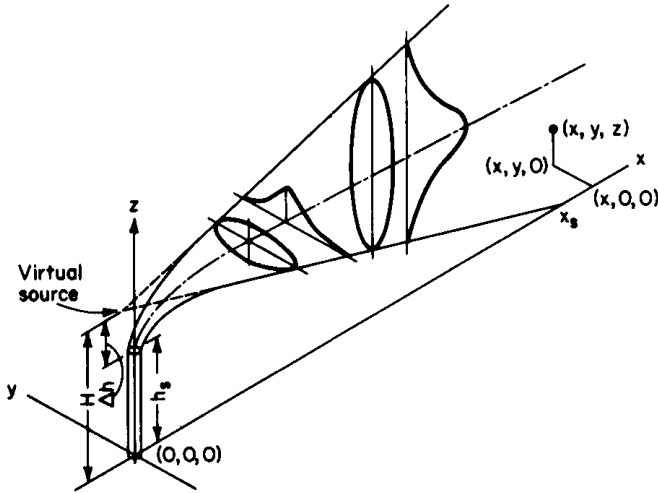
FIGURE 15.21   Gaussian model for dispersion equations.

spreads out in the shape of a cone as it travels downwind. The plume is free to spread horizontally about its axis and upward. However, at the point where the plume hits the ground, the assumption is made that the pollutants are reflected upward. This is handled mathematically by assuming a virtual image source beneath the ground that adds to the concentration downwind beyond the point x_s. Coordinates are x, the downwind direction with wind speed of $\overline{U}$; y, the crosswind direction measured from the plume centerline; and z, the vertical height above the ground.

The downwind concentration C at any point (x, y, z) can be computed from Eqs. (15.78) to (15.82). For downwind distances up to x_s,

$$\frac{C_{(x,y,z)}\overline{U}}{Q} = \frac{1}{2\pi\sigma_y\sigma_z}\exp\left\{-\frac{1}{2}\left[\left(\frac{y}{\sigma_y}\right)^2 + \left(\frac{z - H}{\sigma_z}\right)^2\right]\right\} \tag{15.78}$$

Beyond x_s,

$$\frac{C_{(x,y,z)}\overline{U}}{Q} = \frac{1}{2\pi\sigma_y\sigma_z}\exp\left[-\frac{1}{2}\left(\frac{y}{\sigma_y}\right)^2\right]$$
$$\left\{\exp\left[-\frac{1}{2}\left(\frac{z - H}{\sigma_z}\right)^2\right] + \exp\left[-\frac{1}{2}\left(\frac{z + H}{\sigma_z}\right)^2\right]\right\} \tag{15.79}$$

Beyond x_s, where ground-level concentrations are desired ($z = 0$),

$$\frac{C_{(x,y,0)}\overline{U}}{Q} = \frac{1}{\pi\sigma_y\sigma_z}\exp\left\{-\frac{1}{2}\left[\left(\frac{y}{\sigma_y}\right)^2 + \left(\frac{H}{\sigma_z}\right)^2\right]\right\} \tag{15.80}$$

If plume centerline concentrations only are desired ($y = 0$),

$$\frac{C_{(x,0,0)}\,\overline{U}}{Q} = \frac{1}{\pi\sigma_y\sigma_z} \exp\left[-\frac{1}{2}\left(\frac{H}{\sigma_z}\right)^2 \right] \qquad (15.81)$$

If the effluent is released at ground level ($H = 0$), the downwind plume centerline concentration is obtained by

$$\frac{C_{(x,0,0)}\,\overline{U}}{Q} = \frac{1}{\pi\sigma_y\sigma_z} \qquad (15.82)$$

Any consistent set of units may be used such as C, concentration in g/m³; $\overline{U}$, wind speed, m/s; Q, the pollutant release, g/s; and distances H, x, y, and z and diffusion coefficients σ_y and σ_z in meters.

For very accurate predictions of ground concentration, the values of σ_y and σ_z should be determined experimentally for the meteorological conditions and terrain under consideration. The model represents the real situation best in a neutral atmosphere (lapse rate close to the DALR).

There are six stability categories as defined in Table 15.3 with experimental values for σ_y and σ_z as shown in Figs. 15.22 and 15.23. These values of σ_y and σ_z are most accurate (± 100 percent) for level farm land. Ground structures, wooded areas, and rolling or mountainous terrain will result in larger actual departures from the calculated results. Figure 15.24 may be used with the stability categories of Table 15.3 to determine downwind plume centerline distance at which the maximum ground concentration will occur and its concentration.

The calculated concentrations are typical of those which would be obtained while sampling for 10 min. A sample averaged over a longer time period will give lower values due to normal fluctuations in wind direction. Equation (15.83) may be used to predict the average ground concentrations to be expected up to periods of several hours.

$$C_t = C_0 \left(\frac{t_0}{t_t}\right)^p \qquad (15.83)$$

TABLE 15.3   Stability Categories

Surface wind speed, m/sec	Day			Night	
	Incoming solar radiation			Thin overcast of ≥⅘ cloudiness	≤⅜ cloudiness
	Strong	Moderate	Slight		
<2	A	A–B	B		
2	A–B	B	C	E	F
4	B	B–C	C	D	E
6	C	C–D	D	D	D
>6	C	D	D	D	D

The neutral class, D, should be assumed for overcast conditions during day or night. A—extremely unstable conditions; B—moderately unstable conditions; C—slightly unstable conditions; D—neutral conditions; E—slightly stable conditions; F—moderately stable conditions.

Source: From Turner.[26]
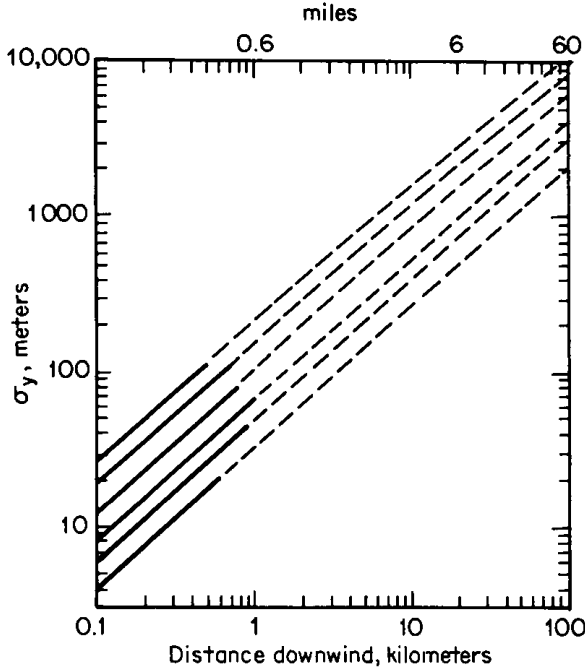
FIGURE 15.22 Horizontal dispersion coefficient as a function of downward distance from the source. Letters refer to stability category; see Table 15.3 (*From Turner.*[26])

where C_t = the concentration for a longer time period
$\quad\quad C_0$ = the concentration estimated for a 10-min period
t_0 and t_t = the respective time periods
$\quad\quad p$ = value between 0.17 and 0.20

The dispersion equations apply specifically to gases and suspended particulates. Owing to gravity fallout closed to the stack, they should not be applied to particulates larger than 20 μm.

Plume Rise.   The preceding dispersion equations require an estimate of the distance the plume continues to rise above the stack:

$$\Delta h = \frac{V_s d}{\overline{U}} \left[1.5 + 2.68 \times 10^{-3} \, pd \left(\frac{T_s - T_a}{T_a} \right) \right] \qquad (15.84)$$

where Δh = the plume rise, m
$\quad\quad V_s$ = stack discharge velocity, m/s
$\quad\quad d$ = stack diameter, m
$\quad\quad \overline{U}$ = wind speed, m/s
$\quad\quad p$ = atmospheric pressure, mbars
$\quad\quad T_s$ = stack discharge temperature, K
$\quad\quad T_a$ = ambient temperature, K

To correct plume rise for atmospheric stability, it is recommended that the value of
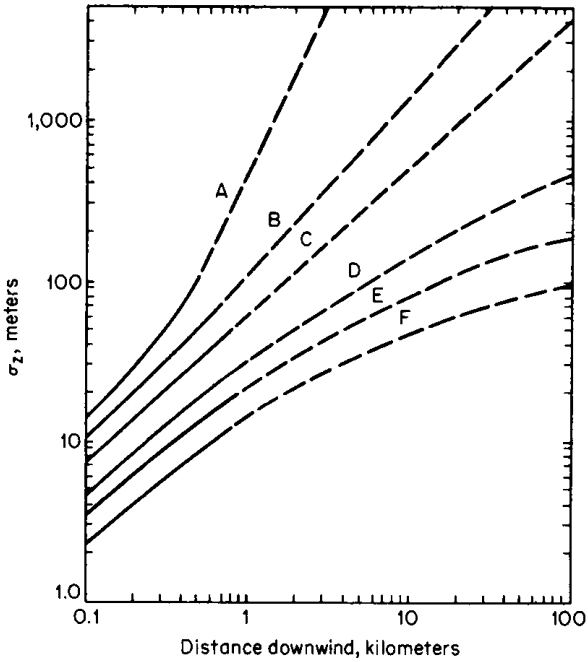
FIGURE 15.23 Vertical dispersion coefficient as a function of downwind distance from the source. Letters refer to stability category; see Table 15.3 (*From Turner.*[26])
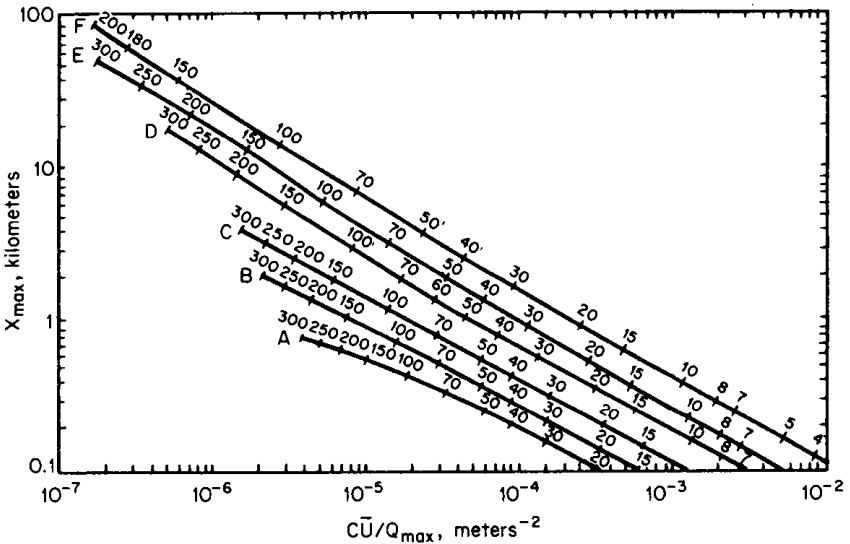
FIGURE 15.24 Distance of maximum concentration as a function of stability and effective height of emission. Letters indicate stability class; numbers, the effective stack height; see Table 15.3 (*From Turner.*[26])

Eq. (15.84) be multiplied by 1.15 for unstable to 0.85 for very stable conditions (1.0 for neutral).

Pollution Potential and Climatology.   When considering a location for an operation with a major air pollution potential, locations well sited to atmospheric dispersion should be considered. Flat, open spaces are preferred. Narrow mountain valleys should be avoided. Locations adjacent to large lakes and oceans can present special problems. A study by Hosler[27] gives indications of geographic areas with high and low potential for air pollution problems. It is desirable to pick an area with a mean mixing depth as great as possible.

14. Air Pollution Measurements.   Sampling may be occasional, intermittent, or infrequent to check on specific problems, or it may be conducted at regular intervals or even continuously to provide frequent monitoring. Positioning of samplers must be carefully considered. For a representative area sample, the intake should not be at a busy intersection, along a dusty road, or directly downwind of a major pollution source. Sampling for gases (and some particulates) is practiced by drawing a measured sample through a series of midget impingers. Water or a suitable solution absorbs the gaseous component.

Source Sampling.   There is a current tendency to use the U.S. EPA Method V sampler,[28] with modifications if needed, for many types of source sampling.
    Process location for sample withdrawal must be carefully selected to obtain a representative sample. The EPA test method V specifies, ''a composite sample shall be taken from at least 12 different locations in the same plane in a duct with each point the centroid of an equal area section.'' Under this circumstance, there should be at least 8 duct diameters of straight pipe upstream of the sampling point and 2 duct diameters downstream.

Velocity Measurement.   A measure of the desirability of a particular sampling location can be obtained from the velocity distribution across the duct. The more uniform the velocity distribution, the better is the chance that the pollutants will be somewhat uniformly distributed. Also, a knowledge of the velocity distribution is necessary in selection of the sample probe orifice, sampling flow rate, and total sampling time. The velocity distribution is usually determined in a round duct on two different diameters at right angles to each other using an S-type pitot tube. The velocity at each point is calculated with Eq. (15.85):

$$V_0 = 2.90(F) \sqrt{\left(\frac{29.92}{P}\right)\left(\frac{1.00}{G_d}\right)(\Delta h_p T_R)} \qquad (15.85)$$

where V_0 = local velocity, ft/s (m/s)
       P = absolute pressure in duct, inHg (mmHg)
       G_d = carrier gas specific gravity (air = 1)
       Δh_p = pitot differential pressure, inH$_2$O (mmH$_2$O)
       T_R = absolute temperature of carrier gas, °R

The value F in Eq. (15.85) is the discharge coefficient for the S-type pitot tube. This should be obtained for the velocity range being measured by calibration in a duct with clean air against a standard pitot tube having a discharge coefficient of unity. Recalibration of the S-type pitot should be performed at intervals, since the value of the discharge coefficient can change with abrasion and wear.

Sampling Technique.   After determining from the velocity traverse that the sampling location is suitable, the diameter of the sampling probe tip and the desired sampling rate must be selected. *Isokinetic sampling* means sampling with the same inlet velocity at the probe tip as the flowing gas stream has at the same site. Sampling isokinetically is necessary to obtain a representative sample of large particles.

Particle-Size Measurement.   Particle-size measurements are usually made either by laboratory analysis or by classification into various size ranges in a flowing gas sample. Where an unclassified sample is collected, particle size is usually measured by microscopic measurement or resuspension and classification in either an air stream or a liquid. With microscopic methods, a thin deposit of solids may be collected on a dry membrane filter or a glass slide using an electrostatic or thermal precipitator sampler. The sizes of particles are measured visually, and the number in each size range is counted. The particles are usually measured by equivalent area or by a predominant dimension. To obtain particle-size distribution on a mass basis, it is necessary to make assumptions about particle volume and density.

Monitoring.   For monitoring pollutant gases, a number of instruments have been developed using infrared and ultraviolet absorption. These are adaptable to SO_2, NO, CO, CO_2, hydrocarbons, and other specific gases. Membrane-type cells involving oxidation-reduction reactions with the pollutant also have been developed.

There is a need to predict plume opacity for compliance with regulations when certain particulate collection efficiencies are anticipated or specified. Ensor and Pilat[29] present a theoretical equation for relating plume opacity and particle concentration and properties:

$$W = -\frac{K\rho \ln (I/I_0)}{L} \tag{15.86}$$

where W = the particle mass concentration, g/m^3
ρ = true particle density, g/cm^3
I/I_0 = ratio of light transmitted through plume to quantity transmitted if there were no emission [opacity = $1 - (I/I_0)$]
L = length of light path, m
K = a proportionality constant, cm^3/m^2, and dependent on particle properties and light wavelength

This equation is most accurate for a process in which particle-size distribution is constant and when the value of K has been determined experimentally by measuring opacity and particle concentration simultaneously. However, predictions of plume opacity can be made with fair accuracy from fundamental data on the mass-median particle size, the standard particle-size deviation, the density of the particles, and their refractive index. Theoretical values of K are given in Ensor and Pilat.[29]

15. Pollution Control Techniques

Minimizing Need for Collection Devices.   There are several means of reducing or eliminating pollutants without specific removal devices. Examples are substitution of low-sulfur fuels, nonvolatile solvents, operation at lower temperatures to reduce NO_x formation or volatilization of a processed material; use of indirect rather than direct-contact heat transfer; heat transfer from radiant panels; pretreatment of nat-

ural raw materials to remove easily airborne fines or volatiles; and wetting or agglomeration of solids to reduce airborne releases during handling.

Application of Control Devices. Table 15.4 lists equipment types useful for reducing emission of gases, odors, and particulates. Some odors can be canceled by other antagonistic odor molecules. This provides a means of eliminating an odor without actually preventing its release.

16. Collection and Removal of Gases

Absorption. This is one of the most frequently used methods for removal of water-soluble gases. Acidic gases (such as HCl, HF, and SIF_4) can readily be absorbed in water efficiently, especially if contact is with water having an alkaline pH. Less-soluble acidic gases (such as SO_2, Cl_2, and H_2S) can be absorbed more readily in a dilute caustic solution such as 5 to 10 percent NaOH. Scrubbing with an ammonium salt solution is also employed; the gas is often contacted with the more alkaline solution first and a neutral or slightly acid solution last to prevent loss of NH_3 to the atmosphere. Lime is an inexpensive alkali but often leads to plugging problems in absorption equipment if the calcium salts have only limited solubility. A better technique is to absorb with an NaOH solution, which is then limed external to the absorption tower. The calcium salts are settled out and the regenerated NaOH returned to the absorption system.

Adsorption. Adsorption processes consist of contacting a gas with a solid. The solids are essentially porous with an affinity for certain substances. Typical materials are activated carbon, activated alumina, silica gels, and molecular sieves. An adsorbent can hold from 8 to 25 percent of its own weight in adsorbed vapors.

TABLE 15.4 Types of Equipment Applicable to the Control of Various Classes of Air Pollutants

| | Pollutant classification | | | |
| | | | Particulate | |
Equipment type	Gas	Odor	Liquid	Solid
Absorption....................	X	X		
Adsorption....................	X	X		
Air dispersion (stacks)...........	X	X	X	X
Condensation	X	X		
Centrifugal (dry)................			X	X
Filtration, bags				X
beds			X	X
fine fibers			X	X
Gravitational settling				X
Impingement (dry)..............			X	X
Incineration....................	X	X	X	X
Precipitation, electric			X	X
thermal............			X	X
Wet collection..................			X	X

Generally, the adsorbate is held in liquid phase, even though physical principles would predict that its physical state should be a vapor. The adsorbate may be held in the pore structure by direct physical attraction or by the formation of chemical bonds.

Condensation.   A number of vapors, especially hydrocarbons with low volatility, can be recovered by condensation. A tubular, water-cooled heat exchanger is adequate for many high-molecular-weight organic vapors. Where the volatility is greater, a refrigerated condenser following the water-cooled condenser may be necessary. In condensing many vapors, where heat transfer is more rapid than mass transfer, fog particles of the condensate (0.5- to 1.5-μm diameter) are apt to form in the bulk gas stream. These particles can result in plume opacity violations as well as a recovery from cooling that is not as great as predicted from vapor pressure-temperature data.

17. Special Gaseous Pollutant Control Methods.   Two widely released gaseous pollutants are SO_2 and NO_x. Major sources of these pollutants are flue gases from combustion operations. Control of these two gases has received wide study, and a number of specialized techniques are available.

Sulfur Dioxide.   One control technique is to substitute a low-sulfur or desulfurized fuel. Desulfurization has been commercialized for petroleum fuels. Several processes have been demonstrated for coal, but it may be several years before they are commercial.

Nitrogen Oxides.   Combustion operations are a major source of NO_x pollutants. O_2 and N_2 from air react at high temperatures in the flame to produce NO. NO reacts more slowly at lower temperatures with O_2 to produce NO_2. In most furnaces, reaction rates to form NO are too slow to producer equilibrium amounts of NO corresponding to flame temperature, but it is not unusual for the flue gases from oil and coal combustion to contain 1000 to 2000 ppm by volume of NO when using 5 to 10 percent excess air.

Control of Pollutants by Incineration.   Incineration is used to destroy combustible vapors such as hydrocarbons (especially unsaturated and aromatic compounds that are photochemically reactive), CO, H_2, H_2S, and mercaptans.

18. Collection and Removal of Particulates.   Six basic principles are used alone or in combination in particulate collectors:

1. Gravity settling
2. Flowline interception
3. Inertial deposition
4. Diffusional deposition
5. Electrostatic deposition
6. Thermal precipitation

Table 15.5 lists these mechanisms and their basic parameters. In addition, sonic agglomeration has been considered but has seldom become commercially practical.

TABLE 15.5 Summary of Mechanisms and Parameters in Aerosol Deposition

Deposition mechanism	Origin of force field	Deposition mechanism measurable in terms of		System parameters
		Basic parameter	Specific modifying parameters	
Flow-line interception*	Physical gradient*	$N_{sf} = \left(\dfrac{D_p}{D_b}\right)$		Geometry (D_{b1}/D_b), (D_{b2}/D_b), etc. ϵ_v α
Inertial deposition	Velocity gradient	$N_{si} = \left(\dfrac{K_m \rho_p D_p^2 V_o}{18\mu D_b}\right)$	$N_{sc} = \left(\dfrac{N_{sf}^2}{N_{si} N_{sd}}\right)$ $= \left(\dfrac{18\mu}{K_m \rho_p D_v}\right)^{\ddagger}$	
Diffusional deposition	Concentration gradient	$N_{sd} = \left(\dfrac{D_v}{V_o D_b}\right)$		
Gravity settling	Elevation gradient	$N_{sg} = \left(\dfrac{u_t}{V_o}\right)$		
Electrostatic precipitation / Attraction / Induction	Electric-field gradient†	$N_{sec} = \left(\dfrac{K_m Q_p \epsilon b}{\mu D_p V_o}\right)$ $N_{sei} = \left(\dfrac{\delta_p - 1}{\delta_p + 2}\right)\left(\dfrac{K_m D_p^2 \delta_o \epsilon b^2}{\mu D_b V_o}\right)$	δ_p, δ_b§	Flow pattern: N_{Re}¶ N_{Ma} N_{Kn} Surface accommodation
Thermal precipitation	Temperature gradient	$N_{st} = \left(\dfrac{T - T_b}{T}\right)\left(\dfrac{\mu}{K_m \rho D_b V_o}\right)$ $\left(\dfrac{k_t}{2k_t + k_{tp}}\right)$	(T_b/T), (T_p/T),§ (N_{Pr}) (k_{tp}/k_t), (k_{tb}/k_t),§ (c_{hp}/c_t), (c_{hb}/c_h)§	

§ Not likely to be significant contributors.
*This has also commonly been termed "direct interception" and in conventional analysis would constitute a physical boundary condition imposed upon particle path induced by action of other forces. By itself it reflects deposition that might result with a hypothetical particle having finite size but no mass or elasticity.
†In cases where the body charge distribution is fixed and known, ϵ_b may be replaced with Q_{bs}/δ_o.
‡This parameter is an alternative to N_{sf}, N_{si}, or N_{sd} and is useful as a measure of the interactive effect of one of these on the other two. It is comparable with the Schmidt number.
¶When applied to the inertial deposition mechanism, a convenient alternative is $(K_m \rho_s/18\rho) = N_{si}/(N_{sf}^2 N_{Re})$.

Source: Perry and Chilton (eds.), *Chemical Engineers' Handbook*, 5th ed., McGraw-Hill, New York, 1973, p. 20–80.

Gravity Settling Chambers.   Owing to low collection efficiency on smaller particles, gravity settling is seldom useful today as a sole collection device. However, such chambers may be used as preclassifiers to remove large abrasive particles ahead of another type of collector. Settling chambers are seldom efficient on particles smaller than 50 μm.

Basically, a *settling chamber* is a large horizontal enlargement in the duct where the gas slows down. Turbulence should be low to prevent reentrainment. The collection efficiency for any size particle can be calculated from the terminal settling velocity of the particle as determined in Fig. 15.25, the distance the particle has to fall to settle out, and the residence time in the chamber. Equation (15.87) can be applied to each individual size particle to calculate collection efficiency for that size. Equation (15.88) gives the smallest size particle that can be collected with 100 percent efficiency.

$$\eta = \frac{u_t L_s}{H_s V_s} \tag{15.87}$$

$$D_{p,\min} = \sqrt{\frac{18\mu H_s V_s}{g_L L_s(\rho_s - \rho_g)}} \tag{15.88}$$

where η = fractional collection efficiency
$\quad u_t$ = particle terminal settling velocity, ft/s (m/s)
$\quad V_s$ = bulk gas velocity, ft/s (m/s)
L_s and H_s = the length and height of the settling path, respectively, ft (m)
$\quad \mu$ = gas viscosity, lb/(s · ft) [kg/(s · m)]
$\quad g_L$ = local acceleration due to gravity, ft/s² (m/s²)
$(\rho_s - \rho_g)$ = difference in particle and gas density, lb/ft³ (kg/m³)

Cyclonic Collectors.   In cyclonic collectors, centrifugal force separates particles from the gas stream. Since centrifugal force can equal many times that of gravity, much smaller particles can be collected. Figure 15.26 shows a typical dust-collection cyclone in which the gas enters tangentially, spirals downward, reverses direction in the cone, and exits through the top in smaller spirals. The dust particles spiral downward along the wall and discharge at the bottom.

The efficiency of a cyclone is computed by integration of the collection efficiency for each individual size particle as obtained from the manufacturer's efficiency curve. It is frequent practice to calculate the particle cut size D_{pc} (the diameter particle that is collected with 50 percent efficiency). Equation (15.89) gives the cut size for a cyclone with dimensions as given in Fig. 15.26.

$$D_{pc} = \sqrt{\frac{9\mu B_c}{2\pi N_c V_c(\rho_s - \rho_g)}} \tag{15.89}$$

where N_c = the number of turns the gas makes in the cyclone, often 5 to 10
$\quad B_c$ = the width of the gas inlet, ft (m)
$\quad V_c$ = inlet gas velocity, ft/s (m/s)

All other symbols are as defined for Eq. (15.88).

Pressure drop through a cyclone depends on cyclone geometry. The manufacturer's calibration curve should be used. The pressure drop of the Fig. 15.26 cyclone
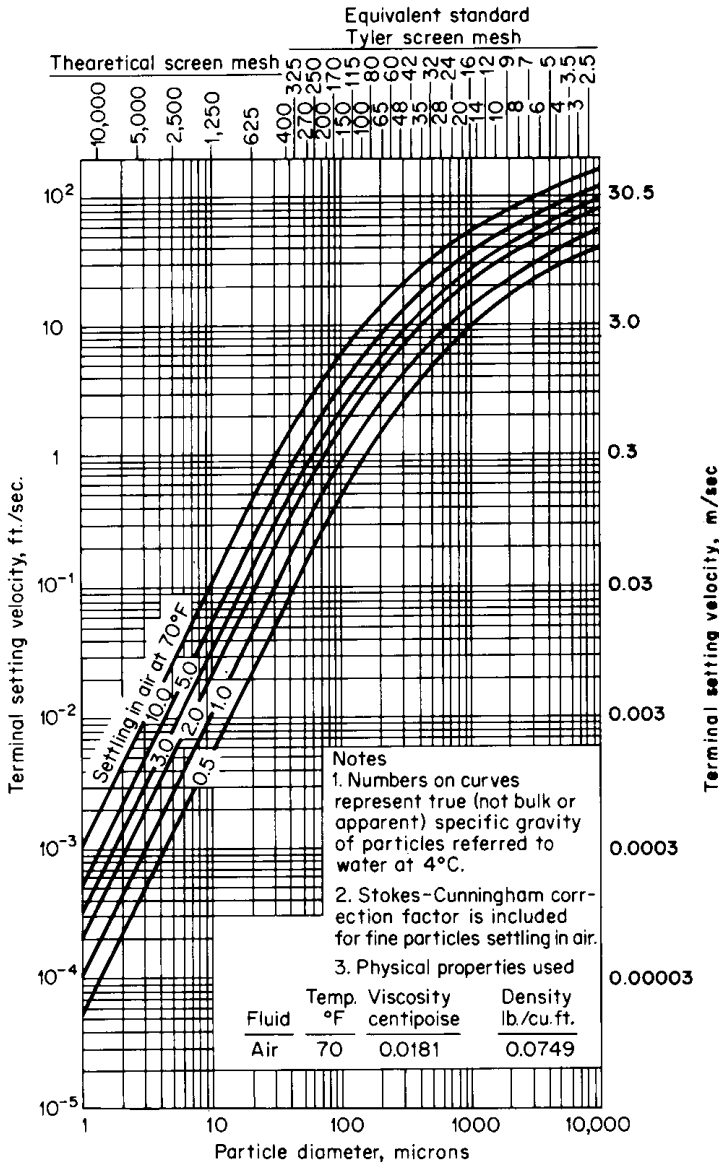
FIGURE 15.25   Terminal settling velocity of spherial particles.

in terms of the number of inlet velocity heads, F_{cv}, is given by Eq. (15.90). The value of K for this cyclone is 16. Other symbols are defined by Fig. 15.26.

$$F_{cv} = KB_cH_c/D_c^2 \qquad (15.90)$$

Bag Filters.   Bag filters for the collection of particulates use a woven fabric or a nonwoven felt bag. Either can be specified for collection efficiencies of 99 percent or better. The collection process is not merely a screening or filtration of the dust, since the openings in the cloth are many times the size of the dust particles col-

$B_c = D_c/4$
$D_e = D_c/2$
$H_c = D_c/2$
$L_c = 2 D_c$
$S_c = D_c/8$
$Z_c = 2 D_c$
$J_c = $ arbitrary,
    usually $D_c/4$

Section A-A
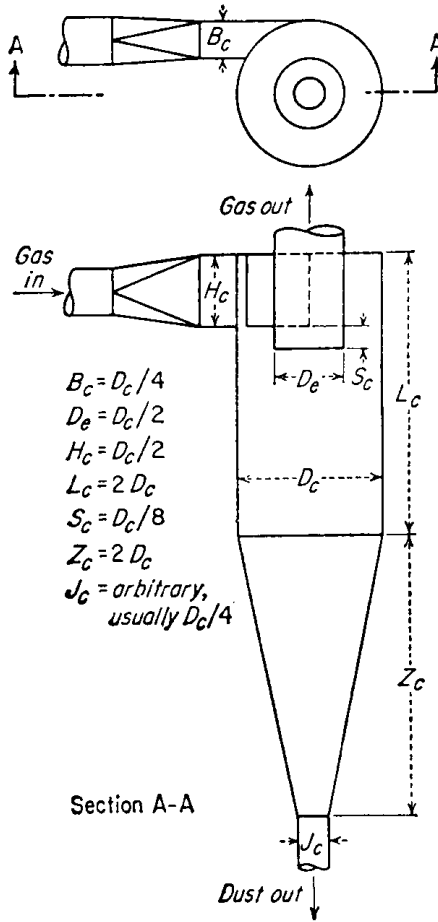
FIGURE 15.26 Cyclone separator proportions.

lected. Efficiency of a new bag may be fairly low for a few moments until it develops a precoat of dust that serves as the filtering layer. Once precoated, the bag usually retains sufficient solids in its pores so that it does not return to its original efficiency.

Pressure drop through the fabric is usually negligible compared to that through the layer of collected dust. Pressure drop through the dust layer can be expressed by

$$\Delta p_i = K_d \mu C_d V_f^2 \theta \tag{15.91}$$

where C_d = the concentration of dust in the gas stream
V_f = the gas velocity through the bag
μ = the gas viscosity
θ = the time since the last bag cleaning cycle

The resistance factor K_d is a function of the dust size, shape, packing density, etc. and is best determined experimentally. Usual practice is to specify a maximum

desired pressure drop across the bag and pick a cleaning cycle such that this pressure drop is not exceeded. Typical operating pressure drops are 2 to 6 inH$_2$O (50.8 to 152 mmH$_2$O).

Electrostatic Precipitators. An electrostatic precipitator collects solid or liquid particulates with high efficiency and low energy utilization. The pressure drop is composed almost entirely of inlet and exit losses. Initial cost of precipitators is high, however, and their on-stream reliability may be less than that of other collection devices.

Equation (15.92) gives the theoretical efficiency for a precipitator:

$$\eta = 1 - e^{-[(u_e A_e)/q]} = 1 - e^{-K_e u_e} \tag{15.92}$$

where η = the weight fractional efficiency
   u_e = the velocity of migration of particle toward collecting electrode
   A_e = the area of collecting electrodes
   q = the total gas flow rate

Any consistent set of units that makes the exponent dimensionless can be used. The efficiency factor K_e is basically a geometric design factor for the precipitator. For a plate-type precipitator, $K_e = L/B_e V_e$. For a tube type, $K_e = 4L/D_t V_e$, where L is the length of a flow passage in the electric field, B_e is the distance between adjacent plates, D_t is the tube diameter, and V_e is the gas velocity through an individual flow passage. The migrational velocity u_c is a function of the particle size, shape, composition, and average electric field strength. Each different size particle has a different migration velocity. For a dust of a given particle-size distribution, an average migrational velocity is often used.

With a given precipitator, it is often desired to know the effect of changes in operating conditions. Using Eq. (15.92) with geometric values for K_e, the effects of changes in gas velocity or residence time can be determined without any knowledge of the value of u_e. However, changing gas temperature or particle size will change the value of u_e. For a constant field strength, the relative changes in u_e can be estimates by Eqs. (15.93) and (15.94).

For a change in temperature,

$$(u_e)_{t_2} = \frac{(u_e)_{t_1}(\mu_g)_{t_1}}{(\mu_g)_{t_2}} \tag{15.93}$$

where μ_g is the gas velocity at temperatures t_1 and t_2.

For a change in median particle size at constant particle standard deviation,

$$u_{e_2} = \frac{u_{e_1} r_1}{r_2} \tag{15.94}$$

where r_1 and r_2 are the radius of the median particle sizes.

Mist Filters. Mists are created by condensation. They consist of small particles, from submicrometer size up to 10 μm. For coarser mists, 5 μm and up, knitted-wire-mesh separator pads, 4 to 6 in (10.2 to 15.2 cm) thick, have collection efficiencies of 98 percent and above with 1 to 2 in (2.5 to 5.1 cm) of water-pressure drop. These separators are most efficient when installed horizontally with upflow of gas, and gravity drainage of liquid. Collection is by impingement.

Wet Scrubbers.   A great variety of liquid scrubbing devices are available for collection of particulates. The major collection mechanism is inertial impaction with direct interception and Brownian diffusion of lesser importance. Some units contact the gas with atomized liquids using either gravity or centrifugal force. Others impinge the gas at high velocity against sheets and films of water, causing the water to shatter into droplets. In either case, the major portion of the collection occurs by collision between small liquid droplets and the particulate.

PETROLEUM AND GAS ENGINEERING*

Petroleum engineering is a broad-based discipline comprising the technologies used for the exploitation of crude oil and natural gas reservoirs. It is usually subdivided into the branches of petrophysical, geological, reservoir drilling, production, and construction engineering. After an oil or gas accumulation is discovered, technical supervision of the reservoir is transferred to the petroleum engineering group, although in the exploration phase the drilling and petrophysical engineers have played a role in the completion and evaluation of the discovery.

PETROPHYSICAL ENGINEERING

The petrophysical engineer is perhaps the first of the petroleum engineering group to become involved in the exploitation of the new discovery. By the use of downhole logging tools and of laboratory analysis of cores made during the drilling operation, the petrophysical estimates the porosity, permeability, and oil content of the reservoir rock that has been sampled at the drill site.

GEOLOGICAL ENGINEERING

The geological engineer, using the petrophysical data, the seismic surveys conducted during the exploration operations, and an analysis of the regional and environmental geology, develops inferences concerning the lateral continuity and extent of the reservoir. However, this assessment usually cannot be verified until additional wells are drilled and the geological and petrophysical analyses are combined to produce a firm diagnostic concept of the size of the reservoir, the distribution of fluids therein, and the nature of the natural producing mechanism. As the understanding of the reservoir develops with continued drilling and production, the geological engineer, working with the reservoir engineer, selects additional drill sites to further develop and optimize the economic production of oil and gas.

---

*Excerpted from the *McGraw-Hill Encyclopedia of Science and Technology,* 8th ed. Copyright © 2003. Used by permission of McGraw-Hill, Inc. All rights reserved.

RESERVOIR ENGINEERING

The reservoir engineer, using the initial studies of the petrophysicist and geological engineers together with the early performance of the wells drilled into the reservoir, attempts to assess the producing rates [barrels of oil or millions of cubic feet (cubic meters) of gas per day] that individual wells and the entire reservoir are capable of sustaining. One of the major assignments of the reservoir engineer is to estimate the ultimate production that can be anticipated from both primary and enhanced recovery from the reservoir. The *ultimate production* is the total amount of oil and gas that can be secured from the reservoir until the economic limit is reached. The *economic limit* represents that production rate which is just capable of generating sufficient revenue to offset the cost of operating the reservoir. The proved reserves of a reservoir are calculated by subtracting from the ultimate recovery of the reservoir (which can be anticipated using available technology and current economics) the amount of oil or gas that has already been produced.

Primary recovery operations are those which produce oil and gas without the use of external energy except for that required to drill and complete the wells and lift the fluids to the surface (pumping). *Enhanced recovery,* or *supplemental recovery,* is the amount of oil that can be recovered over and above that producible by primary operation by the implementation of schemes that require the input of significant quantities of energy. In modern times, waterflooding has been almost exclusively the supplementary method used to recover additional quantities of crude oil. However, with the realization that the discovery of new petroleum resources will become an increasingly difficult achievement in the future, the reservoir engineer has been concerned with other enhanced oil recovery processes that promise to increase the recovery efficiency above the average 33 percent experienced in the United States (which is somewhat above that achieved in the rest of the world). The restrictive factor on such processes is the economic cost of their implementation.

DRILLING ENGINEERING

The drilling engineer has the responsibility for the efficient penetration of the earth by a well bore and for cementing of the steel casing from the surface to a depth usually just above the target reservoir. The drilling engineer or another specialist, the mud engineer, is in charge of the fluid that is continuously circulated through the drill pipe and back up to surface in the annulus between the drill pipe and the bore hole. This must be formulated so that it can do the following: carry the drill cuttings in suspension if circulation stops; from a filter cake over porous low-pressure intervals of the earth, thus preventing undue fluid loss; and exert sufficient pressure on any gas- or oil-bearing formation so that the fluids do not flow into the well bore prematurely, blowing out at the surface. As drilling has gone deeper and deeper into the earth in the search for additional supplies of oil and gas, higher and higher pressure formations have been encountered. This has required the use of positive-acting blowout preventers that can firmly and quickly shut off uncontrolled flow due to inadvertent imbalances in the mud system.

PRODUCTION ENGINEERING

The production engineer, upon consultation with the petrophysical and reservoir engineers, plans the completion procedure for the well. This involves a choice of setting a liner across the formation of perforating a casing that has been extended and cemented across the reservoir, selecting appropriate pumping techniques, and choosing the surface collection, dehydration, and storage facilities. The production engineer also compares the productivity index of the well [barrels per day per pounds per square inch (cubic meters per pascal) of drawdown around the well bore] with that anticipated from the measured and inferred values of permeability, porosity, and reservoir pressure to determine whether the well has been damaged by the completion procedure. Such comparisons can be supplemented by a knowledge of the rate at which the pressure builds up at the well bore when the well is abruptly shut-in. Using the principles of unsteady-state flow, the reservoir engineer can evaluate such a buildup to assess quantitatively the nature and extent of well-borne damage. Damaged wells, like wells of low innate productivity, can be stimulated by acidization, hydraulic fracturing, additional performation, or washing with selective solvents and aqueous fluids.

CONSTRUCTION ENGINEERING

Major construction projects, such as the design and erection of offshore platforms, require the addition of civil engineers to the staff of petroleum engineering departments, and the design and implementation of natural gasoline and gas processing plants require the addition of chemical engineers.

GAS FIELD AND GAS WELL

Petroleum gas, one form of naturally occurring hydrocarbons of petroleum, is produced from wells that penetrate subterranean petroleum reservoirs of several kinds. Oil and gas production are commonly intimately related, and about one-third of gross gas production is reported as derived from wells classed as oil wells. If gas is produced without oil, production is generally simplified, in part at least because the gas flows naturally without lifting, and also because of fewer complications in reservoir problems. As for all petroleum hydrocarbons, the term *field* designates an area underlain with little interruption by one or more reservoirs of commercially valuable gas.

PETROLEUM ENHANCED RECOVERY

Novel technology has been designed to enhance the fraction of the original oil in place in a reservoir. Heightened interest in developing enhanced recovery technology has developed as it has become more certain that over two-thirds of the oil

discovered in the United States, and a still greater percentage in the rest of the world, will remain unrecovered through the application of conventional primary and secondary (waterflood) operations. Thus there is a strong incentive for the development and implementation of advanced technology to recover some of this oil.

REFERENCES*

1. R. H. Perry (ed.), *Engineering Manual,* 3d ed., McGraw-Hill, New York. 1976.

2. *International Critical Tables.*

3. R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Perry's Chemical Engineers' Handbook,* 6th ed., McGraw-Hill, New York, 1984.

4. *Journal of Chemical and Engineering Data.*

5. R. E. Treybal, *Mass Transfer Operations,* 2d ed., McGraw-Hill, New York, 1968.

6. Hala, Wichterle, et al., *Vapor Liquid Equilibrium Data at Normal Pressures,* Pergamon Press, New York, 1968.

7. Smith, Block, and Hickman, in R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Perry's Chemical Engineers' Handbook,* 6th ed., McGraw-Hill, New York, 1984.

8. Hala, Pick, Fried, and Vilim, *Vapor-Liquid Equilibrium,* 2nd ed. Pergamon Press, New York, 1967.

9. Fair, et al., in R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Perry's Chemical Engineers' Handbook,* 6th ed., McGraw-Hill, New York, 1984.

10. Maddox, in R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Perry's Chemical Engineers' Handbook,* 6th ed. McGraw-Hill, New York, 1984.

11. Ellis, *Industrial Chemistry* 28: 483, 1952.

12. Sherwood and Pigford, *Absorption and Extraction,* McGraw-Hill, New York, 1952.

13. Norris, in R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Perry's Chemical Engineers' Handbook,* 6th ed., McGraw-Hill, New York, 1984.

14. McCormick, in R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Perry's Chemical Engineers' Handbook,* 6th ed., McGraw-Hill, New York, 1984.

15. McCabe and Smith, *Unit Operations of Chemical Engineering,* 3d ed., McGraw-Hill, New York, 1976, pp. 454–459.

16. Vermeulen et al., in R. H. Perry, D. W. Green, and J. O. Maloney (eds.), *Perry's Chemical Engineers' Handbook,* 6th ed., McGraw-Hill, New York, 1984.

17. J. J. Collins, *Chemical Engineers' Progress Symposium Service* 63(74):31, 1967.

18. Levenspiel, *Chemical Reaction Engineering,* Wiley, New York, 1962.

19. Walas, *Reaction Kinetics for Chemical Engineers,* McGraw-Hill, New York, 1959.

20. Aris, *The Optimal Design of Chemical Reactors,* Academic Press, New York, 1961.

21. Eckenfelder, "Trickling Filtration Design and Performance," *Journal of the Sanitation Engineering Division, American Society of Civil Engineers* 87(SA4): 33–45, 1961.

22. Germain, "Economic Treatment of Domestic Waste by Plastic-Medium Trickling Filters," *Journal of Water Pollution Control Federation* 38: 192–203, 1966.

*Publication dates given for cited references are for editions containing significant technical breakthroughs for Chapter 15.

23. R. S. Burd, *A Study of Sludge Disposal and Handling* (Publication No. WP-20-4), U.S. Department of the Interior, Federal Water Pollution Control Administration, Washington, D.C., May 1968.

24. U.S. Public Health Service, Publications AP-49, *Particulates,* 1969; AP-50, SO_x, 1969; AP-62, *CO*, 1970, Washington, D.C.

25. U.S. Environmental Protection Agency, Publications AP-63, *Oxidants,* 1970; AP-64, *Hydrocarbons,* 1970; AP-84, NO_x, 1971.

26. D. B. Turner, *Workbook of Atmospheric Dispersion Estimates* (Publication No. AP-26), U.S. Environmental Protection Agency, Washington, D.C., 1970.

27. Hosler, "Low Level Inversion Frequency in the Contiguous US," *Monthly Weather Review* 89: 319–339, 1961.

28. U.S. Environmental Protection Agency, "Method V Sampler," *Federal Register* 36: 24893, 1971.

29. Ensor and Pilat, *Journal of the Air Pollution Control Association* 21: 496–501, 1971.

30. *McGraw-Hill Encyclopedia of Science and Technology,* 6th ed. McGraw-Hill, New York, 2003.

CHAPTER 16

ELECTRICAL ENGINEERING*

BASIC ELECTRICAL DEVICES AND THEIR SYMBOLS

The following symbols are used to represent the basic linear electrical components:

AC Power Source Symbol   Momentary Switch Symbols

Compent Reference
Designator = "SW"
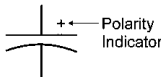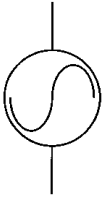
AC Frequency and
Amplitude to be
specified

Momentary On   Momentary Off

Circuit Closed when   Circuit Open when
Button Pressed   Button Pressed

ELECTRIC CIRCUITS AND THEIR CHARACTERISTICS

1. Electric Charge _Q_ and Current _I_.   The entities of electricity are called *electrons*. Each electron carries a *charge* $Q = 1.59 \times 10^{-19}$ coulomb. When electrons are set in motion, they produce an *electric current* denoted by the letter I or i. Electric current is measured in *amperes* and is equal to the *charge* in coulombs passing a given point per second, or

$$I = \frac{dQ}{dt} \tag{16.1}$$

The symbolic notation for current is an arrow →.

2. Electromotive Force emf or _E_.   The agency which can set electrons in motion is called *electromotive force* (abbreviated emf and denoted by the letter E or e). The unit of electromotive force is the volt. There are several sources of emf, of which the following may be mentioned:

1. The motion of a metallic body in a magnetic field. (This is the agency which creates emf in rotating electrical machines.)
2. The change in the value of a magnetic field in the neighborhood of a metallic body. (This is the agency which creates emf in transformers.)
3. Chemical reactions. (These are the cause of emf in storage batteries.)
4. Light. (This is the cause of emf in photoelectric cells.)
5. Heat. (This is the cause of emf in a junction of two unequally heated metals.)
6. Pressure. (This is the cause of emf in certain crystals subject to mechanical pressure.)

In electrical engineering, it is customary to distinguish between two types of emf sources, as follows:

Direct (*continuous* or *unidirectional*) emf, which continues in the same direction. A symbolic notation for direct emf is ⊣⊢ with the light vertical line representing the positive terminal and the small block representing the negative terminal.

Alternating emf, which varies periodically, and the mean value of which during each period is zero. A symbolic notation for alternating emf is a short wavy line: ∼.

3. Circuits and Their Characteristics. Any arrangement in which electric current may flow is called a *circuit*. By definition, a circuit may contain an emf source to set the electrons in motion and a body through which the current may flow. The body through which current flows is known as the *characteristic* (or *parameter*) of the circuit.

4. Types of Circuit Characteristics. There are three fundamental types of circuit characteristics: (1) resistance, (2) inductance, and (3) capacitance.

5. Electric Resistance (*R* or *r*). With the present meager knowledge of the nature of electricity, it is not possible to explain the physical nature of electric resistance. It can be defined, however, through certain empirical relations. These are (1) Ohm's law and (2) Joule's law.

Ohm's Law. If an unvarying emf (Fig. 16.1) is impressed on a homogeneous metallic conductor, and if the temperature of that conductor is held constant, then the ratio of the current I in the conductor to the impressed emf E is constant. This constant ratio may be expressed in two ways:

$$R = \frac{E}{I} \tag{16.2a}$$

$$G = \frac{I}{E} \tag{16.2b}$$

R is known as the *resistance* of the conductor, and $G = 1/R$ is called its *conductance.* The unit of resistance is the ohm. The unit of conductance is the mho.

Equations (16.2) give a mathematical statement of Ohm's law and incidentally define resistance and conductance. The following limitations of Ohm's law must be emphasized:

1. Ohm's law holds true only for homogeneous metallic conductors. It fails to be applicable, for example, in the case of liquid and gaseous conductors, of conducting dielectric materials, or of two dissimilar metals joined or welded together.
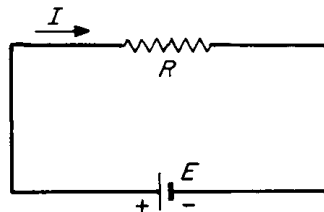
FIGURE 16.1

2. Ohm's law applies only when the source of emf is constant or when no local emf's exist in the circuit.

3. Experiment shows that the resistance of a metallic conductor generally increases with an increase in temperature. Thus, if R_t is the resistance at a temperature $t°C$, and if R_c is the resistance at 0°C, then

$$R_t = R_c(1 + \alpha) \tag{16.3}$$

where α is a positive constant which depends upon the metal and the units used and is known as the *temperature coefficient* of the metal. Temperature coefficients for various materials are given in Tables 16.1 and 16.3.

Joule's Law. The flow of current through a resistance is always accompanied by heat. Joule established the experimental fact that the amount of energy W dissipated in a metallic wire is proportional to the product of the square of the current I and the time t during which that current flows. Moreover, the coefficient of proportionality between W and I^2t is the resistance R of the circuit, or

$$W = RI^2t \tag{16.4}$$

This gives a definition of resistance from an energy viewpoint. Resistance is that characteristic of a circuit which accounts for the existence of heat. An electric circuit is devoid of resistance when the flow of current is not accompanied by loss of energy through heat. Some metals at a temperature 0° abs. offer practically no resistance to the flow of electricity.

According to Ohm's law, $E = RI$. Therefore Eq. (16.4) becomes

TABLE 16.1 Resistivities and Temperature Coefficients of Metals *To obtain the resistivity in microhms for 1 cm² divide by 6.01*

Material	Resistivity at 0°C		Temperature coefficient per °C, at 20°C
	Circular mil-ft	$\mu\Omega/\text{cm}^3$	
Aluminum	17.1	2.85	0.00390
Copper, annealed	9.35	1.56	0.00393
Iron, pure	53.00	8.82	0.00600
Gold	12.36	2.06	0.00365
Lead	115.0	19.14	0.00390
Magnesium	30.0	4.99	0.00381
Mercury	564.00	93.84	0.00072
Nickel	41.60	6.92	0.00500
Platinum	66.00	10.98	0.00370
Silver	8.85	1.47	0.00400
Tantalum	87.60	14.58	0.00330
Tin	78.00	12.98	0.00365
Tungsten (hard drawn)	33.00	5.49	0.00320
Zinc	34.50	5.74	0.00400

TABLE 16.2   Rated Current Carrying Capacity of Different Gauges of Copper Wire

AWG Gauge	Single wire	Bundled wires
6	95 Amps	55 Amps
8	62 Amps	39 Amps
10	50 Amps	31 Amps
12	40 Amps	23 Amps
14	32 Amps	17 Amps
16	22 Amps	13 Amps
18	16 Amps	10 Amps
20	11 Amps	7.5 Amps
22	7 Amps	5 Amps
24	3.5 Amps	2.1 Amps
26	2.2 Amps	1.5 Amps
28	1.4 Amps	.8 Amps
30	.8 Amps	.5 Amps
32	.5 Amps	.3 Amps

Note: for safety do not exceed 50% of the rated value

TABLE 16.3   Resistivities and Temperature Coefficients of Low Temperature-Coefficient Alloys

Material	Resistivity at 20°C		Temperature coefficient per °C, at 20°C
	Circular mil-ft	$\mu\Omega/cm^3$	
Copper-nickel	100–250	16.64–41.59	0.000005–0.0004
Copper-nickel-zinc (nickel silver)	200–290	33.28–48.25	0.0002–0.00027
Iron-nickel	200–700	33.28–116.47	0.00034–0.001
Iron-nickel-chromium	520–720	86.52–119.80	0.00016–0.00072
Copper-manganese-nickel	249–270	41.43–44.93	$0.000025–3 \times 10^{-5}$

$$W = (RI)It = EIt \qquad (16.4')$$

which is the general expression for the energy supplied by a source of emf to a circuit obeying Ohm's law.

Resistance as an Intrinsic Property of Metals.   Although resistance is computed and defined through the current and voltage relations, or through the energy dissipated in an electric circuit, it must not be inferred that it is a function of either E, I, or W. Resistance is independent of all. It is a function of the material of the conductor and its physical dimensions. Experiment shows that the resistance of a straight conductor having a constant length h in the direction of current flow and a constant cross-sectional area A perpendicular to the flow is

$$R = \rho \frac{h}{A} \tag{16.5a}$$

whence
$$G = \frac{1}{R} = \frac{1}{\rho}\frac{A}{h} = \gamma \frac{A}{h} \tag{16.5b}$$

where ρ is known as the *resistivity,* and $\gamma = 1/\rho$ is called the *conductivity* of the metal. ρ and γ are defined as the *resistance* and *conductance,* respectively, of a cube having sides of unit length (a centimeter cube* or an inch cube).

The resistance of a conductor may be computed by the use of Eq. (16.5) only if that conductor satisfies the following conditions:

1. Its length in the direction of current flow must be constant.

2. Its cross section must be uniform and constant.

When these two conditions are not satisfied, there exist two alternative methods of determining resistance: experimentally, through the use of Ohm's or Joule's law, and mathematically, through the use of calculus. The latter method is limited in its scope to conductors having shapes that can be expressed by simple mathematical relations.

For engineering work, resistivity is usually expressed as the resistance of a circular mil-foot.† For rectangular conductors, the square mil-foot is preferable. For scientific work, a centimeter cube is one of the units used, but for good conductors this gives values inconveniently small. Circular-mil-foot resistivities for a number of materials are given in Tables 16.1 and 16.3.

Rated current capacity for different gauge copper wiring is listed in Table 16.2. Rated current will not cause the wiring to increase in temperature by more than 20°C. It is recommended that no more than half the rated current is passed through the conductors for safety.

6. Inductance L.   Inductance (similar in effect to mechanical inertia) is that characteristic of an electric circuit by virtue of which a sudden increase or decrease of the current is checked. Viewed from the standpoint of energy, inductance is that property which causes an electric circuit to store up energy while the current increases and to deliver energy while the current decreases.

The Magnetic Field Accompanying a Current and the Nature of Inductance.   Experiment shows that a circuit carrying an electric current is encircled by a magnetic field in a plane perpendicular to the direction of current flow. The magnetic field may be imagined to consist of lines or tubes of force (Fig. 16.2) which form closed paths around the current. Since current is a directed quantity,

---

*In defining resistivity and conductivity, distinction must be made between centimeter cube and cubic centimeter. The former specifies not only the volume but the dimensions as well. The latter gives the volume irrespective of dimensions. Resistivity is the resistance of a *centimeter cube* and not of a cubic centimeter. The resistance of a cubic centimeter is an indefinite quantity.

†A circular mil is the area of a circle $\frac{1}{1000}$ inch in diameter. This is a convenient unit for measuring the areas of round wires because the area, expressed in circular mils, is equal to the square of the diameter. The area of 1 circular mil, in square mils, is $(1)^2\pi/4 = \pi/4$. The area of a circle of diameter D mils, in square mils, is $D^2\pi/4$. Hence if $\pi/4$ is used as the unit of area, the area of the circle becomes D^2 circular mils. A conductor of 1 circular mil area and of 1 ft length is 1 circular mil-ft. 1 mil = 0.001 in.
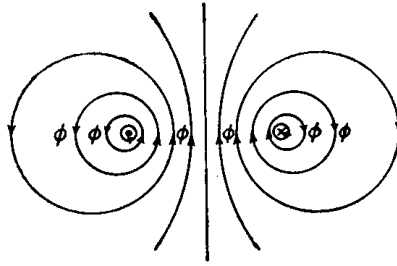
FIGURE 16.2

these tubes of force (Fig. 16.2) will assume reverse directions when they surround oppositely flowing currents. Using a person's right hand, with the thumb pointed outward and the fingers closed, the thumb points in the direction of the magnetic field when the electrical current flow follows the direction of the fingers. This is known as the "right hand rule." Similarly, if a current is directed into the plane of the paper (see X, Fig. 16.2) then the flux surrounding it is clockwise, and vice versa. These facts are illustrated in Fig. 16.2, which shows cross sections of two conductors carrying oppositely directed currents. Current flowing away from the reader is shown by X, while current flowing toward the reader is indicated by a dot. Circles representing the lines of force are drawn around the current. The direction of these lines of force follows the convention stated above. The same relation holds in the relative direction of the flux through a solenoid and its surrounding exciting current.

If the magnetic permeability of the circuit and that of the ambient medium are constant (i.e., if the phenomenon of magnetic saturation does not exist), then the flux linking with the circuit (Fig. 16.2) is proportional to the current, or

$$\phi = \frac{L}{N} i \qquad (16.6)$$

where ϕ = magnetic flux linking with the circuit
　　　i = current in the circuit
　L/N = coefficient of proportionality between ϕ and i (the value of L/N depends on the physical dimensions of the circuit and the units used)

Experiment further shows that if the flux linking a circuit be changed, then an emf e_L is induced in the circuit which is proportional to the time rate of change of flux. In other words,

$$e_L = -N \frac{d\phi}{dt} \qquad (16.7)$$

where N is the number of turns in the circuit which are completely linked by the flux ϕ.

There are, therefore, these two experimental facts:

1. In media of constant permeability, a change in the current produces a proportional change in the flux encircling it [Eq. (16.6)].
2. A change in the flux linking with a circuit induces an emf in that circuit [Eq. (16.7)].

The inevitable conclusion from these two premises is that a change in the current flowing in a circuit produces an emf in that circuit. Experiment shows that such is the case. Mathematically, this deduction may be easily verified by substituting Eq. (16.6) in (16.7). This gives

$$e_L = -L \frac{di}{dt} \tag{16.8}$$

L is known as the *inductance, self-inductance,* or *coefficient of self-induction* of the circuit. It is the coefficient of proportionality between the voltage e_L induced in a circuit and the time rate of change of current in it. The symbol for inductance is a helix (see Fig. 16.3).

The negative sign in Eq. (16.8) indicates that the voltage induced in a circuit, due to a change in current, is such as to oppose the change. Thus, if the current increases, a voltage is induced which tends to check that increase, and vice versa. Lenz was the first to observe this fact. Its validity can be easily seen from the following reasoning.

Suppose that a continuous emf is suddenly impressed on a series circuit having resistance and inductance. The current, heretofore zero, will in time assume a certain definite steady value I. In the interim, the current must increase from zero to I. If each successive increase induces a voltage which tends to increase the current further, then, instead of assuming a finite value I, the current will ultimately become infinite—and this is contrary to physical facts.

Inductance Defined in Terms of Flux Linkages.   Solving Eq. (16.6) for L,

$$L = \frac{N\phi}{i} \quad \left(\text{or } N \frac{d\phi}{di} \right) \tag{16.9}$$

Equation (16.9) states that the inductance of a circuit is equal to the flux per unit current which links with the turns of N of the circuit. The product $(N\phi)$ is often called the *flux linkages*. Hence inductance may be defined as the flux linkages per unit current.

The Limitations of the Concept of Inductance.   In defining inductance by Eq. (16.8), certain restrictions were imposed. These are given below:

1. The concept of inductance applies only to conductors made of materials of constant permeability and surrounded by media of constant permeability. One cannot, for example, speak of the inductance of an iron conductor, nor can one refer to the inductance of a copper wire wound around an iron core unless it is agreed that the term is to be used only where the permeability is constant. This is due to the fact that Eq. (16.6) must hold true.
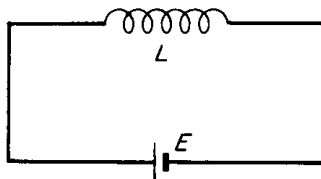
FIGURE 16.3

2. The emf induced in a circuit by virtue of its inductance is invariably due to the change of current carried by that circuit itself and is not due to any other cause. Thus the emf induced in the armature winding of a generator due to its rotation in a magnetic field is not an emf of self-induction. This is why the term *self-inductance* is preferable. It is more descriptive of the phenomenon.

3. Self-inductance disappears if, by any means whatever, the magnetic field due to a current is eliminated. Under these conditions, a change in current does not produce a corresponding change in the flux because ϕ is nonexistent. Although this is not fully realizable, an approach to it may be had if the circuit is wound non-inductively.

Inductance as a Storehouse of Electromagnetic Energy. From an energy viewpoint, *inductance* may be defined as that property of an electric circuit by virtue of which energy is stored in a circuit while the current increases and is given up while the current decreases. This statement will be amplified by describing an experiment and actually computing the stored energy.

Figure 16.4*a* shows a circuit with a continuous emf source E, a variable resistance R which may be reduced to R', and a constant inductance L. Let the resistance be maintained at a value R, and let steady conditions be established. The steady current, according to Ohm's law, is (see Fig. 16.4*b*) $I = E/R$. Next let the resistance suddenly be reduced to R' at the instant t_1. The final steady current is *(see Fig. 16.4b)* $I' = E/R'$.

The current does not increase instantly from I' to I' but rather rises gradually, following curve a (Fig. 16.4*b*). During the time of increase, a voltage of self-induction exists of such a direction as to oppose the impressed voltage. In other words, one part of the impressed voltage is devoted to counterbalance the emf of self-induction and the other part is available to send current through the circuit. That part of the impressed voltage which overcomes e_L is

$$e'_L = -e_L = + L\,\frac{di}{dt} \tag{16.10}$$

If the instantaneous value of the current is i, then the energy supplied by the emf source in time dt at an emf e'_L is

$$dW = e'_L i\,dt = Li\,di \tag{16.11}$$

and the total energy supplied by the source and stored in the circuit by virtue of its inductance is

$$W = \int_I^{I'} Li\,di = \tfrac{1}{2}L(I'^2 - I^2) \tag{16.12}$$
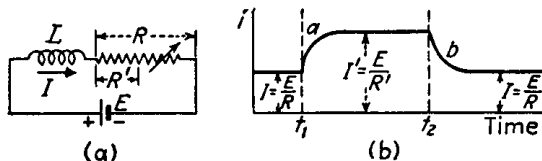
FIGURE 16.4

It can be shown similarly that during the decrease of the current (see curve b, Fig. 16.4b) from I' to I, the magnitude of the energy returned from the circuit to the continuous emf source is also expressed by Eq. (16.12).*

Inductance as an Intrinsic Property of a Circuit.   The reader would ask whether L could possibly be computed from the dimensions of a circuit. The answer is yes only in very simple and special cases. It is outside the scope of this text to deal with this phase of inductance. It will therefore be assumed that the inductance of a circuit is either known or can be determined experimentally.

7. Capacitance C.   When a dielectric separates two conductors, there results a *capacitor* or *condenser.* The name *condenser* or *capacitor* generally signifies a device to store electricity and which is made up of parallel metallic sheets with insulating sheets interleaved between them (see also Secs. 48 to 52). The dielectric may be air or any other nonconducting material. The conductors which the dielectric separates may be two metallic plates, two wires, a wire and the earth, metallic spheres, or a combination of any two or more of the above-mentioned conductors. The symbol for capacitance is two parallel lines (see d and f, Fig. 16.5a). For simplicity of treatment, it will be assumed that the dielectric separating the plates is perfect (for example, vacuum or air) and that the emf is kept low enough to eliminate corona and other disturbing influences.

Relation Between the Charge and the Charging Current.   Figure 16.5a *shows* a circuit consisting of a resistance R, a parallel-plate capacitor C, a ballistic galvanometer G, an oscillograph 0, a source of continuous emf E, and a double-throw switch S. Upon moving S to point a, the emf E is impressed on the capacitor C. The ballistic galvanometer shows a deflection indicating that a quantity of electricity Q has passed from the source to the capacitor. This charge of electricity is nothing but a momentary current which continues to flow for a short time and then ceases altogether as shown by the curve i_c (Fig. 16.5b).

Electric current may be regarded as consisting of a transfer of electricity along a conductor. Consequently, the amount (or quantity) of electricity which passes through any cross section of a conductor in time t is proportional to the magnitude of the current and the time of flow. Thus the readings of the ballistic galvanometer and the oscillograph may be coordinated through the following equations:
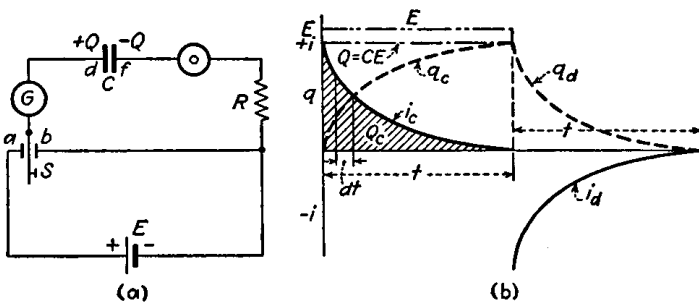
FIGURE 16.5

---

*The expression for returned energy is the negative of Eq. (16.12). The negative sign means that the emf source is receiving rather than supplying energy.

$$Q = \int_0^t i \, dt \tag{16.13}$$

or

$$i = \frac{dq}{dt} \tag{16.14}$$

Equation (16.13) states that the total charge acquired by the condenser is the area under the curve i_c (Fig. 16.5b). Equation (16.14) states that the magnitude of the current flowing in the circuit (Fig. 16.5a) is at every instant equal to the rate at which the quantity of electricity is carried along the wires. This is graphically the slope of the curve q_c and is represented by the curve i_c (Fig. 16.5b).

The Charging EMF Is Equal to the Potential across the Plates of a Charged Capacitor. Let the switch S be open after a time t (Fig. 16.5b) when the current has become zero. No spark is observed when the switch is open, because no current has been interrupted. Upon examining the plates d and f, it will be found that the plate d has a positive charge $+Q$ and f has an equal negative charge $-Q$. Moreover, measurement of the potential difference E_c between the plates shows that the voltage is actually equal to the impressed emf E, or

$$E = E_c \tag{16.15}$$

Under these conditions the capacitor is said to be fully charged.

Equality of the Charge and Discharge. Let the switch S now be moved to the point b. The capacitor begins to discharge through the resistance R. The galvanometer shows a deflection equal in magnitude to that previously indicated, but opposite in direction. This indicates that the quantity of electricity Q_d obtained during the discharge is equal to the quantity Q_c supplied during the charge, or

$$Q_c = Q_d \tag{16.16}$$

A capacitor may therefore be defined as a device for storing electricity—not in the form of electromagnetic energy, as in the case of inductance, but as electrostatic energy (the electric charge).

Relation Between the Charging EMF and the Charge. Experiment *shows* that the charge acquired by a capacitor whose conducting terminals are separated by a perfect dielectric is proportional to the emf across its plates. This proportionality between E_c and Q may be expressed in the following two ways:

$$C = \frac{Q}{E_c} \tag{16.17a}$$

$$S = \frac{E_c}{Q} \tag{16.17b}$$

or

C is known as the *capacitance* of a condenser (or capacitor), and $S = 1/C$ is known as its *elastance*. Equations (16.17) are similar to Eqs. (16.2) and express what may be termed *Ohm's law for the electrostatic circuit*. Equation (16.17a) gives a definition of capacitance. It is simply the ratio Q/E_c.

Capacitance as an Intrinsic Property of a Condenser.   The *capacitance* **C** *for per*fect dielectrics is independent of both Q and E and is simply a function of the dimensions and material of the capacitor. Experiment shows that for capacitors made of plates having equal areas placed parallel to each other and separated by a homogeneous perfect dielectric of area A and thickness h, the capacitance C is*

$$C = Kk_e \frac{A}{h} \tag{16.18}$$

where K = a constant known as the *relative permittivity of the material,* which is the ratio of the permittivity k of a material to the permittivity k_e of vacuum, or $K = k/k_a$
$\quad\;\;$ k_a = the *permittivity of vacuum* and may be taken as $8.842 \times 10^{-14}\,\text{F/cm}^3$
$\quad\;\;$ A = area of one plate or of a cross section of the dielectric in square centimeters
$\quad\;\;$ h = distance between plates in centimeters

Capacitance Viewed from an Energy Standpoint.   The energy supplied to a capacitor is

$$dW = ei\,dt = e\,dq \tag{16.19}$$

But by Eq. (16.17)

$$dq = C\,de \tag{16.20}$$

$$\therefore W = \int_0^{E_c} Ce\,de = \tfrac{1}{2}CE_c^2 = \tfrac{1}{2}CE^2 \tag{16.21}$$

whence
$$C = \frac{2W}{E^2} \tag{16.22}$$

Thus the capacitance of a condenser is a measure of its capacity to store electric energy at a given applied potential.

8. Types of Circuits.   There are three general types of circuits: (1) series circuits, (2) parallel circuits, and (3) series-parallel circuits.

   1. *Series Circuit:* A *series circuit* (Fig. 16.6) is one which has the same current I in all its parts. Thus Fig. 16.6*a* consists of four resistances connected in series across the same emf *E*. They are in series because one and the same current I flows through all the resistances. Similarly, Fig. 16.6*b*. comprises four inductances connected in series. Figure 16.6*c* consists of four capacitances in series, while Fig. 16.6*d* has two resistances, two inductances, and two capacitances in series.

---

*Equation (16.18) is true under the following limitations: (1) it neglects fringing of the electrostatic flux at the edges of the plates, (2) it assumes a perfect contact between the plates and the dielectric, and (3) it assumes a nonanomalous perfect dielectric in which no hysteroviscosity or conduction effects can be observed.
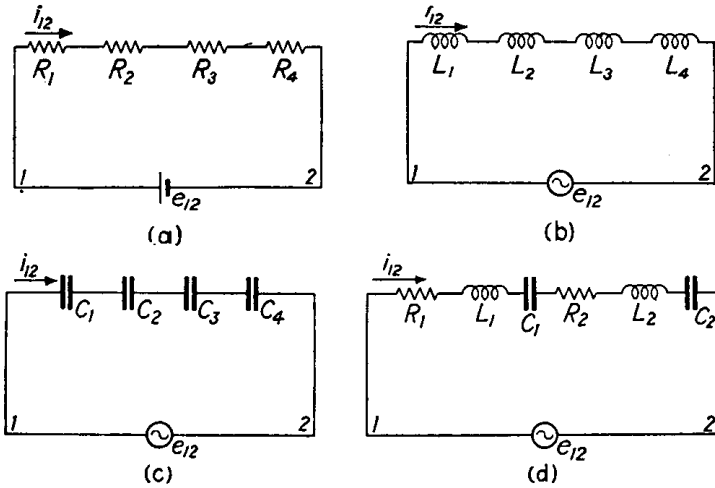
FIGURE 16.6

2. *Parallel Circuit:* A *parallel circuit* (Fig. 16.7) is one which has the same emf across all its parts. Thus the circuit in Fig. 16.7a has the same emf E across all the characteristics R_1, R_2, R_3, R_4. The currents in these characteristics i_1, i_2, i_3, and i_4, respectively, may or may not be the same. Figure 16.7b is a circuit consisting of four inductances in parallel, while Fig. 16.7c consists of four capacitances in parallel. Figure 16.7d consists of two resistances, two inductances, and two capacitances, all connected in parallel.

3. *Series-Parallel Circuit* (*or Network*): A *series-parallel circuit* (or *network*) is a combination of circuit characteristics which includes both series and parallel (Fig. 16.8).
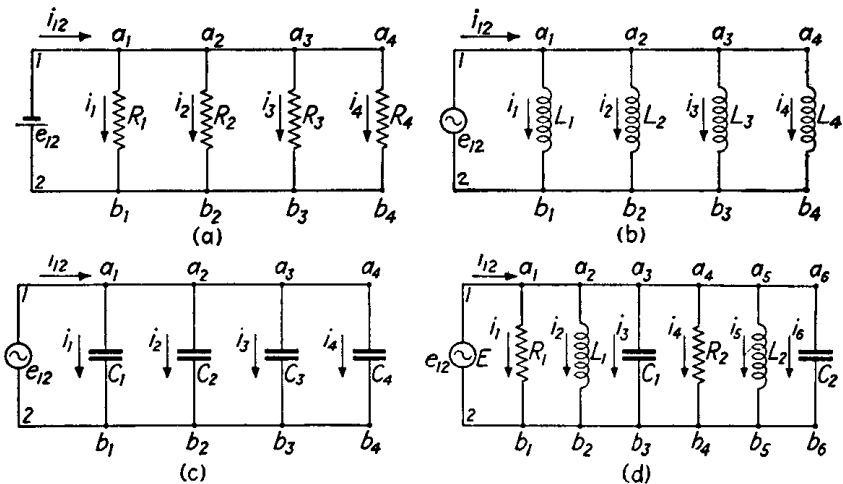
FIGURE 16.7
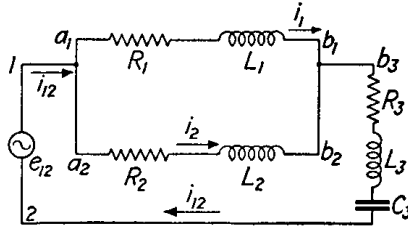
FIGURE 16.8

9. Kirchhoff's Laws

Branch.   A *branch* is a part of a circuit which carries one current. A branch may consist of one characteristic or of several characteristics connected in series. Thus all the circuits in Fig. 16.6 have but one branch each. Figure 16.7a to c has four branches each: a_1b_1, a_2b_2, a_3b_3, and a_4b_4. Figure 16.7d has six branches: a_1b_1, a_2b_2, a_3b_3, a_4b_4, a_5b_5, and a_6b_6. Figure 16.8 has two branches: a_1b_1, and a_2b_2.

Junction.   A *junction* is a point at which two or more branches meet. Thus Fig. 16.7a to c has two junctions each: $(a_1a_2a_3a_4)$ and $(b_1b_2b_3b_4)$. Figure 16.8 has two junctions: (a_1a_2) and (b_1b_2). Observe that any two or more points between which no circuit characteristic or emf exists are electrically identical and form one junction. Thus in Fig. 16.7 the points a_1, a_2, a_3, a_4 are electrically identical and form one junction a.

Loop.   A *loop is any* closed path in a circuit. A loop may be *simple* or *compound*. A *simple loop* is one which encloses no branches. A compound loop is one enclosing branches. Thus Fig. 16.6a to d has one simple loop each. Figure 16.7a to c has the following simple loops: $1a_1b_12$, $a_1a_2b_2b_1$, $a_2a_3b_3b_2$, and $a_3a_4b_4b_3$. Figure 16.7a to c also has the following compound loops: $1a_2b_22$, $a_1a_2a_3b_3b_2b_1$, $a_1a_2a_3a_4b_4b_3b_2b_1$, $1a_1a_2a_3b_3b_2b_12$, $1a_1a_2a_3a_4b_4b_3b_2b_12$.

Potential Drop.   The potential drop in any characteristic of a circuit is equal to the emf which is required to make a current flow in that characteristic. Thus if e_r, e_L, and e_c are the potential drops in a resistance, inductance, and capacitance, respectively (see Secs. 5, 6, and 7),

$$e_r = Ri \tag{16.23a}$$

$$E_L = L\frac{di}{dt} \tag{16.23b}$$

$$e_c = \frac{q}{C} = \frac{1}{C}\int i\,dt \tag{16.23c}$$

EMF-Current Notations.   The direction of the emf and current which it causes will be as follows: If E_{12} is an emf applied to the points 1 and 2 of a circuit, that emf will cause a current to flow from point 1 to 2 through the characteristics of

that circuit. The direction of flow of the current is indicated by an arrow (see Figs. 16.6 to 16.8) pointing from point 1 to point 2 of the characteristics.

Kirchhoff's EMF Law. The sum of the potential drops around any loop is equal to the sum of the emf's in that loop, provided the notations are as in the preceding subsection. Any current which flows in a direction opposite to that specified in the preceding subsection causes a negative drop. Moreover, in any one loop, oppositely directed currents cause drops of opposite signs.

Kirchhoff's emf law for some of the loops in Figs. 16.6 to 16.8 is expressed in Eqs. (16.24) to (16.29):

Fig. 16.6*a*:

$$e_{12} = i_{12}R_1 + i_{12}R_2 + i_{12}R_3 + i_{12}R_4 = i_{12}(R_1 + R_2 + R_3 + R_4) \quad (16.24)$$

Fig. 16.6*b*:

$$e_{12} = L_1 \frac{di_{12}}{dt} + L_2 \frac{di_{12}}{dt} + L_3 \frac{di_{12}}{dt} + L_4 \frac{di_{12}}{dt} \quad (16.25a)$$

$$= \frac{di_{12}}{dt} (L_1 + L_2 + L_3 + L_4) \quad (16.25b)$$

Fig. 16.6*c*:

$$e_{12} = \frac{q}{C_1} + \frac{q}{C_2} + \frac{q}{C_3} + \frac{q}{C_4} = q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \frac{1}{C_4}\right) \quad (16.26a)$$

$$= \left(\int i_{12}\, dt\right)\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \frac{1}{C_4}\right) \quad (16.26b)$$

Fig. 16.6*d*:

$$e_{12} = i_{12}R_1 + L_1 \left(\frac{di_{12}}{dt}\right) + \left(\int i_{12} \frac{dt}{C_1}\right)$$

$$+ i_{12}R_2 + L_2 \left(\frac{di_{12}}{dt}\right) + \left(\int i_{12} \frac{dt}{C_2}\right) \quad (16.27a)$$

$$= (R_1 + R_2)i_{12} + (L_1 + L_2) \left(\frac{di_{12}}{dt}\right) + \left(\frac{1}{C_1} + \frac{1}{C_2}\right)\int i_{12}\, dt \quad (16.27b)$$

Fig. 16.7*a*:

Loop $1a_1b_12$: $\qquad\qquad\qquad\qquad e_{12} = i_1R_1$ $\qquad\qquad\qquad$ (16.28*a*)

Loop $a_1a_2b_2b_1a_1$: $\qquad\qquad\qquad 0 = i_1R_1 - i_2R_2$ $\qquad\qquad\qquad$ (16.28*b*)

Loop $a_2a_3b_3b_2a_2$: $\qquad\qquad\qquad 0 = i_2R_2 - i_3R_3$ $\qquad\qquad\qquad$ (16.28*c*)

Loop $a_3a_4b_4b_3a_3$: $\qquad\qquad\qquad 0 = i_3R_3 - i_4R_4$ $\qquad\qquad\qquad$ (16.28*d*)

Fig. 16.8:

Loop $1a_1b_1b_2 2$:

$$e_{12} = i_1 R_1 + L_1 \left(\frac{di_1}{dt}\right) + R_3 i_{12} + L_3 \left(\frac{di_{12}}{dt}\right) + \int \frac{i_{12}\, dt}{C_3} \qquad (16.29a)$$

Loop $a_1 b_1 b_2 a_2$:

$$0 = i_1 R_1 + L_1 \left(\frac{di_1}{dt}\right) - i_2 R_2 - L_2 \left(\frac{di_2}{dt}\right) \qquad (16.29b)$$

Kirchhoff's Current Law. The sum of currents entering a junction is equal to the sum of currents leaving that junction.

Note: Entrance to a junction is denoted by an arrow directed toward the junction; exit is indicated by an arrow directed away from that junction.

Kirchhoff's current law for the junction a, Figs. 16.7 and 16.8, is expressed in Eqs. (16.30).

Fig. 16.7a to c $\qquad\qquad\qquad i_{12} = i_1 + i_2 + i_3 + i_4 \qquad\qquad\qquad (16.30a)$

Fig. 16.7d: $\qquad\qquad\qquad i_{12} = i_1 + i_2 + i_3 + i_4 + i_5 + i_6 \qquad\quad (16.30b)$

Fig. 16.8: $\qquad\qquad\qquad i_{12} = i_1 + i_2 \qquad\qquad\qquad\qquad\qquad (16.30c)$

10. Use of Kirchhoff's Laws. Kirchhoff's emf and current laws are used to determine the current flowing in each branch of a given circuit when the emf and characteristics of that circuit are specified. It can be shown that in order to solve for the currents, the number of Kirchhoff's emf equations must be equal to the number of simple loops in a circuit, while the number of Kirchhoff's current equations must be equal to the number of junctions less one.

DIRECT-CURRENT CIRCUITS

11. Circuit with One Resistance. Consider a circuit (Fig. 16.1) in which direct voltage E is impressed on a resistance R. In order to determine the current I, use is made of Ohm's law. Thus $E = IR$, and

$$I = \frac{E}{R} \qquad (16.31)$$

12. Circuit with One Inductance. The circuit in Fig. 16.3 consists of a direct voltage E impressed on an inductance L. Equation (16.10) is used to determine the current i:

$$E = L\frac{di}{dt} \qquad (16.32)$$

Solving Eq. (16.32) for i,

$$i = \frac{1}{L} \int E \, dt + k \qquad (16.33a)$$

$$i = \frac{E}{L} t + k \qquad (16.33b)$$

Equation (16.33b) shows that the current i increases indefinitely when a direct voltage E is impressed on an inductance. Indeed, an inductance acts virtually as a short circuit for a constant direct emf.

13. Circuit with One Capacitance. The circuit in Fig. 16.9 consists of a direct voltage E impressed on a capacitance C. Here, by Eq. (16.17a),

$$q = CE \qquad (16.34a)$$

$$i = \frac{dq}{dt} = \frac{C \, dE}{dt} \qquad (16.34b)$$

$$i = 0 \qquad (16.34c)$$

Equation (16.34c) states that the *steady* current in Fig. 16.9 is zero. In other words, a condenser acts virtually as an open circuit when a direct emf is impressed on it.

Since an inductance acts as a short circuit and a condenser acts as an open circuit whenever a direct emf is impressed on them, they will be considered as nonadmissible characteristics in this section. The section will be devoted, therefore, to series, parallel, and series-parallel circuits of resistances only.

14. Series Circuit with Several Resistances. In the circuit in Fig. 16.10a, there are n resistances $R_1, R_2, \ldots, R_n$ connected in series across a direct voltage E. By Kirchhoff's law,
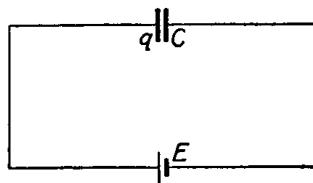
FIGURE 16.9
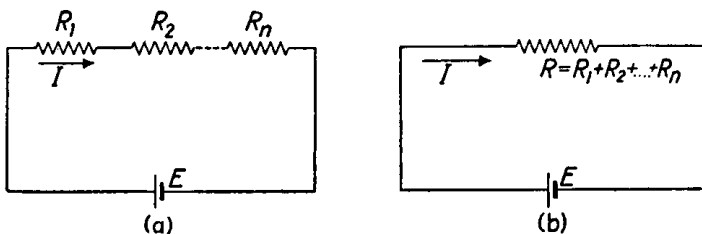
FIGURE 16.10

$$E = IR_1 + IR_2 + \cdots + IR_n \tag{16.35a}$$

$$= I(R_1 + R_2 + \cdots + R_n) \tag{16.35b}$$

$$= IR \tag{16.35c}$$

where $\qquad R = R_1 + R_2 + \cdots + R_n \tag{16.35d}$

Equations (16.35c) and (16.35d) state that the series circuit (Fig. 16.10a) may be replaced by a simple circuit comprising one resistance R (Fig. 16.10b) whose value is equal to the sum of the series resistances.

15. Parallel Circuits with Several Resistances. In the circuit in Fig. 16.11a, there are n resistances in parallel across the same voltage E. By Ohm's law,

$$I_1 = \frac{E}{R_1} \quad I_2 = \frac{E}{R_2} \quad I_n = \frac{E}{R_n} \tag{16.36}$$

Applying Kirchhoff's current law to the junction a,

$$I = I_1 + I_2 + \cdots + I_n \tag{16.37a}$$

$$= E \frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_n} \tag{16.37b}$$

$$= \frac{E}{R} \tag{16.37c}$$

where $\qquad \dfrac{1}{R} = \dfrac{1}{R_1} + \dfrac{1}{R_2} + \cdots + \dfrac{1}{R_n} \tag{16.37d}$

Equations (16.36) give the currents in the various branches of the parallel circuit. Equation (16.37a) states that the total current I supplied by the emf source is the sum of these currents. Equations (16.37c) and (16.37d) state that the parallel circuit (Fig. 16.11a) may be replaced by a simple circuit (Fig. 16.11b) comprising one resistance whose value is expressed by Eq. (16.37d).

16. Series-Parallel Circuits (Networks). The determination of the currents in a network of resistances (Fig. 16.12) is accomplished by the application of Kirchhoff's laws. Thus, applying Kirchhoff's emf law to the loops $1a_1b_12$ and $a_1b_1b_2a_2$,
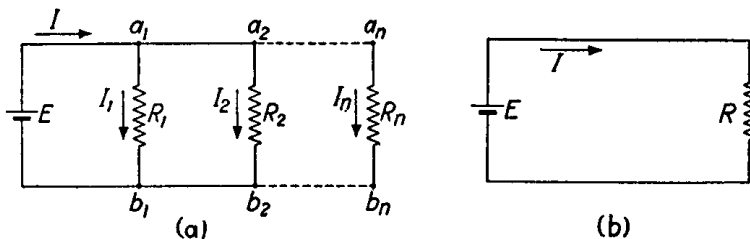
FIGURE 16.11
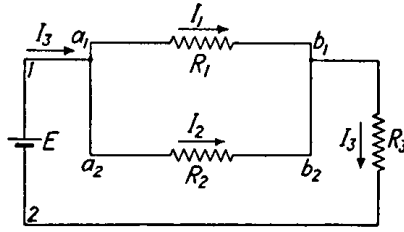
FIGURE 16.12

$$E = I_1R_1 + I_3R_3 \qquad (16.38a)$$

$$0 = I_1R_1 - I_2R_2 \qquad (16.38b)$$

Applying Kirchhoff's current law to the junction a,

$$I_3 = I_1 + I_2 \qquad (16.39)$$

Equations (16.38) and (16.39) give three independent relations among the currents I_1, I_2, and I_3 from which these currents can he determined. Thus write these equations as follows:

$$R_1I_1 + 0I_2 + R_3I_3 = E \qquad (16.40a)$$

$$R_1I_1 - R_2I_2 + 0I_3 = 0 \qquad (16.40b)$$

$$I_1 + I_2 - I_3 = 0 \qquad (16.40c)$$

Using determinants (see Sec. 13), let

$$\Delta = \begin{vmatrix} R_1 & 0 & R_3 \\ R_1 & -R_2 & 0 \\ 1 & 1 & -1 \end{vmatrix} = (R_1R_2 + R_1R_3 + R_2R_3) \qquad (16.41)$$

$$\therefore I_1 = \frac{\begin{vmatrix} E & 0 & R_3 \\ 0 & -R_2 & 0 \\ 0 & 1 & -1 \end{vmatrix}}{\Delta} = \frac{R_2E}{R_1R_2 + R_1R_2 + R_2R_3} \qquad (16.42a)$$

$$I_2 = \frac{\begin{vmatrix} R_1 & E & R_2 \\ R_1 & 0 & 0 \\ 1 & 0 & -1 \end{vmatrix}}{\Delta} = \frac{-ER_1}{R_1R_2 + R_1R_3 + R_2R_3} \qquad (16.42b)$$

$$I_3 = \frac{\begin{vmatrix} R_1 & 0 & E \\ R_1 & -R_2 & 0 \\ 1 & 1 & 0 \end{vmatrix}}{\Delta} = \frac{E(R_2 - R_1)}{R_1R_2 + R_1R_3 + R_2R_3} \qquad (16.42c)$$

17. Power and Energy in Direct-Current Circuits. Power in a dc circuit is defined as follows:

$$P = EI \tag{16.43}$$

where E is the emf impressed on the circuit, and I is the total current furnished by that emf. If E is expressed in volts and I in amperes, P is in watts.

The power in any branch k of the circuit is defined through Joule's law as follows:

$$P_k = I_k^2 R_k \tag{16.44}$$

where P_k is the power in branch k, I_k is the current flowing through branch k, and, R_k is the resistance of branch k.

The energy supplied to a dc circuit in an interval of time of length t is

$$W = Pt = EIt \tag{16.45}$$

If P is in watts and t is in seconds, W is expressed in watt-seconds or joules. The energy supplied to each branch in an interval t is

$$W_k = P_k t = I_k^2 R_k t \tag{16.46}$$

But, by the principle of conservation of energy, the total energy supplied must be equal to the sum of the energies supplied to all the branches. Hence, if a circuit has n branches,

$$W = EIt = \sum_{k=1}^{k=n} I_k^2 R_k t \tag{16.47a}$$

or

$$EI = \sum_{k=1}^{k=n} I_k^2 R_k \tag{16.47b}$$

Equation (16.47a) is the law of conservation of *energy* in a dc circuit, and Eq. (16.47b) is the law of conservation of power in a dc circuit.

SINGLE-PHASE ALTERNATING-CURRENT CIRCUITS

18. Sine Functions

Definition. A sine function of x is of the form

$$e = E_m \sin (x + \alpha) \tag{16.48}$$

Graph. A graph of the sine function given in Eq. (16.48) is shown in Fig. 16.13, wherein E_m is the *maximum value* or *amplitude* of the sine function, e_0 is the *initial value* of the sine function or its value corresponding to $x = 0$, and α is the initial angle of the sine function.
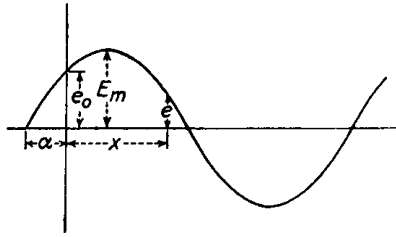
FIGURE 16.13

Periodicity. Sine functions are periodic and have a period of 2π. In other words, a sine function repeats itself every 2π radians, or

$$e = E_m \sin (x + \alpha) = E_m \sin (x + \alpha + 2n\pi) \tag{16.49}$$

where n is any integer.

Cycle. The set of values which a sine function assumes in the course of a period constitute a *cycle*. The term *cycle* is often used in ac circuits and is denoted by the symbol $\sim$.

19. Average and Effective Values. Frequent reference is made in the text to the average and effective values of a sine function. These are now defined.

Average Value E_a. In general, the average value of any function between two limits θ and β is the value of the mean ordinate between these limits. The mathematical expression for this statement is

$$E_a = \left(\frac{1}{\theta - \beta}\right) \int_\beta^\theta e \, dx \tag{16.50}$$

Applying Eq. (16.50) to a sine function, Eq. (16.48),

$$E_a = \left(\frac{1}{\theta - \beta}\right) \int_\beta^\theta E_m \sin (x + \alpha) \, dx \tag{16.51a}$$

$$= \left(\frac{E_m}{\theta - \beta}\right) [\cos (\beta + \alpha) - \cos (\theta + \alpha)] \tag{16.51b}$$

It is evident from an inspection of Eq. (16.51) and Fig. 16.14 that the average value of a sine wave varies depending on the limits θ and β. For many purposes, however, the sine wave is integrated from its zero value to an immediately succeeding π value. Thus in Fig. 16.14 the limits are taken from $x = -\alpha$ to $x = (\pi - \alpha)$. Substituting these limits in Eq. (16.51) and simplifying,

$$E_a = \frac{E_m}{\pi} (\cos 0 - \cos \pi) = 2 \frac{E_m}{\pi} \tag{16.52}$$
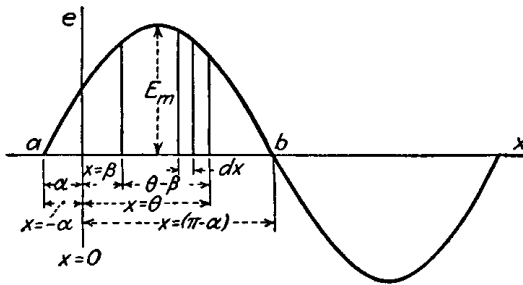
FIGURE 16.14

Effective Value E.   The effective value E of a periodic function between the limits β and θ is the square root of the area under the squared curve divided by the base.

Thus let the curve in Fig. 16.15 be obtained by squaring the ordinates of the curve in Fig. 16.14. The effective value of the sine wave between the limits θ and β is the square root of the quotient obtained by dividing the value of the shaded area by the base $\theta - \beta$. The mathematical expression for this operation is

$$E = \left(\int_{\beta}^{\theta} \frac{e^2 \, dx}{\theta - \beta} \right)^{1/2}$$
(16.53)

Applying Eq. (16.53) to a sine function having the form of Eq. (16.48),

$$E = \left(\int_{\beta}^{\theta} \frac{E_m^2 \sin^2(x + \alpha) \, dx}{\theta - \beta} \right)^{1/2}$$
(16.54)

Here again the effective value of a sine function will vary depending on the limits β and θ. For most purposes, however, the limits are taken as 0 and 2π. Substituting these limits in Eq. (16.54),

$$E = E_m \left[\frac{1}{2\pi} \int_0^{2\pi} \sin^2(x + \alpha) \, dx \right]^{1/2} = \frac{E_m}{\sqrt{2}}$$
(16.55)

Amplitude and Form Factors.   The two values defined above have given rise to two factors known as the *amplitude* and *form factors.*

The amplitude factor k_a is the ratio of the maximum to the effective value of a periodic wave, or
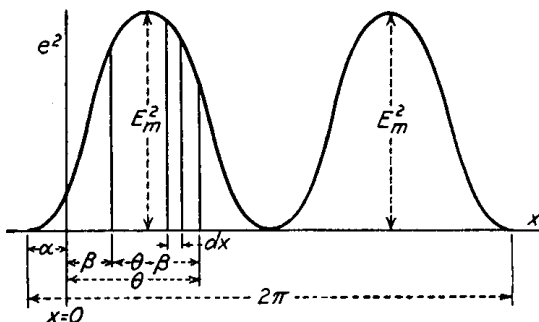
FIGURE 16.15

$$k_a = \frac{E_m}{E}$$ (16.56)

For sine waves,

$$k_a = \sqrt{2}$$ (16.57)

The form factor k_f is the ratio of the effective value to the average value of a periodic wave, or

$$k_f = \frac{E}{E_a}$$ (16.58)

For sine waves,

$$k_f = \frac{\pi}{2\sqrt{2}} \cong 1.11$$ (16.59)

20. Lead, Lag, and Phase. *Lead* means to be ahead, and *lag* means to be behind. Graphically (see curve *a*, Fig. 16.16), a sine wave leads the origin if its zero value (which is nearest to the origin) occurs *before* the point $x = 0$. A curve lags the origin if its zero value (which is nearest to the origin) occurs *after* the point $x = 0$ (see curve *c*, Fig. 16.16). Finally, a sine function is in phase with the origin if its zero value *coincides* with $x = 0$ (see curve *b*, Fig. 16.16). The *angle of lead or lag* (the *phase angle* or the *initial angle*) is measured from the point $x = 0$ to the nearest zero value of the sine wave. A positive angle is measured to the left, and a negative angle to the right.

The terms *lead* and *lag* are relative. Thus if curve *a* in Fig. 16.16 leads the origin, then the origin lags curve *a*. These two statements have the same meaning and are interchangeable. Again, if curve *a* leads the origin by an angle α and curve *c* lags it by $-\beta$, then curve *a* leads *c* by an angle $(\alpha + \beta)$. When it is desired to give the phase angle without any reference to lead or lag, the term *out of phase* is used. Thus in Fig. 16.16 the curves *a* and *b* are $\alpha°$ out of phase and the curves *a* and *c* are $(\alpha + \beta)°$ out of phase.

It is very convenient to be able to state, by a mere inspection of the function and without actually plotting it, whether a sine function leads, lags, or is in phase with the origin. The value of the initial angle α [Eq. (16.48)] determines whether the nearest zero value of a sine function occurs before or after the point $x = 0$. In general, a sine function leads the origin (curve *a*, Fig. 16.16) if the initial angle is such that $0 < \alpha < \pi$ and lags (curve *c*, Fig. 16.16) if the initial angle is such that
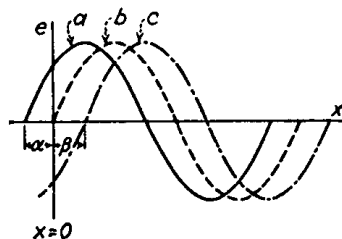
FIGURE 16.16

$\pi < \alpha < 2\pi$. If $\alpha = 0$ or $\alpha = 2n\pi$, the function is in phase with the origin. The use of either term, *lead* or *lag*, is justifiable when $\alpha = (2n + 1)\pi$.

21. Sine Functions and Revolving Vectors. Figure 16.17a shows a vector of length E_m and initial argument α. Let this vector revolve at a uniform angular velocity ω such that in time t it moves through an angle x. Then by definition

$$\omega = \frac{x}{t} \quad \text{or} \quad x = \omega t \tag{16.60}$$

If the period (time required to make one complete revolution, or 2π radians) is denoted by T, then

$$\omega = \frac{2\pi}{T} \tag{16.61}$$

Consequently,*
$$x = \omega t = \frac{2\pi t}{T} = 2\pi f t \tag{16.62}$$

where
$$f = \frac{1}{T} \tag{16.63}$$

The quantity f is known as the *frequency of the revolving vector* and is the number of revolutions or cycles performed per second.

Now consider the projections of the revolving vector on the vertical axis. From the geometry of Fig. 16.17a, these are

$$e = E_m \sin (x + \alpha) = E_m \sin (\omega t + \alpha) = E_m \sin (2\pi f t + \alpha) \tag{16.64}$$

But Eq. (16.64) is simply a sine function which may be represented by the sine wave (Fig. 16.17b). It can be concluded, therefore, that a revolving vector of modulus E_m and initial argument a moving at a uniform angular velocity ω has projections along the vertical axis which correspond exactly with the instantaneous values e of a sine function of amplitude E_m, argument ωt, and initial angle α.

(a)   (b)
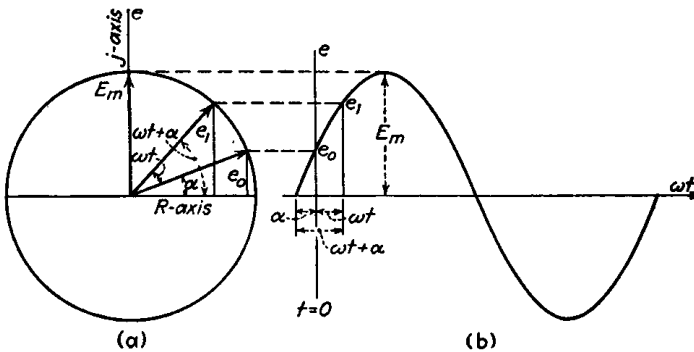
FIGURE 16.17

---

*Note that x may be expressed in radians or degrees depending on whether π is taken as 3.1416 rad or 180°.

22. The Use of Stationary Vectors in Alternating-Current Circuits. Let two sine waves (Fig. 16.18) be generated by two revolving vectors having the same angular velocity ω. Let the initial angles of the first and second be α and $-\beta$, respectively. It is evident that as long as the vectors move at the same velocity, the angle between them will always be $(\alpha + \beta)$. Now in many applications of sine functions to ac circuits this phase angle is sought. Since this angle is the same no matter whether the vectors revolve or are stationary, and since stationary vectors lend themselves easily to analytical treatments through complex numbers, the revolving vector is replaced by a stationary vector. Again, under steady conditions, many of the phenomena that occur in alternating electric circuits are such that what is true of any one instant is also true of all other instants. This furnishes a further justification for replacing revolving vectors by stationary ones which may be drawn for any arbitrary instant.

In representing sine functions by stationary vectors, the following facts should be remembered:

1. Stationary vectors may be used if, and only if, all the sine functions entering into the problem are of the same frequency f.

2. The phenomena must be such that what is true of one instant is also true of all instants (steady-state conditions must prevail).

3. The vector is so plotted that it makes an angle, with the reference axis, equal to the initial angle of the sine function. This angle is the argument of the stationary vector. It is, hence, the direction angle between the stationary vector and the reference axis. Thus a vector leads the reference axis if it has an argument α such that $0 < \alpha < \pi$. On the other hand, a vector lags the reference axis if its argument is such that $-\pi < \alpha < 0$. A vector is in phase with the R axis if $\alpha = 0$.

4. The angle that the vector makes with the reference axis is measured counterclockwise from the R axis to the vector. If the angle is measured clockwise from the R axis to the vector, it is considered a negative angle. Thus the angle β (Fig. 16.18a) is a negative angle.

5. The length of the vector is made proportional to the effective value E, instead of the amplitude E_m, of the sine function. This departure from the fundamental representation described above may be, at first, confusing. Its justification, however, for practical purposes becomes apparent when it is noted that meters read effective values rather than amplitudes, and most of the computations for electric circuits are based on effective values. Again, for sine waves, the relation between the effective

(a)   (b)
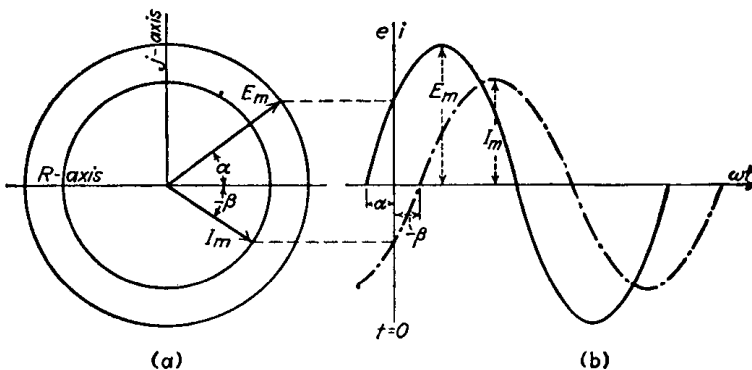
FIGURE 16.18

value and the amplitude is definite [see Eq. (16.55)]. Hence a change of scale makes the same vector represent either E or E_m.

6. The vectors representing sine functions possess all the properties of vectors. Since all vectors used in this text are coplanar, complex numbers can be used to define their moduli and arguments.

In order to make these facts clear, the instantaneous expressions for the sine curves drawn in Fig. 16.18b, and the complex expressions for the stationary vectors representing them, are given below:*

$$e = E_m \sin (\omega t + \alpha) \qquad (16.65a)$$

$$i = I_m \sin (\omega t - \beta) \qquad (16.65b)$$

$$\hat{E} = E\epsilon^{j\alpha} = \frac{E_m}{\sqrt{2}} \epsilon^{j\alpha} \qquad (16.66a)$$

$$\hat{I} = I\alpha^{-j\beta} = \frac{I_m}{\sqrt{2}} \epsilon^{-j\beta} \qquad (16.66b)$$

It should be noted that Eqs. (16.65) and (16.66) are reversible. In other words, given Eqs. (16.65), Eqs. (16.66) can easily be written, and vice versa.

23. Some Advantages of Vector Representation. The representation of sine functions by vectors possesses advantages which simplify the algebraic operations and render the physical phenomena clearer to the student. The demonstration of the latter aspect is reserved to succeeding pages. The following example shows how the addition and subtraction of sine functions are simplified through the use of vectors.

Let it be required to add the three sine waves

$$a = A_m \sin (\omega t + \alpha) \qquad (16.67a)$$

$$b = B_m \sin (\omega t + \beta) \qquad (16.67b)$$

$$c = C_m \sin (\omega t + \gamma) \qquad (16.67c)$$

Vectorially these waves become

$$\hat{A} = A\epsilon^{j\alpha} = a' + ja'' \qquad (16.68a)$$

$$\hat{B} = B\epsilon^{j\beta} = b' + jb'' \qquad (16.68b)$$

$$\hat{C} = C\epsilon^{j\gamma} = c' + jc'' \qquad (16.68c)$$

whence

$$\hat{D} = \hat{A} + \hat{B} + \hat{C} = (a' + b' + c') + j(a'' + b'' + c'') = d' + j'' = D\epsilon^{j\theta} \quad (16.69)$$

---

*Throughout this section, the symbol ϵ is used to represent the base of the Napierian or natural logarithms in order to avoid confusion between this term and instantaneous voltage, which is represented by the letter e.

Then the sum of' the three sine waves is a sine wave having an amplitude $D_m = D\sqrt{2}$ and an initial angle θ. The instantaneous value of this wave is

$$d = D_m \sin (\omega t + \theta) \tag{16.70}$$

Caution. Since the product of two sine functions yields a constant term and a double-frequency sine term, and since the frequency is entirely omitted in the representation of a sine function by a stationary vector, the student is hereby warned to avoid the use of vectors in multiplying two sine functions.

24. Cosine Functions. Cosine functions are encountered rather frequently. A separate treatment of these functions becomes superfluous when it is remembered that

$$\cos (x + \alpha) = \sin \left(x + \alpha + \frac{\pi}{2} \right) \tag{16.71}$$

Thus a cosine function can always be converted to a sine function by adding $\pi/2$ to the initial angle of the cosine function.

25. Circuit Containing Resistance. Consider a circuit (Fig. 16.19) which has a resistance R and a sine emf

$$e = E_m \sin (\omega t + \alpha) \tag{16.72}$$

The drop across the resistance is equal to the applied emf, or

$$e_r = Ri = E_m \sin (\omega t + \alpha) \tag{16.73}$$

Hence the current i is

$$i = \frac{E_m}{R} \sin (\omega t + \alpha) = I_m \sin (\omega t + \alpha) \tag{16.74a}$$

where

$$I_m = \frac{E_m}{R} \tag{16.74b}$$

Equation (16.74) states that the current in a circuit with resistance is in phase with the voltage and has an amplitude $I_m = E_m/R$.
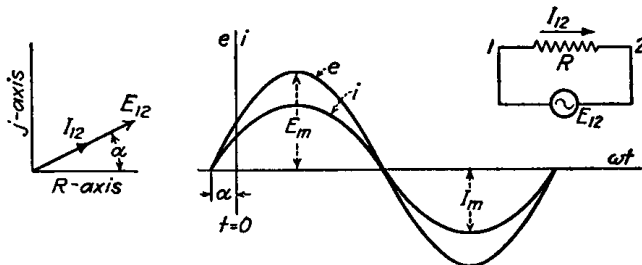
FIGURE 16.19

The vectors representing the voltage and current in a circuit are shown in Fig. 16.19 and may be expressed in complex form as follows:

$$\hat{I} = I\epsilon^{j\alpha} \quad \hat{E} = \hat{I}R = E\epsilon^{j\alpha} \tag{16.75}$$

It may thus be concluded that in a circuit subjected to a sine emf the vector potential drop across the resistance is

$$\hat{E}_r = \hat{I}R \quad \text{or} \quad \hat{I} = \hat{E}_r G \tag{16.76}$$

where $G = 1/R$ is called the *conductance* of the circuit.

26. Circuits Containing Inductance.

In the circuit in Fig. 16.20 let a current be

$$i = I_m \sin(\omega t - \beta) \tag{16.77}$$

Then by Eq. (16.10) the drop across the inductance which is equal to the applied emf is

$$e_L = L\frac{di}{dt} = L\,\omega I_m \cos(\omega t - \beta) = X_L I_m \sin\left(\omega t - \beta + \frac{\pi}{2}\right) \tag{16.78a}$$

$$= E_m \sin\left(\omega t - \beta + \frac{\pi}{2}\right) = E_m \sin(\omega t + \alpha) \tag{16.78b}$$

where
$$E_m = X_L I_m = \omega L I_m \quad \text{and} \quad \alpha = \frac{\pi}{2} - \beta \tag{16.78c}$$

Observe that in this case the voltage wave [compare Eqs. (16.77) and (16.78a)] leads the current wave by $\pi/2$ radians. The current is said to be in *lagging quadrature* with the voltage. The vectors representing the current and voltage, respectively, are shown in Fig. 16.20. Their vector expressions are

$$\hat{I} = I\epsilon^{-j\beta} \tag{16.79a}$$

$$\hat{E} = E\epsilon^{j[(\pi/2)-\beta]} = X_L I \epsilon^{j(\pi/2)}\epsilon^{-j\beta} = jX_L I\epsilon^{-j\beta} = jX_L \hat{I} \tag{16.79b}$$

The angle $\phi = \alpha - \beta = \pi/2$ between the voltage and current waves, as well as between the voltage and current vectors (see Fig. 16.20), is called the *power-factor angle* of the circuit. X_L is called the *inductive reactance* of the circuit and is defined by Eq. (16.78c) as follows:
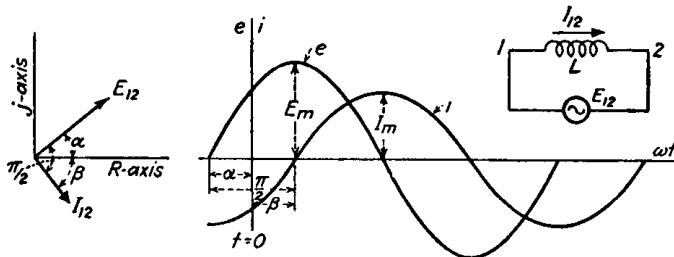
FIGURE 16.20

$$X_L = \omega L = 2\pi f L \tag{16.80}$$

The vector drop across an inductance may therefore be defined, in magnitude and phase position, as follows:

$$\hat{E}_L = jX_L\hat{I} \tag{16.81a}$$

$$= \frac{\hat{I}}{-jB_L} \tag{16.81b}$$

B_L is called the *inductance susceptance* of the circuit and is defined as follows:

$$B_L = \frac{1}{X_L} = \frac{1}{2\pi f L} \tag{16.82}$$

27. Circuits Containing Capacitance. Let the emf impressed on the capacitance C of Fig. 16.21 be

$$e = E_m \sin(\omega t + \alpha) \tag{16.83}$$

The charge accumulating on the capacitor (condenser) C at any instant is, by Eq. (16.17a),

$$q = Ce = CE_m \sin(\omega t + \alpha) \tag{16.84}$$

Hence the current flowing through the circuit is

$$i = \frac{dq}{dt} = C\omega E_m \cos(\omega t + \alpha) = B_c E_m \sin\left(\omega t + \alpha + \frac{\pi}{2}\right) \tag{16.85a}$$

$$= I_m \sin\left(\omega t + \alpha + \frac{\pi}{2}\right) = I_m \sin(\omega t + \beta) \tag{16.85b}$$

where $\quad I_m = B_c E_m = \omega C E_m \quad$ and $\quad \beta = \left(\alpha + \dfrac{\pi}{2}\right) \tag{16.85c}$

Observe that in this case the current i leads the voltage by $\pi/2$ radians or is in *leading quadrature* with the voltage. The vectors representing the current and voltage, respectively, are given in Fig. 16.21 and may be expressed as follows:
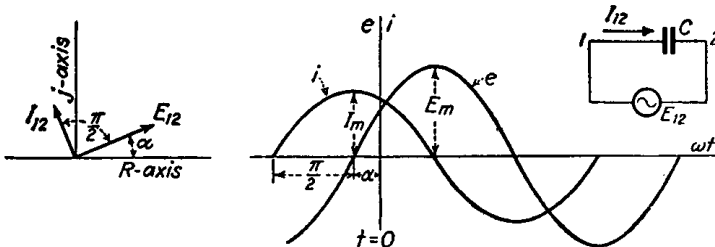
FIGURE 16.21

$$\hat{I} = I\epsilon^{j\beta} = I\epsilon^{j(\alpha+\pi/2)} = B_c E\epsilon^{j\alpha}\epsilon^{j(\pi/2)} = jB_c E\epsilon^{j\alpha} = jB_c\hat{E} \qquad (16.86a)$$

$$\hat{E} = E\epsilon^{j\alpha} \qquad (16.86b)$$

The angle $(\alpha - \beta) = -\pi/2$, between the voltage and current waves, as well as between the voltage and current vectors (see Fig. 16.21), is the *power-factor angle* of the circuit. $B_c = \omega C = 2\pi f C$ is called the *capacitive susceptance* of the circuit, and $jb_c = j\omega C$ is called the *complex capacitive susceptance* of the circuit.

The potential drop across a capacitance may be defined as follows:

$$\hat{E}_c = \frac{\hat{I}_c}{jB_c} = -jX_c\hat{I}_c \qquad (16.87)$$

where X_c and B_c are known, respectively, as the capacitive reactance and capacitive susceptance of the circuit and are defined as

$$X_c = \frac{1}{2\pi f C} \qquad (16.88a)$$

$$B_c = 2\pi f C \qquad (16.88b)$$

28. Series Circuits with _R_, _L_, and _C_.   An emf e is impressed on a series circuit (Fig. 16.22). Let

$$e = E_m \sin(\omega t + \alpha) \qquad (16.89)$$

A current i flows in the circuit whose value is

$$i = I_m \sin(\omega t + \beta) \qquad (16.90)$$

It is required to ascertain the relation between the effective values E and I and to determine the phase angle ϕ between the voltage and current waves or vectors.
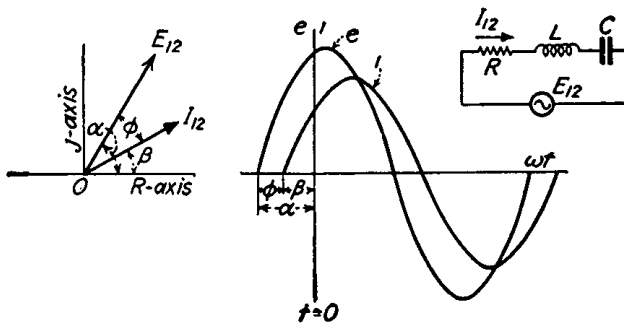
FIGURE 16.22

Solution. The voltage drops across R, L, and C are, by Eqs. (16.76), (16.81), and (16.87),

$$\hat{E}_r = \hat{I}R \tag{16.91a}$$

$$\hat{E}_L = j\hat{I}X_L \tag{16.91b}$$

$$\hat{E}_c = -j\hat{I}X_c \tag{16.91c}$$

By Kirchhoff's emf law,

$$\hat{E} = \hat{E}_r + \hat{E}_L + \hat{E}_c = \hat{I}[R + j(X_L - X_c)] \tag{16.92a}$$

$$\hat{I} = \frac{\hat{E}}{R + j(X_L - X_c)} = \frac{\hat{E}}{\hat{Z}} = \frac{E\epsilon^{j\alpha}}{Z\epsilon^{j\phi}} = \frac{E}{Z}\,\epsilon^{j(\alpha-\phi)} = I\epsilon^{j\beta} \tag{16.92b}$$

$$\hat{Z} = R + j(X_L - X_c) \tag{16.93a}$$

where

$$Z = \sqrt{R^2 + (X_L - X_c)^2} \tag{16.93b}$$

$$\phi = \tan^{-1}\left(\frac{X_L - X_c}{R}\right) \tag{16.93c}$$

Z is known as the *impedance* of the series circuit (Fig. 16.22).

Equation (16.92b) states that the effective value of the current and its phase angle in the series circuit in Fig. 16.22 are, respectively,

$$I = \frac{E}{Z} \tag{16.94a}$$

$$\beta = (\alpha - \phi) \tag{16.94b}$$

The power-factor angle is, by Eq. (16.92b),

$$\phi = \alpha - \beta \tag{16.95}$$

28a. Parallel Circuits with R, L, and C. An emf e is impressed on a parallel circuit (Fig. 16.23). Let

$$e = E_m \sin(\omega t + \alpha) \tag{16.96}$$
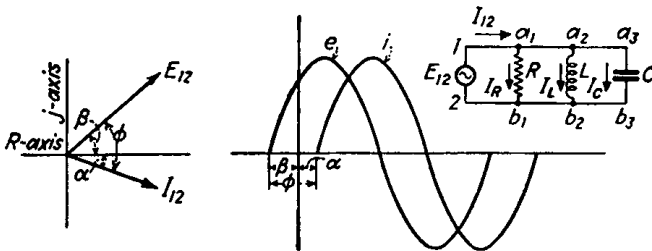
FIGURE 16.23

Since all the characteristics are across the same emf, from Eqs. (16.76), (16.81), and (16.87),

$$\hat{I}_r = \frac{\hat{E}}{R} = \hat{E}G \tag{16.97a}$$

$$\hat{I}_L = \frac{\hat{E}}{jX_L} = -j\hat{E}B_L \tag{16.97b}$$

$$\hat{I}_c = \hat{E}B_c \tag{16.97c}$$

Applying Kirchhoff's current law to the junction a (Fig. 16.23),

$$\hat{I} = \hat{I}_r + \hat{I}_L + \hat{I}_c$$

$$= \hat{E}[G + j(B_c - B_L)] = \hat{E}\hat{Y} = E\epsilon^{j\alpha}Y\epsilon^{j\phi} = EY\epsilon^{j(\alpha+\phi)} = I\epsilon^{j\beta} \tag{16.98}$$

where
$$\hat{Y} = G + j(B_c - B_L) \tag{16.99a}$$

$$Y = \sqrt{G^2 + (B_c - B_L)^2} \tag{16.99b}$$

$$\phi = \tan^{-1}\left(\frac{B_c - B_L}{G}\right) \tag{16.99c}$$

Y is called the *admittance* of the parallel circuit in Fig. 16.23.

Equation (16.98) states that the effective value of the current and its phase angle in the parallel circuit (Fig. 16.23) are, respectively,

$$I = EY. \tag{16.100a}$$

$$\beta = \alpha + \phi \tag{16.100b}$$

The power-factor angle ϕ is., by Eq. (16.99c) and by Fig. 16.23,

$$\phi = \beta - \alpha \tag{16.101}$$

29. Voltage and Current Resonance

Voltage Resonance.   Refer to the series circuit (Fig. 16.22) and observe by Eqs. (16.91) that

$$\dot{E}_L + \hat{E}_c = j\hat{I}(X_L - X_c) \tag{16.102}$$

When the sum of the voltage drops across the inductance and capacitance is nil, the series circuit in Fig. 16.22 is said to be in *resonance.* Obviously, this occurs when

$$X_L - X_c = 0 \quad \text{or} \quad X_L = X_c \tag{16.103}$$

In order to determine the resonant frequency f_r, substitute in Eq. (16.103) for X_L and X_c their values as given by Eqs. (16.80) and (16.88), and obtain

$$2\pi f_r L = \frac{1}{2\pi f_r C} \tag{16.104a}$$

or

$$f_r = \frac{1}{2\pi\sqrt{LC}} \tag{16.104b}$$

Current Resonance. Refer to the parallel circuit (Fig. 16.23) and observe, by Eq. (16.97), that

$$\hat{I}_L + \hat{I}_c = j\hat{E}(B_c - B_L) \tag{16.105}$$

When the sum of the Currents in the inductance and capacitance of a parallel circuit is nil, that circuit is said to be *in resonance*. This occurs when

$$B_c - B_L = 0 \quad \text{or} \quad B_L = B_c \tag{16.106}$$

In order to determine the resonant frequency f_r, substitute in Eq. (16.106), for B_L and B_c, their values as given in Eqs. (16.82) and (16.88), and obtain

$$2\pi f_r C = \frac{1}{2\pi f_r L} \tag{16.107a}$$

or

$$f_r = \frac{1}{2\pi\sqrt{LC}} \tag{16.107b}$$

Comparing Eqs. (16.104b) and (16.107b), it may be concluded that the frequency at which resonance occurs in a series or parallel circuit is the same.

POLYPHASE ALTERNATING-CURRENT CIRCUITS

30. Three-Phase EMF Sources. A three-phase emf source is one having three distinct emf's. Such a three-phase source is said to be *balanced* when (1) the three emf's have the same effective values, and (2) the vectors representing these emf's are equally displaced from each other by $2\pi/3$ radians or 120 degrees. The three emf sources, their waves, and the vectors representing them are shown in Fig. 16.24. Commercial three-phase systems are approximately balanced, and this condition is assumed in the following sections.
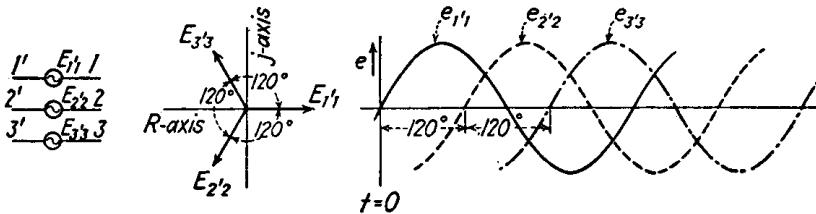
FIGURE 16.24

30a. Methods of Connecting Three-Phase Sources and Three-Phase Loads. Three-phase sources may be connected either in star (Y) or in mesh (Δ). Each of these types of connection is defined below.

Star or Y Connection. Let the terminals 1′, 2′, and 3′ of the emf sources be connected together as in Fig. 16.25. Call the common terminal 0. The voltages then become

$$\hat{E}_{1'1} = \hat{E}_{01} \tag{16.108a}$$

$$\hat{E}_{2'2} = \hat{E}_{02} \tag{16.108b}$$

$$\hat{E}_{3'3} = \hat{E}_{03} \tag{16.108c}$$

The vectors representing these emf's are also shown in Fig. 16.25. The Y connection is so named because of the shape of this vector diagram.

The emf's given in Eq. (16.108) are known as *phase emf's*. If a voltmeter is connected across the terminals 12, 23, and 31 (Fig. 16.25), respectively, it will read what is known as the *line emf's* E_{12}, E_{23}, and E_{31}. The vector relation between the line and phase emf's is obtained directly from Fig. 16.25. Thus

$$\hat{E}_{12} = \hat{E}_{02} - \hat{E}_{01} \tag{16.109a}$$

$$\hat{E}_{23} = \hat{E}_{03} - \hat{E}_{02} \tag{16.109b}$$

$$\hat{E}_{31} = \hat{E}_{01} - \hat{E}_{03} \tag{16.109c}$$

Adding Eqs. (16.109), it is seen that the sum of the line voltages is zero, or

$$\hat{E}_{12} + \hat{E}_{23} + \hat{E}_{31} = 0 \tag{16.110}$$

Again, in the vector diagram (Fig. 16.25), drop a perpendicular from 0 to any of the line voltages. Then, designating the magnitudes of the line and phase voltages by E_L and E_p, respectively,

$$0.5E_L = E_p \sin 60° = E_p \frac{\sqrt{3}}{2} \tag{16.111a}$$

or $$E_L = \sqrt{3}E_p \tag{16.111b}$$
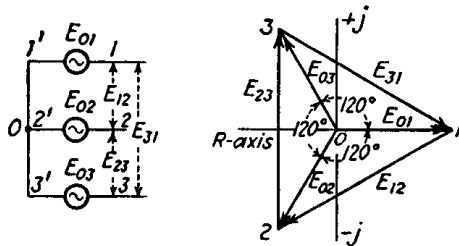
FIGURE 16.25

This is an extremely important relation between the effective values of the phase and line voltages.

The Mesh or Delta (Δ) Connection.   Since the vectors representing the emf's in Fig. 16.24 are all equal and equally displaced from each other by 120°, their sum is zero. Hence the termihals of the emf's may be joined in series (Fig. 16.26) and there will be no residual voltage in the combination. Such a connection is known as the Δ (delta) connection. It is evident, by reference to Fig. 16.26, that the line voltages are

$$\hat{E}_{12} = \hat{E}_{2'2} \tag{16.112a}$$

$$\hat{E}_{23} = \hat{E}_{3'3} \tag{16.112b}$$

$$\hat{E}_{31} = \hat{E}_{1'1} \tag{16.112c}$$

Figure 16.26. also shows the vector diagram of the voltages in Eqs. (16.112).

Star and Mesh Loads.   Consider three equal impedances of value $Z\epsilon^{j\phi}$. These impedances also may be connected in star (Y) or mesh (Δ), as shown in Fig. 16.27a and b.

31. Types of Polyphase Circuits.   There are four types of three-phase circuits designated by the type of connection of source and load, as shown in Fig. 16.28 and described in Table 16.4.

32. Current-EMF Relations.   It can be shown that all the types of three-phase circuits can be converted to a star-star circuit. Hence the emf-current relations for a star-star balanced circuit only are given by
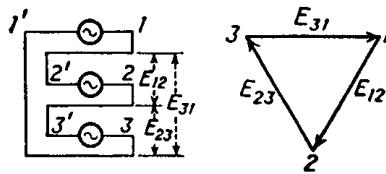
FIGURE 16.26
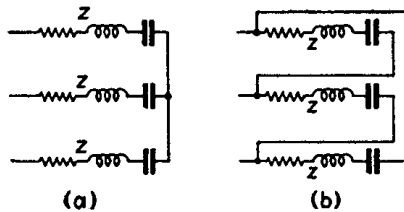
FIGURE 16.27   (a) Impedance connected in star; (b) impedance connected in Δ
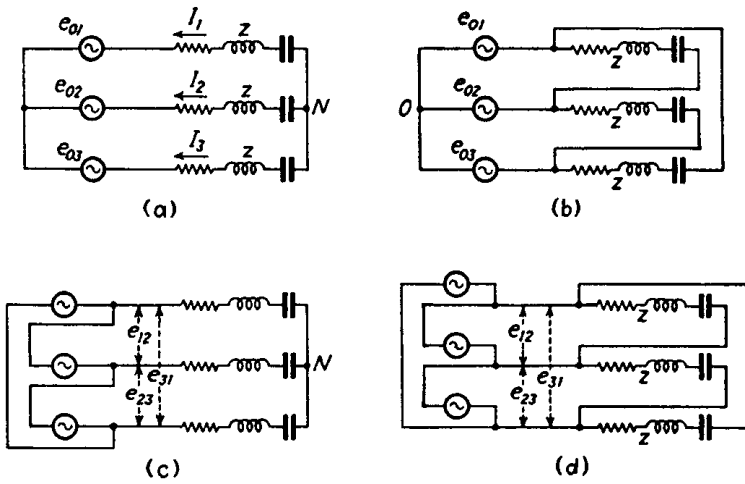
FIGURE 16.28

TABLE 16.4 Types of Polyphase Circuits in Fig. 16.28

Source	Load	Name	Figure
Star	Star	Star-star	16.28a
Star	Mesh	Star-mesh	16.28b
Mesh	Star	Mesh-star	16.28c
Mesh	Mesh	Mesh-mesh	16.28d

$$\hat{I}_1 = \frac{E_{01}}{\hat{Z}} \tag{16.113a}$$

$$\hat{I}_2 = \frac{\hat{E}_{02}}{\hat{Z}} \tag{16.113b}$$

$$\hat{I}_3 = \frac{\hat{E}_{03}}{\hat{Z}} \tag{16.113c}$$

33. Power in AC Circuits. The instantaneous power in an ac circuit is defined as

$$p = ei = E_m \sin (\omega t + \alpha) I_m \sin (\omega t + \alpha \pm \phi) \tag{16.114}$$

where e is the emf applied to the circuit, and i is the total current furnished by this emf source. Using the fact that $\sin x \sin y = \frac{1}{2}[\cos (x - y) - \cos (x + y)]$, Eq. (16.114) becomes

$$p = \frac{E_m I_m}{2} [\cos (\pm \phi) - \cos (2 \omega t + 2\alpha \pm \phi)] \qquad (16.115a)$$

$$= EI [\cos (\phi) - \cos (2 \omega t + 2\alpha \pm \phi) \qquad (16.115b)$$

The average power in an electric circuit is defined as

$$P = \frac{1}{T} \int_0^T p \, dt \qquad (16.116)$$

Substituting Eq. (16.115b) in (16.116) and integrating,

$$P = \frac{EI}{T} \int_0^T \cos \phi \, dt - \int_0^T \cos (2\omega t + 2\alpha \pm \phi) \, dt \qquad (16.117a)$$

$$= EI \cos \phi + 0. \qquad (16.117b)$$

Thus the *average power* in an ac circuit is

$$P = EI \cos \phi \qquad (16.118)$$

The term cos ϕ is known as the *power factor* (P.f.) of the circuit; the product EI is called the *apparent power* (P_a) and is measured in voltamperes (VA) or kilovoltamperes (kVA); the product $EI \phi$ is called the *reactive power* (P_r), also measured in voltamperes or kilovoltamperes. Thus

$$\text{P.f.} = \cos \phi \qquad (16.119)$$

$$P_a = EI \qquad (16.120)$$

$$P_r = EI \sin \phi \qquad (16.121)$$

THE MAGNETIC CIRCUIT

34. Fundamental Concepts. A piece of iron is said to be *magnetized,* to possess *magnetism,* to be in a *magnetic state,* or to be a *magnet* if it is capable of attracting or repelling other pieces of iron placed near it. Magnetism appears to be concentrated at the ends of a magnet. Two magnets placed side by side (Fig. 16.29) may either attract or repel each other. Magnets placed as in Fig. 16.29a will attract each other. The reversal of one of them (Fig. 16.29b) results in repulsion between thcm. It is necessary, therefore, to distinguish between the positive (north) and negative (south) poles of a magnet.
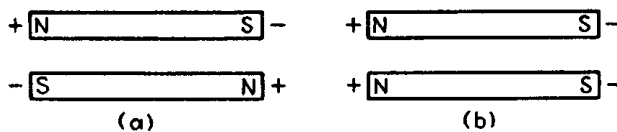
FIGURE 16.29

A *magnetic needle* is a magnetized piece of steel having small mass and so suspended that it can move freely in any of three mutually perpendicular directions. Such a needle is used to detect the existence of magnetism, because a magnetic needle orients itself in a certain direction whenever it is placed near a magnctized object. The magnetic state is not confined to iron; it may exist in air. Thus the space (Fig. 16.30) surrounding a coil carrying a current *I* is in a *magnetic state,* because a magnetic needle placed near such a coil orients itself at each point in the direction tangent to the lines shown in Fig. 16.30. The earth itself is endowed with magnetism and is a huge magnet, which accounts for the fact that a magnetic needle orients itself according to north and south. The terms *north* and *south* as applied to the ends of a magnet originated because the north pole of a magnetic needle points northward. According to this terminology, the geographic north pole of the earth, since it attracts the north pole of a magnetic needle, is in reality the earth's magnetic south pole.

The region in which a magnetic state exists is a *magnetic field.* The magnetism at every point in a magnetic field is characterized by two properties, magnitude and direction. These two properties are defined by one term, *induction,* which is a vector. The magnitude of magnetism at any point is proportional to the force exerted on a magnetic needle placed at that point. The direction is the direction in which the needle is oriented. Curves (Fig. 16.30) whose tangents at every point correspond with the direction of magnetism are known as *lines of force* or *lines of flux.* These lines always form *closed paths.*

35. Graphical Representation of a Field. The induction of a magnetic field varies from point to point in space. Hence it may be represented by a vector whose length is proportional to the magnitude of the magnetism and whose direction is the direction of the lines of force. Such representation, while describing the induction at any one point, does not give a picture of the field simultaneously at several points. Two-dimensional figures are therefore generally used, giving a map of the field in a plane (Fig. 16.31). The *direction* of induction is here represented by the lines of force (flux), while its *magnitude* is indicated by the density of the lines of flux. Thus the number of lines of force drawn across a unit length is a measure of the magnitude of the induction at any point in the plane. This idea of replacing the magnitude of induction by flux density permeates all engineering literature and is a very useful concept. *Induction* may now be defined as a vector having at every point a magnitude equal to the flux density and a direction corresponding to that of a magnetic needle placed at this point.
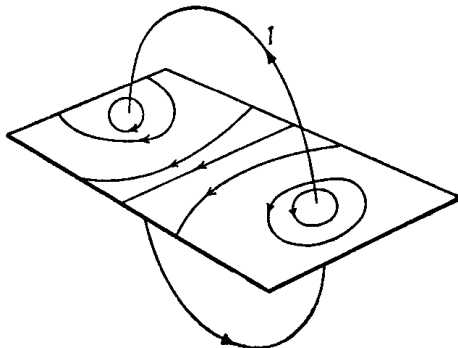
FIGURE 16.30
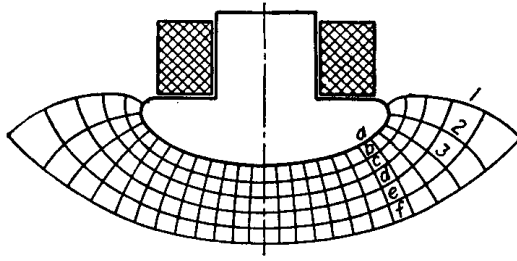
FIGURE 16.31

36. Magnetomotive Force. So far attention has been paid to the magnetic field without inquiring into its cause. The cause of such a field is, or may always be traced to, an electric current. The question as to why there is a field in a magnetized piece of iron or in magnetite when apparently no current exists will be taken up subsequently. Suffice it to say here that an electric current (Fig. 16.30) is always accompanied by a magnetic field whose direction is perpendicular to the direction of the current and whose flux density at any point varies with the strength of the current and the configuration of the current path.

Figure 16.30 shows that the current and field which it produces form links like the links of a chain. Moreover, the directions of the current and of the lines of flux are such that if a right-hand screw is rotated in the direction of current flow, the progress of the screw corresponds to the direction of the flux lines.

A few simple experiments will now be outlined in an effort to determine the relation between cause and effect in a magnetic field. Consider two currents I_1 and I_2 (Fig. 16.32) in two separate wires of identical shape. The fields due to I_1 and I_2 are represented by solid and dotted lines, respectively. If the loops are large and the wires are thin and placed very close to each other, the two fields, for all practical purposes, will merge and may be considered as one.

Let $I_1 = I$, $I_2 = 0$, and the flux density at a point $= \mathfrak{B}_x$. If now $I_1 = 2I$ and $I_2 = 0$, the flux density becomes $2\,\mathfrak{B}_x$. It may be concluded, therefore, that the flux density at any point is directly proportional to the current. Next, let $I_1 = 0.5I$ and $I_2 = 0.5I$. The flux density due to I_1 is $0.5\mathfrak{B}_x$, and that due to I_2 is also $0.5\mathfrak{B}_x$. Hence the total density is $0.5\mathfrak{B}_x + 0.5\mathfrak{B}_x = \mathfrak{B}_x$. But these two currents flowing through two separate loops are electrically equivalent to one current $0.5I$ flowing through *two turns.* Hence it may be concluded that flux density is proportional not only to
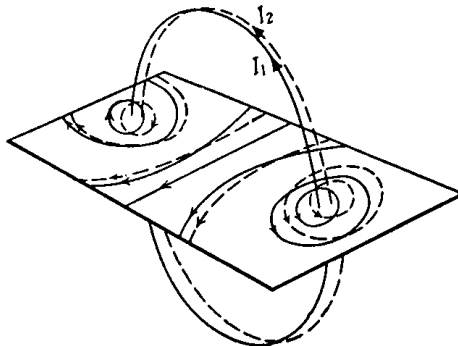
FIGURE 16.32

the current but also to the number of turns this current makes. In other words, the cause of magnetic flux is the *ampere-turns* linking it.

By analogy to electromotive force, which is the cause of electric flow, the term *magnetomotive force* is used to describe the cause of magnetic flow or flux. Magnetomotive force is measured in ampere-turns in the m.k.s. (meter-kilogram-second) system of units.*

37. Magnetic Circuit.   Just as an electric circuit consists of an emf and a current, so a magnetic circuit consists of a magnetomotive force (mmf or $\mathcal{M}$) and a flux. The study of the numerical relations between cause and effect in an electric circuit constitutes the subject of *electric-circuit analysis.* Similarly, the study of the numerical relations between cause and effect in a magnetic circuit is known as *magnetic-circuit analysis.* It must here be emphasized that while in most practical problems of electric-circuit analysis the flow of electricity is uniform and uniformly distributed over the path of flow, such is not the case in the majority of problems of magnetism. In general, a magnetic field has different values of induction at different points. The study of the magnetic circuit, therefore, presents extremely complicated problems. The following discussion is restricted generally to very simple circuits. Where complications do arise, simplifying assumptions will be made.

37a. Magnetic Circuit with Uniform Circular Field.   The lines of flux form closed paths. Hence the direction of a line of force can never be constant over all its path. The flux density, however, may be constant thrdughout the whole field. Thus the simplest field is one having constant flux density and variable direction. The magnetic circuit (Fig. 16.33) in which the flux density has constant magnitude (induction) and circular direction will be studied, with attention only to numerical relations, regardless of the direction of flux.
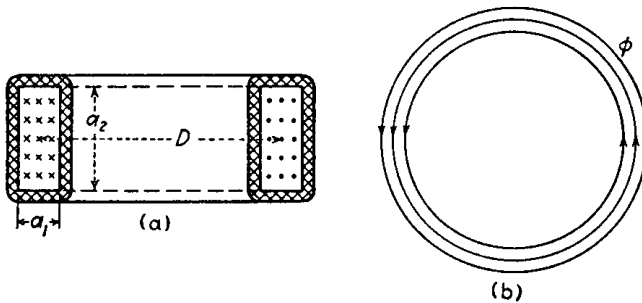
FIGURE 16.33

---

*The m.k.s. system of units was originally proposed by G. Giorgi in 1902. It is based on the meter, the kilogram, and the second as fundamental units. The simplicity of the system for engineering work can be seen from the following table, which gives the values of seven practical electrical units in terms of the c.g.s. units and the m.k.s. units:

	Length	Mass	Time	Coulomb	Ampere	Volt	Ohm	Henry	Farad	Weber
C.g.s.	1 cm	1 g	1 s	10^{-1}	10^{-1}	10^{8}	10^{9}	10^{-9}	10^{-9}	10^{8}
M.k.s.	1 m	1 kg	1 s	1	1	1	1	1	1	

The mmf producing the field in Fig. 16.33 consists of one layer having N turns of very thin wire wound over a circular form of rectangular cross section and carrying a continuous current L The mmf is, therefore,

$$\mathfrak{M} = kNI \tag{16.122}$$

If the diameter of the ring is assumed to be large in relation to its radial thickness a_1, this mmf produces a uniform field of constant flux density $\mathfrak{B}$. Moreover, since $\mathfrak{B}$ is proportional to $\mathfrak{M}$,

$$\mathfrak{M} = k\mathfrak{B} \tag{16.123a}$$

when

$$k = k_1 v \pi D \tag{16.123b}$$

k_1 is a factor affected only by the units used and has nothing to do with the physical dimensions or the material of the magnetic ring.

Substituting (16.123b) in (16.123a),

$$\mathfrak{M} = k_1 v \pi D \mathfrak{B} \tag{16.124}$$

Now let ϕ be the total flux, and let $A = a_1 a_2$ be the cross-sectional area of the toroid. Since the flux density $\mathfrak{B}$ is constant,

$$\phi = A\mathfrak{B} \quad \therefore \quad \mathfrak{B} = \frac{\phi}{A} \tag{16.125}$$

hence

$$\mathfrak{M} = \frac{k_1 v \pi D}{A} \phi = \mathfrak{R}\phi \tag{16.126a}$$

where

$$\mathfrak{R} = \frac{k_1 v \pi D}{A} \tag{16.126b}$$

$\mathfrak{R}$ in Eq. (16.126) is known as the *reluctance* of the flux path and is analogous to resistance in the case of electric circuits, and v is called the *reluctivity* of the flux path, or the reluctance of a path having unit length and unit sectional area.

The analogy, between $\mathfrak{M} = \mathfrak{R}\phi$ in a magnetic circuit with uniform circular field and $E = RI$ in an electric circuit is complete. Hence Eq. (16.126a) is often referred to as *Ohm's law for the magnetic circuit.* Reference will hereafter be made to the term $\mathfrak{R}\phi$ as the *mmf drop.*

Equation (16.12a) may be written in another form:

$$\phi = \frac{1}{\mathfrak{R}} \mathfrak{M} = \mathfrak{P}\mathfrak{M} \tag{16.127a}$$

where

$$\mathfrak{P} = \frac{1}{\mathfrak{R}} = \frac{1}{k_1} \mu \frac{A}{\pi D} \tag{16.127b}$$

$\mathfrak{P}$ is known as the *permeance* of the flux path, and $\mu = 1/v$ is its *permeability.* By definition, μ is the permeance of a flux path of unit length and unit sectional area.

38. Point Relations.   The mmf (Fig. 16.33) is entirely consumed in forcing the flux against the reluctance of the magnetic path. The flux density is uniform, and the magnetic path has a uniform reluctance. Since the mmf drop $\mathfrak{R}\phi$ is uniformly

distributed throughout the whole length of the lines of force, the mmf drop per unit length may be defined as

$$\mathcal{K} = \frac{\mathcal{R}\phi}{\pi D} \tag{16.128a}$$

$$= \left(\frac{k_1 v \pi D}{A}\right)\left(\frac{\mathcal{B}A}{\pi D}\right) = k_1 v \mathcal{B} \tag{16.128b}$$

$\mathcal{K}$ is known as the *mmf drop per unit length,* the *field gradient,* or the *field intensity.* It is a vector whose magnitude is proportional to that of $\mathcal{B}$ and whose direction is along $\mathcal{B}$. Both $\mathcal{B}$ and $\mathcal{K}$ are quantities defined at every point of a magnetic field. Hence Eq. (16.128b) expresses a *point relation* and holds true at every point of a field, irrespective of whether the field is uniform or not.

39. Units.   The relations developed above are general and hold true independently of the units used. For purposes of numerical computations, however, it is necessary to use a consistent set of units, as listed in Table 16.5.

   With this choice of units, the constants k_1, μ, and v in Sec. 37a, for air or vacuum, become

$$k_1 = 1$$

$$\mu_0 = (4\pi)10^{-9} = 1.257 \times 10^{-8} \text{ henry per cm}^3 \tag{16.129}$$

$$= 3.195 \times 10^{-8} \text{ henry per in}^3$$

$$v_0 = (\tfrac{1}{4}\pi)10^9 = 0.795 \times 10^3 \text{ yrneh per cm}^3 = 0.313 \times 10^3 \text{ yrneh per in}^3$$

and the relations of Secs. 37 and 38, for an air core, become

$$\mathcal{M} = \mathcal{R}_0\phi \quad (a) \qquad \mathcal{K} = v_0\mathcal{B} \quad (b) \qquad \mathcal{R}_0 = v_0\frac{l}{A} \quad (c)$$

$$\phi = \mathcal{M}\mathcal{P}_0 \quad (a') \qquad \mathcal{B} = \mu_0\mathcal{K} \quad (b') \qquad \mathcal{P}_0 = \mu_0\frac{A}{l} \quad (c') \tag{16.130}$$

The relations in Eqs. (16.130b) and (16.130b′) are point relations and hold true irrespective of whether the field is uniform or not. The remaining equations in Eqs. (16.130) apply only to a uniform air field wherein l is the length of the flux path and A is the cross-sectional area perpendicular to the direction of flux.

40. Series and Parallel Reluctances and Permeances.   Two permeances are in parallel if the same mmf acts on them. Two reluctances are in series if the same flux path passes through them. Expressions for the equivalent permeance of a parallel circuit and the equivalent reluctance of a series circuit are developed as follows.

Equivalent Permeance.   Consider a set of *n* permeances in parallel. Let the flux through each of them be

TABLE 16.5 Physical Quantities, Symbols, and Units

Symbol	Physical quantity	Equation	Dimension	Name of unit
l	Length	Fund. unit	[L]	Centimeter
t or T	Time	Fund. unit	[T]	Second
μ	Permeability	$\mu = \dfrac{\mathcal{B}}{\mathcal{H}}$	$[RTL^{-1}]$	Henry per cm. cube
I, i	Current	Fund. unit	[I]	Ampere
R, r	Resistance	Fund. unit	[R]	Ohm
E, e	E.m.f.	$E = RI$	[RI]	Volt
A	Area	$A = l_1 l_2$	$[L^2]$	Sq. cm.
V	Volume	$V = l_1 l_2 l_3$	$[L^2]$	Cu. cm.
L	Inductance	$e = -L\dfrac{di}{dt}$	[RT]	Henry
ν	Reluctivity	$v = \dfrac{1}{\mu}$	$[R^{-1}T^{-1}L]$	Yrneh per cm. cube
ϕ	Magnetic flux	$e = \dfrac{d\phi}{dt}$	[RIT]	Weber
$\mathcal{B}$	Flux density	$\mathcal{B} = \dfrac{\phi}{A}$	$[RITL^{-2}]$	Weber per sq. cm.
$\mathcal{M}$	M.m.f.	$\mathcal{M} = NI$	[I]	Ampere-turn
$\mathcal{H}$	Magnetic intensity	$\mathcal{H} = \dfrac{\mathcal{M}}{l}$	$[IL^{-1}]$	Amp.-turns per cm.
P	Power	$P = EI$	$[RI^2]$	Watt
W	Energy	$W = PT$	$[RI^2T]$	Joule
$\mathcal{R}$	Reluctance	$\mathcal{R} = \dfrac{vl}{A}$	$[R^{-1}T^{-1}L^{-1}]$	Yrneh
$\mathcal{P}$	Permeance	$\mathcal{P} = \dfrac{\mu A}{l}$	[RTL]	Henry

$$\phi_1 = \mathcal{M}\mathcal{P}_1 \quad \phi_2 = \mathcal{M}\mathcal{P}_2 \quad \phi_n = \mathcal{M}\mathcal{P}_n \tag{16.131}$$

The total flux in the circuit is

$$\phi = \phi_1 + \phi_2 + \cdots + \phi_n = \mathcal{M}(\mathcal{P}_1 + \mathcal{P}_2 + \cdots \mathcal{P}_n) = \mathcal{M}\mathcal{P} \tag{16.132a}$$

where

$$\mathcal{P} = \mathcal{P}_1 + \mathcal{P}_2 + \cdots + \mathcal{P}_n \tag{16.132b}$$

Thus in a parallel circuit the equivalent permeance is the sum of the parallel permeances.

Equivalent Reluctance.   Consider next a set of reluctances in series. Let the mmf drop across each of them be

$$\mathcal{M}_1 = \mathcal{R}_1\phi \quad \mathcal{M}_2 = \mathcal{R}_2\phi \quad \mathcal{M}_n = \mathcal{R}_n\phi \tag{16.133}$$

Then the total mmf of the circuit is

$$\mathfrak{M} = \mathfrak{M}_1 + \mathfrak{M}_2 + \cdots + \mathfrak{M}_n = (\mathfrak{R}_1 + \mathfrak{R}_2 + \cdots + \mathfrak{R}_n)\phi = \mathfrak{R}\phi \quad (16.134a)$$

where $$\mathfrak{R} = \mathfrak{R}_1 + \mathfrak{R}_1 + \mathfrak{R}_2 + \cdots + \mathfrak{R}_n \quad (16.134b)$$

Thus in a series circuit the equivalent reluctance is the sum of the reluctances.

41. Magnetic Properties of Matter. According to its magnetic properties, matter falls into either one of the following classes:

1. Diamagnetic matter
2. Paramagnetic matter
   a. Ferromagnetic
   b. Nonferromagnetic

Diamagnetic matter includes substances whose magnetic permeability is less than μ_0. Among these may be mentioned antimony and bismuth. The permeability of bismuth, which is one of the most highly diamagnetic substances known, is 1.254×10^{-8} H/cm^3, which is not very different from $\mu_0 = 1.257 \times 10^{-8}$.

Paramagnetic matter includes substances whose magnetic permeability is equal to or greater than μ_0. Air, wood, copper, brass, etc. have a permeability which is approximately μ_0. Iron and iron alloys and nickel and cobalt and their alloys have permeabilities varying from μ_0 to $10^5 \mu_0$, depending on the flux density used and the composition of the alloy. This latter class of paramagnetic matter which possesses high permeability is referred to as *ferromagnetic* in order to distinguish it from *nonferromagnetic* matter, whose permeability is very close to that of vacuum.

The absence of materials of very high reluctivity accounts for the fact that there is no known magnetic insulator. Such an insulator, if it existed, would have a very low permeability or very high reluctivity. The reluctance of a path made of this material would be high, and hence ϕ in the relation $\mathfrak{M} = \mathfrak{R}\phi$ would be low. In order to appreciate this fact, compare the ratio of the resistivity of insulators to that of conductors with the ratio of the reluctivity of paramagnetic or diamagnetic materials to that of ferromagnetic materials as follows:

Resistivity of silver: $\rho_1 = 1.63 \times 10^{-6}$
Resistivity of gutta-percha: $\rho_2 = 3.7 \times 10^{14}$
Reluctivity of permalloy: $v_1 = 2 \times 10^{-5}$
Reluctivity of bismuth: $v_2 = 8 \times 10^7$

$$\frac{\rho_2}{\rho_1} = \frac{3.7 \times 10^{14}}{1.63 \times 10^{-6}} = 2.27 \times 10^{20}$$

$$\frac{v_2}{v_1} = \frac{8 \times 10^7}{2 \times 10^{-5}} = 4 \times 10^{12}$$

From this comparison it is evident that while the variation in resistivity covers a very wide range, this is not true of reluctivity.

42. Magnetization or B-H Curves of Steel and Its Alloys.

Some magnetization curves are shown in Fig. 16.34. Some observations relative to these curves should be pointed out because of their significance in the solution of magnetic-circuit problems.

1. The curves are not straight lines. This means that the permeability of iron μ is not constant but is a function of the flux density $\mathcal{B}$ or the magnetic intensity $\mathcal{H}$.

2. On account of (1), the flux ϕ in a magnetic circuit does not obey the principle of superposition. In other words, if $\mathcal{M}_1$ and $\mathcal{M}_2$ produce fluxes ϕ_1 and ϕ_2, respectively, in an iron path, and if $\mathcal{M}_1 = a\mathcal{M}_2$, then $\phi_1 \neq a\phi_2$. This fact is of extreme importance and has far-reaching consequences, some of which will be pointed out in the succeeding section.

43. Magnetic Circuits with Iron.

Three types of magnetic circuits may be recognized: (1) series magnetic circuit (Fig. 16.35a), (2) parallel circuit (Fig. 16.35b), and (3) series-parallel circuit (Fig. 16.35c). Two types of problems occur in each of these types of circuit: (A) given the flux ϕ; one must find the mmf $\mathcal{M}$, and (B) given the mmf $\mathcal{M}$, one must find the flux ϕ.

The solutions of these two problems are discussed below for each type of circuit, with, in each case, the following assumptions: (1) the flux path follows the path marked b_k (Fig. 16.35) and consists of arcs of circles and straight lines, (2) there exists no leakage flux, so that in any series circuit the total flux is the same in each part of the path, (3) there exists no fringing of flux at the air gaps where two iron paths meet, and (4) the flux density is uniform in any one element of the circuit.

Case A₁: Given ϕ, Required $\mathcal{M}$(Series Circuit). Referring to Fig. 16.35a, let the various drops in the series circuit be $\mathcal{H}_1 b_1$, $\mathcal{H}_2 b_2$, . . . , $\mathcal{H}_6 b_6$. Then the required mmf is
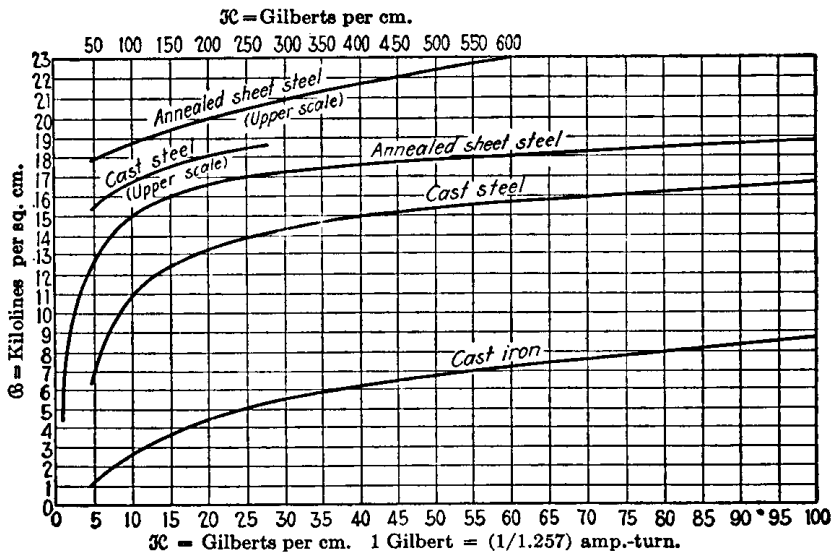
FIGURE 16.34 Magnetization or $\mathcal{B}$–$\mathcal{H}$ curves

FIGURE 16.35

$$\mathfrak{M} = \mathfrak{K}_1 b_1 + \mathfrak{K}_2 b_2 + \cdots + \mathfrak{K}_6 b_6 \tag{16.135}$$
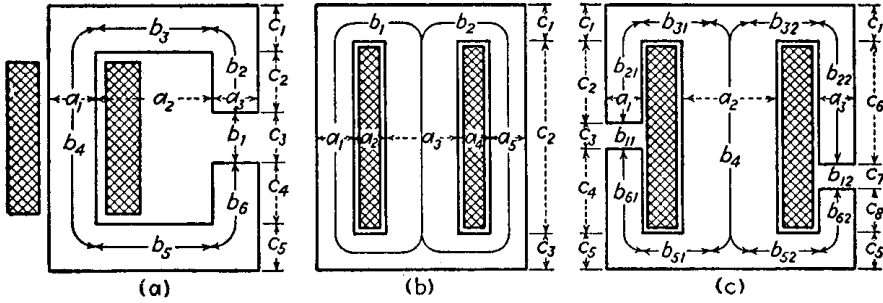
Since ϕ is given, the flux density in the various sections of the series circuit may be found from the relation $\mathfrak{B}_k = \phi/A_k.$. Knowing $\mathfrak{B}_k$, the value of $\mathfrak{K}_k$ may be found from the magnetization curves (Fig. 16.34) for the iron sections and from the relation $\mathfrak{B} = \mu_0 \mathfrak{K}$ for the air sections. Thus all the $\mathfrak{K}$'s in Eq. (16.135) are known, and $\mathfrak{M}$ can be computed.

Case A_2: Given ϕ, Required $\mathfrak{M}$(Parallel Circuit). Referring to Fig. 16.35b, let it be assumed that the total flux ϕ divides into two parts ϕ_1 and ϕ_2 which traverse the paths b_1 and b_2, respectively. Then the drops through the two parallel branches of the magnetic circuit are equal, or

$$\mathfrak{M} = \mathfrak{K}_1 b_1 = \mathfrak{K}_2 b_2 \tag{16.136}$$

Two subcases may be considered:

1. ϕ, ϕ_1, ϕ_2 are all known. Here $\mathfrak{B}_1$ and $\mathfrak{B}_2$, and therefore $\mathfrak{K}_1$ and $\mathfrak{K}_2$, are known. Thus $\mathfrak{M}$ in Eq. (16.136) becomes known and the problem is solved.

2. ϕ is known, but ϕ_1 and ϕ_2 are unknown. Here $\mathfrak{M}$ cannot be found in a straightforward manner. It has to be determined by trial. The work is somewhat simplified through the following considerations: Since $\mathfrak{M} = \mathfrak{K}_1 b_1 = \mathfrak{K}_2 b_2$,

$$\mathfrak{K}_1 = \frac{b_2}{b_1} \mathfrak{K}_2 \tag{16.137a}$$

also

$$\mathfrak{B}_1 = \frac{\phi}{A_1} \quad \text{and} \quad \mathfrak{B}_2 = \frac{\phi_2}{A_2} = \frac{\phi - \phi_1}{A_2} \tag{16.137b}$$

In order to solve the problem, assume values of ϕ_1 and compute $\mathfrak{B}_1$ and $\mathfrak{B}_2$. Find $\mathfrak{K}_1$ and $\mathfrak{K}_2$ from the magnetization curves (Fig. 16.34). If these satisfy Eq. (16.137a), the problem is solved. Otherwise try again.

Case A_3: Given ϕ, Required $\mathfrak{M}$ (Series-Parallel Circuit). This case (Fig. 16.35c) also divides itself into two parts.

1. The flux in each of the branches (Fig. 16.35c) is known. Here the flux densities $\mathfrak{B}_k$ and therefore the intensities $\mathfrak{K}_k$ become known, and the problem simply reduces to adding the various drops $\mathfrak{K}_k b_k$.

2. Only the total flux ϕ is known, and the branch fluxes ϕ_k are not known. Here a straightforward solution becomes impossible. The problem can be solved only by assuming values of ϕ_1 and ϕ_2, computing the various drops in the series-parallel branches, and testing the result by the following relation:

$$\mathfrak{M} = \Sigma\mathfrak{H}_{k1}b_{k1} = \Sigma\mathfrak{H}_{k2}b_{k2} \tag{16.138}$$

where $\mathfrak{H}_{k1}$ are the various intensities in branch 1, and $\mathfrak{H}_{k2}$ are the various intensities in branch 2.

Case B_1: Given $\mathfrak{M}$, Required ϕ (Series Circuit). This case can be solved only by trial. Thus, referring to Fig. 16.35a,

$$\mathfrak{M} = \Sigma\mathfrak{H}_k b_k \tag{16.139}$$

The solution consists of assuming values of ϕ and finding the flux densities $\mathfrak{B}_k$. From the magnetization curves (Fig. 16.34), the corresponding $\mathfrak{H}_k$ can be obtained. The substitution of these values in Eq. (16.139) should yield the given $\mathfrak{M}$ if the correct value of ϕ has been assumed. Various values of ϕ are assumed until Eq. (16.139) is satisfied.

A special case arises when the series circuit has just two elements, one of which is iron with uniform flux density $\mathfrak{B}_i$ and the other of which is an air gap having the same flux density ($\mathfrak{B}_a = \mathfrak{B}_i$). Here Eq. (16.139) reduces to

$$\mathfrak{M} = \mathfrak{H}_i b_i + \mathfrak{H}_a b_a \tag{16.140a}$$

$$= \mathfrak{H}_i b_i + \mu_0 \mathfrak{B}_a b_a = \mathfrak{H}_i b_i + \mu_0 \mathfrak{B}_i b_a \tag{16.140b}$$

Equation (16.140b) is the equation of a straight line in the two unknowns $\mathfrak{H}_i$ and $\mathfrak{B}_i$. The other equation involving these two unknowns is the magnetization curve of the iron. Hence the required values of $\mathfrak{B}_i$ and $\mathfrak{H}_i$ are the coordinates of the point of intersection of the straight line [Eq. (16.140b)] and the magnetization curve (Fig. 16.34). Having determined $\mathfrak{B}_i$, the value of ϕ can be obtained from the expression $\phi = \mathfrak{B}_i A_i$, where A_i is the cross-sectional area of the iron.

Case B_2: Given $\mathfrak{M}$, Required ϕ (Parallel Circuit). Referring to Fig. 16.35b,

$$\mathfrak{M} = \mathfrak{H}_1 b_1 = \mathfrak{H}_2 b_2 = \cdots = \mathfrak{H}_k b_k \tag{16.141a}$$

whence

$$\mathfrak{H}_k = \frac{\mathfrak{M}}{b_k} \tag{16.141b}$$

Having obtained $\mathfrak{H}_k$, determine $\mathfrak{B}_k$ from the magnetization curve and obtain ϕ from the relation $\phi = \mathfrak{B}_k A_k$.

Case B_3: Given $\mathfrak{M}$, Required ϕ (Series-Parallel Circuit). The solution of this problem is accomplished by trial. Thus various values of ϕ are assumed, and the corresponding values of the flux densities $\mathfrak{B}_k$ are found. Then from the magnetization curves the values $\mathfrak{H}_k$ are determined. The relation which must be satisfied by each branch of the series-parallel circuit is $\mathfrak{M} = \mathfrak{H}_k b_k$.

Conclusion. It will be observed that the only magnetic circuits which possess a straightforward solution are those discussed under cases A_1 and B_2. The remaining circuits are solvable only by trial. This emphasizes the difference between electric and magnetic circuits—a difference which is, in a great measure, due to the fact

that magnetic materials do not obey the law of flux superposition. In other words, the magnetization curves (Fig. 16.34) are not straight lines.

44. Hysteresis.   If the magnetizing force (Fig. 16.36) is gradually lowered from its maximum value, the $\mathcal{B}$–$\mathcal{H}$ curve will not be retraced. When the magnetizing force is reduced to zero, there will still be a considerable value of magnetic density $\mathcal{B}$. This is known as *residual magnetism.* The negative magnetizing force necessary to bring the density to zero is the *coercive force.* By continuing to build up the magnetomotive force in the negative direction and then again reversing it, as in the curve of Fig. 16.36, the *hysteresis loop* or *hysteresis curve* is produced. The area of this loop is a measure of the energy consumed in carrying the iron through the cycle of magnetization.

When a piece of iron is put through a cycle of positive and negative magnetization, a certain amount of energy is lost in the form of heat. This loss is measured by the area of the hysteresis curve (Fig. 16.36). Whenever iron is subjected to the magnetization effect of an alternating current, this cycle of hysteresis loss occurs during every period of the alternating current. Experirnentally, it can be shown that the hysteresis loss P_h, in watts, varies as a power h (less than 2) of the magnetic density $\mathcal{B}_m$, as the volume V of the iron, as its magnetic quality k, and as the frequency f. Hence the loss may be expressed by the following relation:

$$P_h = kfV\mathcal{B}_m^h 10^{-7} \tag{16.142}$$

The exact data on all iron losses for the steel to be used should be obtained from the maker. The loss is lowered by alloying steel with silicon, thus forming what is known as silicon steel.

45. Eddy Currents.   Whenever a conductor is placed in a varying magnetic field, emf's are set up in the conductor. These cause local currents known as *eddy currents* which produce heat and are a source of loss. To keep these losses down, the iron
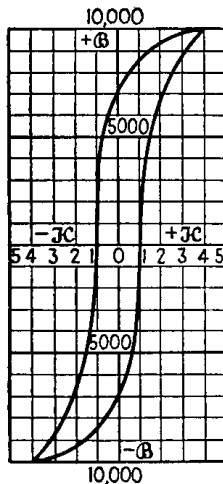
FIGURE 16.36   Hysteresis loop.

of a magnetic circuit, which carries the variable flux, is made up of thin laminations. The eddy-current loss P_{ee} in thin laminations is

$$P_{ee} = kf^2 V \mathscr{B}_m^2 t^2 10^{-6} \tag{16.143}$$

where P_{ee} is the eddy-current loss in watts, t is the thickness of the laminations, and the other terms are as defined in Eq. (16.142).

Equation (16.143) shows the way in which the eddy-current loss varies with the various factors involved. Uncertainty with regard to the value of k, however, makes it impracticable to use the formula for quantitative data; hence test results are generally furnished by the manufacturer.

46. Magnetic Pull or Traction; Magnetic Mechanisms.   When there is a cut at right angles across the iron of a magnetic circuit and the two faces are in close contact, the magnetic pull which is exerted between the faces is equal to $\mathscr{B}^2 A/(8\pi \times 981)$ g or $\mathscr{B}^2 A/11$, 183,000 lb (lb $\times$ 4.45 = N) ($\mathscr{B}$ in gausses or lines per square centimeter, A in square centimeters). In such devices as magnetic clutches, face-plates, brakes, and hoists, this expression for traction makes the design comparatively simple. Where, however, the pull is called on to move part of the mechanism, the presence of leakage around the necessary gap complicates the calculation. The shorter the gap, the more nearly will the pull come up to that given by the formula. It is, therefore, desirable to make use of levers and other devices to shorten the required distance of travel of the magnet armature. One method of doing this is to use a conical pole piece, with an armature through which a corresponding conical hole is bored. A desirable form is the iron-clad magnet, in which one pole consists of an iron cylinder surrounding the coil, while the other pole constitutes the core of the coil. The armature is then a circular disk (see Fig. 16.37).

47. Magnetic Force Produced by Current-Carrying Conductors.   Since whenever magnetic fields come in contact with each other a force is exerted between them, and since all current-carrying conductors are surrounded by magnetic fields, it follows that a force will be exerted by every current-carrying conductor on any other similar conductor and also on any magnetic pole in its vicinity. Conversely, any conductor carrying current and situated in a magnetic field will experience a pull at right angles to the conductor and the field. This pull is equal to $10.2\mathscr{B}I \times 10^{-8}$ kg/cm length when the conductor carries a current I and is in a field of $\mathscr{B}$ gauss. The operation of the electric motor is based on this fact. Two conductors of length l lying parallel to each other at a distance of d cm and carrying currents I_1 and I_2, respectively, will experience a pull of
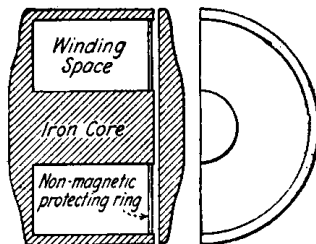
FIGURE 16.37   Iron-clad magnet.

$$\frac{(2.04 \times 10^{-8}I_1I_2l)}{d} \quad (\text{kg} \times 9.81 = N)$$

When the conductors are surrounded by iron, the pull will, of course, be much greater. These magnetic forces may be great enough to do serious damage to electrical equipment, especially in the event of short circuits.

ELECTROSTATIC CIRCUIT

48. The Electrostatic Charge. When an emf E (Fig. 16.5a) is applied to the plates of a condenser (see Sec. 7) with a capacity C, an electrostatic field is established between these plates owing to the accumulation of negative and positive charges on their surfaces. The amount of charge Q that accumulates on the plates depends on (1) the potential E impressed on the plates, (2) the area A of the plates, (3) the distance d between the plates, (4) the kind of dielectric occupying the space between the plates, and (5) the time that elapses from the instant of impressing E to the instant of measuring the charge Q.

In most commercial dielectrics, the variation of charge with each of these factors can be ascertained only by experiment. Indeed, there seems to be no general law which all dielectrics do obey. In order to simplify this discussion, assume a perfect dielectric, i.e., one for which there exists a linear relation between Q and E, and Q and A, as well as between Q and $1/h$. Moreover, such a perfect dielectric will have an infinite resistance so that no conduction current will flow through it. Finally, in such a dielectric, Q is independent of time, so that the dielectric will be assumed to take its full charge instantly. With the foregoing assumptions (which, except for time, hold true for air and vacuum), the following relation can be set forth as an empirical result for a two-plate condenser:

$$Q = CE = (kk_a)\left(\frac{A}{h}\right)E \tag{16.144}$$

where k = dielectric constant (permittivity) of the material
 k_a = dielectric constant of air
 A = area of one plate
 h = distance between the plates
 E = applied potential

From Eq. (16.144) the constant C, which is known as the *capacitance* of a condenser, may be determined from the dimensions of the condenser. Thus

$$C = kk_a\frac{A}{h} \tag{16.145}$$

49. Elastance as the Reciprocal of Capacitance. Another form of writing Eq. (16.144) is as follows:

$$E = \frac{Q}{C} = SQ = \left(\frac{1}{kk_a}\right)\frac{h}{A}Q \tag{16.146}$$

where

$$S = \frac{1}{C} = \left(\frac{1}{kk_a}\right)\frac{h}{A} = \sigma\sigma_a\frac{h}{A} \qquad (16.147)$$

The symbol S stands for the elastance of a condenser and corresponds to resistance in the electrodynamic circuit. From Eq. (16.147) one observes that S is directly proportional to the distance between the plates of a condenser and inversely proportional to the area of any one plate.

50. Units. The preceding laws are true irrespective of the choice of units. In order to apply them to numerical problems, however, a definite system of units is necessary. The practical system of units, as given in Table 16.4, are recommended. The constants σ_a and k_a for this system of units have the following values:

$$\sigma_a = 11.3 \times 10^{12} \text{ darafs per cm}^3 = 4.45 \times 10^{12} \text{ darafs per in}^3$$

$$k_a = 0.08842 \times 10^{-12} \text{ farad per cm}^3 = 0.2244 \times 10^{12} \text{ farad per in}^3$$

51. Charge Density and Potential Gradient. In the electrodynamic circuit, the concept of current density is helpful in the solution of problems, so here the concept of the charge density D is introduced, which will be defined as follows:

$$D = \underset{\Delta A \to 0}{\text{limit}} \left(\frac{\Delta Q}{\Delta A}\right) \qquad (16.148a)$$

When the charge Q is equally distributed over the area of the dielectric, Eq. (16.148a) becomes

$$D = \frac{Q}{A} \qquad (16.148b)$$

The voltage gradient G is defined as

$$G = \underset{\Delta h \to 0}{\text{limit}} \left(\frac{\Delta E}{\Delta h}\right) \qquad (16.149a)$$

and for a linear variation of E with h, Eq. (16.149a) becomes

$$G = \frac{E}{h} \qquad (16.149b)$$

The following interesting and useful relationship exists between G and D:

$$D = \frac{Q}{A} = \frac{CE}{A} = \left(\frac{kk_aA}{h}\right)\left(\frac{E}{A}\right) = kk_aG \qquad (16.150a)$$

$$G = \left(\frac{1}{kk_a}\right)D = (\sigma\sigma_a)D \qquad (16.150b)$$

52. Dielectric Strength. The *strength* of a dielectric is defined as the voltage it can stand per unit thickness (without breaking down) under certain specified conditions of temperature, humidity, time of application of potential, etc. The mechanism of breakdown in dielectrics is still unknown and offers a great field for research. But whatever the mechanism may be, it is a commonly observed fact that

dielectrics do break down when the potential applied to them exceeds a certain fixed value known as the *dielectric strength* of the material. This phenomenon finds its analogue in the breakdown of materials such as steel due to the application of a stress exceeding their elastic limit.

Values of dielectric constants and dielectric strengths of various insulating materials are given in Tables 16.6 and 16.7. These values are to be used with great caution, however, because the conditions under which they were obtained may vary materially from the conditions of a given problem.

SOURCES OF EMF: GENERATORS

53. Electromagnetic Induction. Faraday discovered that when relative motion exists between a metallic conductor and a magnetic field, an emf is induced in the conductor. The magnitude of this emf is

$$e = \mathcal{B}vl \qquad (16.151)$$

where e = voltage induced between the terminals of the conductor
  $\mathcal{B}$ = flux density of the magnetic field in a direction perpendicular to the axis of the conductor
  l = length of the conductor
  v = relative velocity of the conductor and the magnetic field

The essential limitations of Eq. (16.151) are as follows:

1. $\mathcal{B}$, v, and I shall be directed along three mutually perpendicular axes, as shown in Fig. 16.38.

TABLE 16.6 Dielectric Constants

Material	Dielectric constant k	Material	Dielectric constant k
Glass (easily fusible)	2.0 to 5.0	Air and other gases	1.0
Glass (difficult to fuse)	5.0 to 10.0	Alcohol, amyl	15.0
Gutta-percha	3.0 to 5.0	Alcohol, ethyl	24.3 to 27.4
Ice	3.0	Alcohol, methyl	32.7
Marble	8.3	Bakelite	4.5 to 5.5
Mica	2.5 to 6.6	Benzine	1.9
Paper with turpentine	2.4	Benzol	2.2 to 2.4
Paper or jute impregnated	4.3	Micarta	4.1
Porcelain	5.7 to 6.8	Olive oil	3.0 to 3.2
Rubber	2.4	Paraffin	2.3
Rubber, vulcanized	2.5 to 3.5	Paraffin oil	1.9
Shellac	2.7 to 4.1	Petroleum	2.0
Silk	1.6	Turpentine	2.2
Sulphur	4.0	Water	81

Source: From C. E. Magnusson, *Alternating Currents,* McGraw-Hill, New York.

TABLE 16.7   Dielectric Strengths

Material	Strength, kv. per mm.	Material	Strength, kv. per mm.
Air at atmospheric pressure		Melted paraffin	7.5 to 20.0
(76 cm.)	3.0	Mica	25.0 to 220.0
Boiled linseed oil	8 to 19	Micanite	33 to 40
Dry wood	0.4 to 0.6	Paraffined paper	40.0 to 60.0
Fiber, vulcanized	8.0 to 18.0	Paraffin oil	16 to 21
Fish paper	10.0 to 15.0	Turpentine	3.5 to 16
Kraft paper	4.0 to 6.0	Varnished cambric	45.0 to 70.0
Marble	2.0 to 4.0	Varnished silk	45.0 to 70.0
Maple, oiled or paraffined	3.0 to 45.0	Vulcanized rubber	9.0 to 10.0

Source: From C. E. Magnusson, *Alternating Currents,* McGraw-Hill, New York.
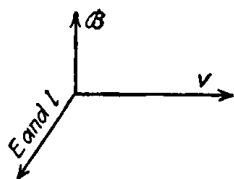
FIGURE 16.38
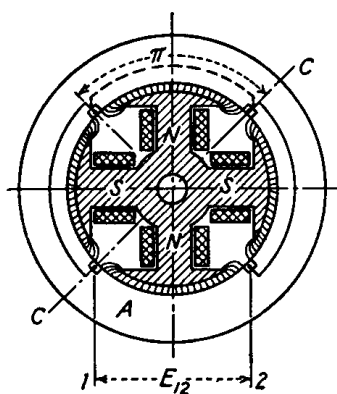
FIGURE 16.39

2. $\mathcal{B}$ shall be uniform along the entire length of the conductor, but not necessarily uniform along the direction of v.

3. Since the velocity v is relative, it is immaterial whether the conductor or the flux performs the motion. As a matter of fact, in most commercial alternators the conductor is stationary while the field revolves. In continuous-current machines, however, the flux is stationary and the conductor revolves.

54. Single-Phase Alternators.   Sine voltages are produced by machines called *alternators* which utilize Faraday's law of electromagnetic induction. In order to make this statement more real, the following brief description of such a machine is given.

Figure 16.39 shows a cross section of four magnetic poles marked N (north pole) and S (south pole). These poles are mounted on a shaft and are free to revolve inside an iron cylinder A, known as the *armature* of the machine. Four slots, 90° apart, are cut in the armature, and two insulated coils, each consisting of one turn of copper wire, are placed in these slots and are connected in series. Let the struc-
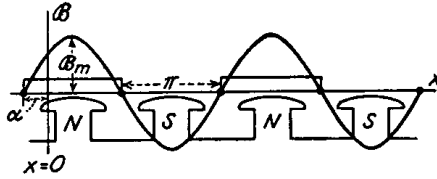
FIGURE 16.40

ture (Fig. 16.39) be cut radially along the line *CC* and the armature and poles be developed so that their relative positions are unaltered. This gives Fig. 16.40, to which the following facts relate:

1. The flux emanating from a north pole is considered positive, and the flux entering into a south pole is negative. This is indicated by drawing the flux-density wave (sine wave with amplitude $\mathcal{B}_m$) so that its positive values occur opposite a north pole and its negative values opposite a south pole.

2. The flux-density curve is assumed to be a sine function of space along the air gap. Therefore, its positive maximum occurs opposite the midpoint of a north pole, and its negative amplitude is opposite the middle of a south pole. The mathematical expression for this wave is

$$\mathcal{B} = \mathcal{B}_m \sin (x + \alpha) \tag{16.152}$$

3. From paragraph (2), it follows that a pole pitch (the distance from one point on a north pole to a similar point on an adjacent south pole) is equal to π radians measured along the abscissa of the flux wave.*

Now let v be the relative velocity of the pole structure and l the length per conductor. Then the emf between the terminals of each conductor is, according to Eq. (16.151),

$$e = \mathcal{B}_m v l \sin (x + \alpha) = E_m \sin (x + \alpha) \tag{16.153}$$

where $$E_m = \mathcal{B}_m v l \tag{16.154}$$

Let the field rotate through an electrical angle x in time t, and let its velocity be such that it takes T seconds to move over a distance of 2π electrical radians. Then

$$\frac{x}{2\pi} = \frac{t}{T}$$

$$\therefore x = \frac{2\pi t}{T} = 2\pi f t = wt \tag{16.155}$$

Substituting this value of x in Eq. (16.153),

---

*When distance along the periphery of the armature is expressed in radians, measured along the abscissa of the flux curve (Fig. 16.40), the term *electrical radians* is used. Distinction should clearly be drawn between an electrical radian and a mechanical radian. Thus there always exist π electrical radians (Figs. 16.39 and 16.40) between any two corresponding points of a north pole and an adjacent south pole, irrespective of the number of poles in a machine. On the other hand, considering that there are 2π mechanical radians in a circle, the number of mechanical radians between any two adjacent pole centers is $2\pi/p$, where p is the number of poles in a machine.

$$e = E_m \sin (wt + \alpha) = E_m \sin (2 \pi ft + \alpha)$$

$$= E_m \sin \left(\frac{2 \pi t}{T} + \alpha \right) \tag{16.156}$$

Thus there is a sine voltage of amplitude of E_m and frequency f.*

55. Polyphase Alternators.

Two-Phase Afternator. A *two-phase alternator* (Fig. 16.41a) differs from a single-phase machine in that it comprises two independent windings instead of just one. Thus let the armature in Fig. 16.41a have eight equidistant slots. Let these slots be divided into two groups by marking them alternately 1 and 2. Let insulated conductors be placed in these slots and be so connected that all conductors occupying slots 1 are connected in series and all those occupying slots 2 are also connected in series, so as to form two independent windings with terminals 11′ and 22′, respectively. If the flux distribution along the air gap is sinusoidal, then the voltage induced in any one conductor also will be sinusoidal. Furthermore, the sum of the emf's induced in the first group of conductors is equal to the terminal voltage $E_{1'1}$, and the sum of the emf's induced in the second group of conductors is equal to the terminal voltage $E_{2'2}$.

An inspection of Fig. 16.41a shows that at the instant when the first group of conductors is under the centers of the poles, the second group of conductors falls in the neutral spaces between two adjacent poles. In other words, when the flux cutting the first group of conductors is a maximum, the flux cutting the second group is zero, and vice versa. Since the induced voltage in a conductor is proportional to the density of the flux which cuts that conductor ($e = \Re vl$), when the voltage induced in the first group of conductors is a maximum, that induced in the second group is zero, and vice versa. There exists, therefore, a phase displacement in time of 90° (Fig. 16.42a) between the induced voltage $e_{1'1}$ and $e_{2'2}$. The first group of conductors (Fig. 16.41a) forms a winding which is called *phase 1,* and

(a)   (b)
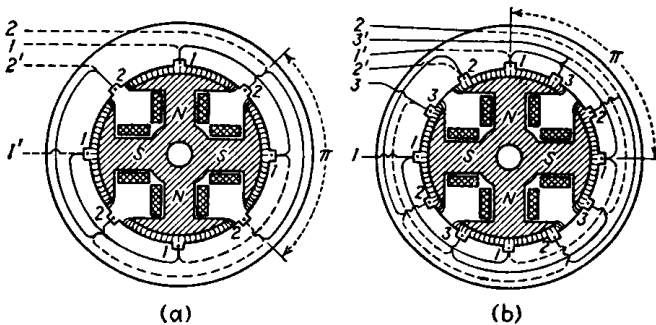
FIGURE 16.41

---

*It is desirable to develop a relation between the frequency of the emf generated and the number of revolutions that the pole structure makes. Consider an alternator having p poles and making S revolutions per second. Each conductor passes through $p/2$ cycles of flux per revolution because it takes two poles to make a complete cycle of flux (see Fig. 16.40). Thus, in 1 s it passes through $pS/2$ cycles. Consequently $f = pS/2$ cycles per second.
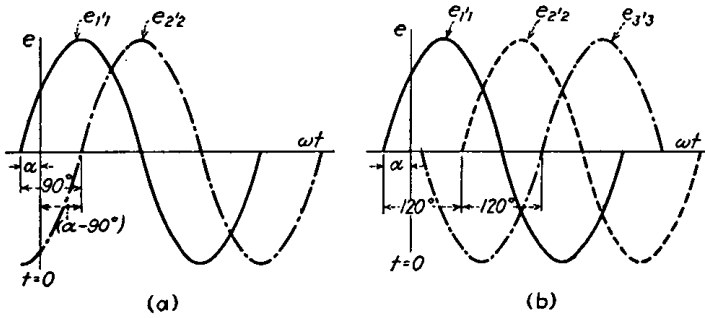
FIGURE 16.42

the second group forms another winding which is known as *phase 2*. The machine is a two-phase alternator.

Three-Phase Alternator.   The *three-phase system* has come to be almost universally used. Let each pole pitch (covering π electrical radians; Fig. 16.41*b*) be divided into three equal parts such that the points of division are 60 electrical degrees apart. Let slots be cut at these points of division, and let insulated conductors be placed in these slots. Figure 16.41*b* thus has 12 equidistant slots. Divide these slots into three groups by marking them consecutively 1, 3, 2 (not 1, 2, 3). Connect in series all the conductors occupying slots 1 and obtain a winding with the terminals 1′1. Do the same with the conductors occupying slots 2 and slots 3, respectively, and obtain two windings with terminals 2′2 and 3′3. The alternator is then said to be a *three-phase alternator* having three separate windings interconnected to give three terminals. The induced voltages $e_{1'1}$, $e_{1'2}$, and $e_{3'3}$, measured at these terminals, are displaced from each other (in time) by 120°, as shown in Fig. 16.42*b*.

Although alternators are generally constructed three-phase, the reader can see how an electric machine may be wound for any number *m* of phases by having an independent winding for each phase. The conductors constituting these separate windings are generally so displaced in space that there exist $2\pi/m$ electrical radians between any two corresponding conductors of two consecutive phases. Rotary converters furnish a practical illustration of electric machines built for six or more phases.

56. Direct-Current Generators

Construction.   In Sec. 53 the statement was made that the relative velocity of the conductor and flux is the cause of emf generation. In Sec. 54 a machine was described in which the emf is generated by the rotation of the field structure while the armature (conductor structure) is held stationary. Let it now be assumed that the armature structure will revolve while the field will remain stationary. Such a machine is shown in Fig. 16.43*a*; the voltage induced in the conductors is a line voltage.

Commutator.   Assume that in addition to the rotating armature there is a structure which contains pieces of copper C which are insulated from each other and to which the terminals of the armature coils are connected. Such a structure is called a *commutator*. Let stationary brushes be in contact with the commutator. As the
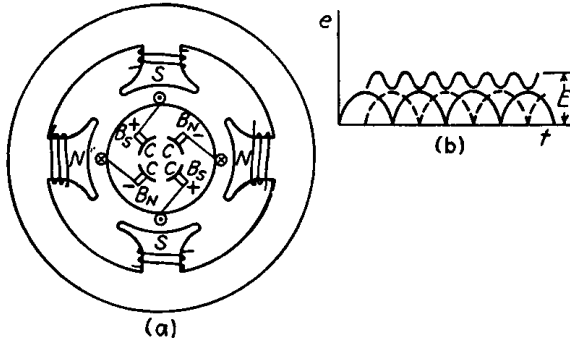
FIGURE 16.43   Direct-current generator.

armature revolves, the emf induced in any one coil side changes direction, but since each brush makes contact only with that coil side which is under a definite pole, the emf at the brush remains constant. The net result of such an arrangement is to produce, at any one brush, an emf which is unidirectional, as shown in Fig. 16.43*b*. Thus the actual emf produced by such a machine is (except for the ripples, which become less pronounced as more conductors are utilized) a dc voltage.

57. Types of DC Generators.   Alternators require a separate source of direct emf to excite their field windings. This is not true of dc generators, whose field excitation may be taken directly from the emf generated in the dc machine. The methods of connecting the field winding or windings to the terminal voltage of a machine vary, and this variation gives rise to differing operating characteristics. Accordingly, dc generators are divided into the following classes:

1. A *shunt generator* (Fig. 16.44*a*) is one that has its field winding connected across the armature. Its field strength, therefore, varies with the armature volts. Since the *IR* drop through the generator varies with the load, for a fixed position of the field rheostat the generator will have a drooping voltage curve (see Fig.
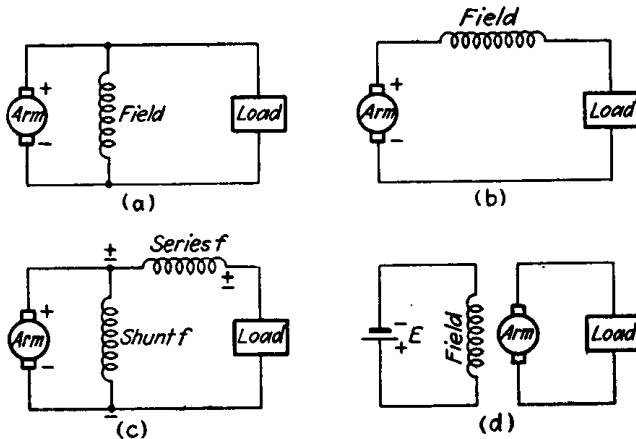
FIGURE 16.44

16.45). The amount of the drop in voltage from no load to full load depends on the magnetic densities, the amount of armature reaction, and the *IR* drop. For commutating-pole generators, the drop is about 15 to 20 percent of the line voltage. The shunt type of generator is not used to any great extent.

2. A *series-wound generator* (Fig. 16.44b) is one which has its field in series with the armature. Its excitation, therefore, varies with the load. The voltage rises rapidly with the load (see Fig. 16.45). Generators of this type are rarely built.

3. A *compound-wound generator* (Fig. 16.44c) is a combination of a shunt and series winding. The no-load voltage is obtained by the shunt winding. The series winding simply takes care of the *IR* drop and armature reaction. The voltage curve is called a *compound characteristic* (see Fig. 16.45). For slow-speed generators this curve (see Fig. 16.45) has a large hump; for moderate- and high-speed machines the curve is very flat. For some classes of service it is desired to *overcompound* the generator, i.e., to have a full-load voltage higher than the no-load voltage. This is to compensate for line drop under load.

4. A *differential generator* (Fig. 16.44c) has the series field in opposition to the shunt field, and thus there is more droop to the voltage curve (Fig. 16.45). This type is used in electric-shovel work, arc welding, and electric traction. By varying the percentage of differential and shunt fields, the load characteristic can be given any desired shape.

5. A *separately excited generator* (Fig. 16.44d) is usually a shunt-wound machine but with the field excited from an external source. This is necessary when the generator is operating over a variable voltage that starts from zero, especially if the generator has to carry a heavy load at low voltage. Under certain conditions, the series fields may be separately excited to obtain definite voltage curves.

The behavior of these various types of dc generators with load is shown in Fig. 16.45, where the terminal voltage of the machine is plotted as a function of the load current.

58. Three-Wire Generators.   Three-wire generators are built for use on the Edison three-wire system. They are usually 250-V and fundamentally standard dc generators with additions. There are two methods of connecting the third or neutral
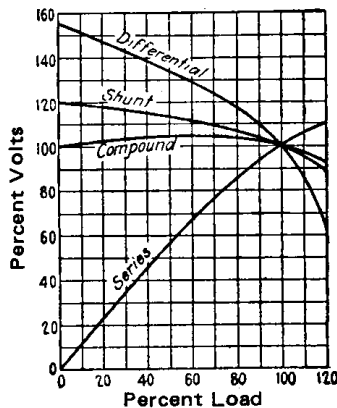
FIGURE 16.45   Characteristic curves of dc generators.

wire. One method is to build in a revolving compensator whose ends are connected to opposite sides of the dc armature winding, usually on the back end, and connect the center of its winding to a collector ring. The second and usual method is to connect two collector rings to the opposite sides of the armature winding and have an external compensator connected between them. The neutral wire is connected to the center of the compensator winding. In order to have the correct *IR* drop and field ampere-turns, alternate series and commutating fields are connected on the opposite sides of the armature (see Fig. 16.46).

SOURCES OF EMF: ELECTRIC BATTERIES

59. Definitions. *Electric battery cells* are devices for generating or storing electric energy by means of electrochemical action. Primary cells are those in which, by chemically transforming one or another of a group of materials, electric energy is generated. The *secondary cell* is one which, by means of a reversible electrochemical action, converts electric into chemical energy, which is then available to be returned again as electric energy. Such a cell is also called a *storage cell.* Two or more cells connected, usually in series, constitute a *battery.* The terminals of a cell are also called *poles.* The *positive terminal* or pole is that from which the current flows to the external circuit. The active chemical portions of the cell, connected to the terminals and immersed in the electrolyte, are the *electrodes.* The electrode connected to the negative terminal of the cell is called the *anode* because current flows from it into the electrolyte. The electrode into which the current flows in the cell is the *cathode.* In many primary cells, the anode is metallic zinc. *Polarization* is a secondary reaction occurring at the surface of one or both of the electrodes, which introduces a back emf and reduces that available at the terminals. A polarized cell, if allowed to stand on open circuit, will regain its normal emf. *Depolarizers* are means, usually chemical, for preventing polarization.

60. Wet Primary Cells. The development of storage cells and dry cells has greatly decreased the importance of primary cells using a fluid as an electrolyte. Before the development of the dynamo generator, such cells formed the only practicable means for the generation of electric energy. Tables 16.8 and 16.9 give the data for several of the most important wet cells.

61. Dry Cells. The only commercial dry cell makes use of the chemicals of the Leclanché wet cell, the ammonium chloride solution being held in an absorbent material which separates the electrodes. This cell is moist rather than dry and is
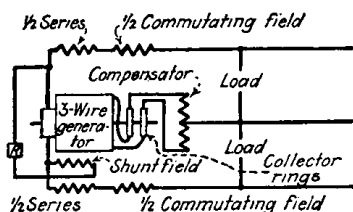
FIGURE 16.46 Three-wire generator connections.

TABLE 16.8 Wet Cells

Name	Electrolyte	Positive	Negative	Depolarizer
Chromic acid	H_2SO_4 or NaCl	Carbon	Zinc	CrO_2
Daniell	H_2SO_4 or $ZnSO_4$	Copper	Zinc	$CuSO_4$
Gravity	$ZnSO_4$ or H_2SO_4	Copper	Zinc	$CuSO_4$
Edison-Lalande	NaOH	CuO + Cu	Zinc	CuO
Leclanché	NH_4Cl	Carbon	Zinc	MnO_2

TABLE 16.9 Wet Cells

Name	E.m.f., volts	Resistance, ohms	End result
Chromic acid	2.00	5 to 4	$Cr_2(SO_4)_3 + 3ZnSO_4$
Daniell	1.07 to 1.14	0.3 to 30	$Cu + ZnSO_4$
Gravity	1.00	0.1 to 6	$Cu + ZnSO_4$
Edison-Lalande	0.75	0.02 to 0.1	$Cu + Na_2ZnO_2$
Leclanché	1.50	1 to 5	$Mn_2O_3 + 2NH_3 + ZnCl_2$

useless when completely dried out. In the usual form, a cylindrical container of zinc is connected to the negative terminal, and a central rod of carbon to the positive. Surrounding the carbon rod is a mixture of carbon and manganese oxide used as a depolarizer. Between this and the inside of the zinc cylinder is the absorbent material, saturated with ammonium chloride and the zinc chloride which is formed during discharge. The open-circuit voltage, 1.5 V, is reduced to about 1 V on closed circuit, principally by polarization. Even a good dry cell, when unused, becomes worthless, owing to drying out and local action, after about a year. The most common size of large cell has a 2½-in (6.4-cm) diameter and a 6-in (15-cm) length. On short-circuit tests such a cell should give about 20 A or more. The cylindrical unit cell most used in flashlights is 1⁵⁄₁₆ × 2¼ in (3.3 × 5.7 cm); two or three such cells are generally used in series. The data in Table 16.10 show typical dry-cell capacities. For an end voltage of 1 V, the values were about two-thirds of the above. For 0.6-V end voltage, they were about 50 percent greater. If the service is intermittent, the capacity may be increased as much as two or three times, but if the test extends over many days, the total capacity is decreased by local action.

TABLE 16.10 Capacity of Dry Cells [6 in (15 cm)]

(End voltage, 0.8 volt)		
Amp	Hr	Amp-hr
½	16	8
¼	51	12
⅛	143	18
¹⁄₁₆	414	26
¹⁄₃₂	1078	34

62. Storage Cells.   These are cells which are reversible as to their electrochemical action. That is, charging a discharged cell with electric current will bring it back to its original chemical conditions and thus make it available for another discharge. In order that the plates may keep their original form, it is necessary that the active material should not be soluble in the electrolyte. Such cells are also called *secondary cells* or *accumulators,* as distinguished from primary cells. Two types of storage cells are in general use: the lead cell and the alkaline cell.

63. The Lead Cell.   The electrodes of this cell consist of lead grids or frames which carry the active material. When the battery is completely charged, the active material on the positive plate is principally lead peroxide (PbO_2), and on the negative plate spongy metallic lead. During discharge both materials are converted into lead sulfate ($PbSO_4$). The positive plate is reddish brown and the negative plate is grayish in color. There is always one more negative than positive plate, because the negative is more robust and withstands better the effect of the reactions on one side only. The electrolyte is chemically pure sulfuric acid and water. For stationary batteries, the specific gravity varies from 1.150 to 1.160 on discharge and rises to 1.200 or 1.215 when fully charged. In portable and vehicle batteries, the specific gravity may be 1.300 when charged. The smaller cells are usually carried in hard rubber or glass jars, sealed at the top except for small ventilating holes. Large cells have lead-lined wooden tanks. During the life of the cell, the active material slowly disintegrates and falls from the plates and thus eventually terminates the usefulness of the cell. This sediment collects in the lower part of the container, so that it is necessary to support the plate a considerable distance from the bottom to prevent the sediment from coming in contact with, and short-circuiting, the plates. Thin separators of wood, plastic, or rubber are generally used between the plates.

There are, in general, two methods of applying the active material to the plate. For the *pasted* plate, the lead oxide is made into a paste and forced into the interstices of a suitable grid of pure lead which forms the plate. Such plates are assembled and placed in sulfuric acid. A forming current is then used to convert the positive plate into lead peroxide and the negative plate into spongy lead. In the *formed* or Planté plate, the active material is formed from pure lead, sometimes from the supporting frame itself, by chemical means. The Planté type is usually heavier, stronger, and more durable. The pasted type is generally used for portable and vehicle batteries.

64. Operation of Lead Storage Cell.   During the charge, the emf rises from about 2.2 to 2.6 V, though this latter figure may range in extreme cases from 2.4 to 2.8. The discharge starts at about 2 V and is usually carried to 1.8 V or, with heavy current, as low as 1.7 V. The state of charge or discharge can be determined by the voltage only when normal current has been passing through the battery for a period of several minutes. Nothing concerning the condition of the cell is shown by the open-circuit voltage. The state of charge can be determined accurately by a hydrometer reading, and the completion of charging is indicated by the violent emission of gas, due to decomposition of the water.

The capacity rating of a cell is usually based on a discharge continued during a certain number of hours. For stationary batteries, this is usually 8 h, but for vehicle batteries it may be 3 or even as low as 2 h. A battery is not injured by a wide variation of discharge rates. When discharged (or charged) at a high rate, its ampere-hour capacity for the discharge in question is greatly decreased. For a 5-h

discharge, the capacity of a cell is about ⅞ of the 8-h-rate capacity, and for a 1-h discharge, this becomes ½.

While heavy discharge rates can generally be handled without injury to the battery, frequent discharge below 1.8 V, and especially complete discharge, causes the formation of an insoluble lead sulfate, which seriously damages the battery. This sulfate may, to some extent, be reconverted into the normal active material by long-continued slow charging. Since, when the battery is discharged, the sulfuric acid is weakened, there is danger that the battery will freeze under such conditions. For a specific gravity of 1.15, normal when a battery is discharged, freezing occurs at about 5°F (-15°C); for 1.2, at -16°F (-26.7°C); and for 1.22, at -30°F (-34.4°C). The efficiency of a lead cell at a given discharge rate is the watthours received on discharge, divided by the watthours used on the previous charge. The charge must be at a specified rate, and the starting and stopping states of the battery must be the same.

65. Alkaline Storage Cells.   The electrolyte of the Edison cell is a 21 percent solution of potassium hydrate in water; it also contains a little lithium hydrate. The active electrode materials for the charged cell are nickel peroxide for the positive plate and finely divided iron for the negative plate. Nickel flakes are added to the nickel peroxide to increase its conductivity, and small amounts of other chemicals are present to promote the normal reactions. When discharged, the active material of both plates becomes a mixture of iron and nickel hydrates and oxides. The active materials on both plates are held in their perforated containers of nickel steel, which form parts of the plates. Alternate positive and negative plates are mounted on insulated steel rods, and the elements so combined are placed in a nickel-steel container. These containers must, therefore, be insulated from each other in mounting. The chemical reactions at the two plates balance each other, so that there is no change in the electrolyte except some loss of water by decomposition and evaporation. In charging, the emf, starting at about 1.5 V, rises rapidly to 1.7 V and then gradually to 1.8 V when fully charged. The discharge, starting at 1.4 V, falls rapidly to 1.3 V, then gradually to 1.1 V on the completion of the discharge. The preceding voltages are considerably decreased at high discharge rates. With four times normal discharge rates, the operating range. of voltage will be from 1.0 to 0.8 V. Seven hours is the normal charging rate for most Edison cells. High discharge rates can be used without injury. Over-charging or overdischarging does not seriously injure the cells, which are much more robust than the lead cells. The specified capacity of a cell is normally based on a 5-h discharge rate. The capacity of a cell increases during its early use, sometimes as much as 30 percent above the rating. A 40°F (22°C) rise in temperature decreases the voltage about 0.1 V, and the capacity of the cell is greatly decreased by low temperature.

At normal discharge rate, the watthour efficiency is about 72 percent, and at a 1-h discharge rate, about *58* percent. The internal resistance of the alkaline cells is considerably higher than that of the lead cells, so that there is an immediate drop of about 8 percent with normal discharge currents.

The Hubbell cell differs from the Edison cell principally in the use of cadmium instead of iron for the negative plate. It is used for miners' lamps.

66. Lithium Storage Cells.   Lithium solid-state batteries have been used for many years as a long-life power backup for computer system clocks and nonvolatile memories. These cells have extremely long shelf lives and their typical

output voltage of more than 2.5 volts is useful with CMOS "non-volatile" memories.

The disadvantages of lithium storage cells are their inability to source large amounts of current, their inability to be recharged and their toxicity. Lithium storage cells have been known to explode if attempts are made to recharge them.

67. NiCad and NiMH rechargeable cells. Nickel-cadmium ("NiCad") rechargeable cells have been available for many years as an alternative to dry or alkaline cells for many applications. Early cells were known as "Nife" cells and were used as long-term power storage. The cell's typical output voltage of 1.2 volts is less than other options. Concerns with NiCad cells are their toxicity as well as their propensity to form internal dendrites that will decrease their operating life over time.

Recently, nickel-metal hydride ("NiMH") cells have become available. These cells offer the same output characteristics of NiCads but do not have the same downsides. NiMH cells require different charging equipment than NiCads.

Different Battery Type Operation
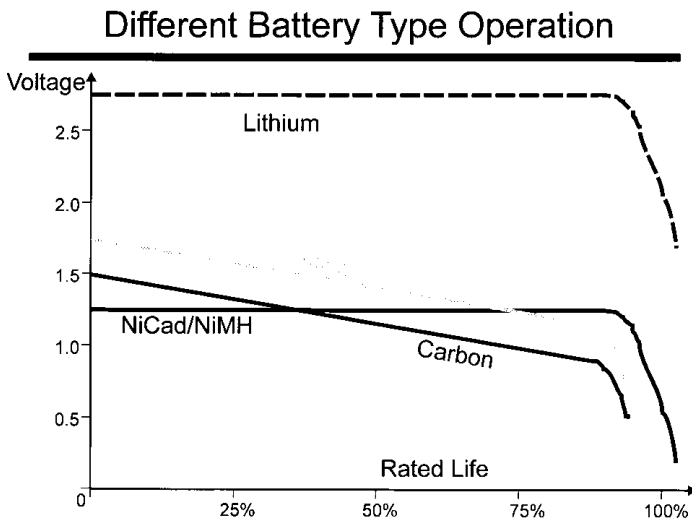
68. Battery-Regulating Equipment. Owing to the large change in voltage between charged and discharged conditions, voltage-regulating equipment is needed. An *end-cell switch* is similar to the faceplate of a rheostat, but the contacts are connected in between the cells, so that a successively increasing number can be used, as the voltage per cell drops during discharge. This switch should be designed so that in passing from cell to cell, while the circuit is not opened, no cell is short-circuited through the contacts. Instead of cutting in regular cells, this switch may cut out *counter-emf cells.* These are cells of low capacity, the voltage of which is always used in opposition to the voltage of the main battery, so that the end cells are always charged during use. A booster generator may be used to obtain increased voltage required by the battery during charge.

TRANSFORMERS*

69. General Considerations. Figure 16.47 shows an iron core on which two windings are placed. The *primaly winding* has n_1 turns; the *secondary winding* has n_2 turns. If an alternating voltage e is impressed on the primary, a current i_1 will flow. This current causes flux ϕ in the core of the transformer. Moreover, since the voltage and hence the current in the primary are alternating, the flux ϕ will alternate. But the flux threads both the primary and secondary windings. Hence in both primary and secondary windings a voltage e_2 is induced of such magnitude that

$$e_1 = n_1 \frac{d\phi}{dt} \tag{16.157a}$$

$$e_2 = n_2 \frac{d\phi}{dt} \tag{16.157b}$$

Dividing Eq. (16.157a) by Eq. (16.157b),

$$\frac{e_1}{e_2} = \frac{n_1}{n_2} \tag{16.158}$$

Virtually, e_i is equal to the applied voltage e. Moreover, Eq. (16.158) states that any desired ratio of primary to secondary voltage can be obtained merely by varying the ratio of the number of turns n_1 and n_2 in the primary and secondary. A *transfonner* is therefore a device which is used to transform the voltage from high to low, or vice versa.

According to the law of conservation of power,

$$E_1 I_1 = E_2 I_2 \tag{16.159a}$$

$$\therefore \frac{I_2}{I_1} = \frac{E_1}{E_2} = \frac{n_1}{n_2} \tag{16.159b}$$

Thus while the ratio of primary to secondary voltage varies directly as the number of turns, the ratio of primary to secondary current varies inversely as the number of turns.
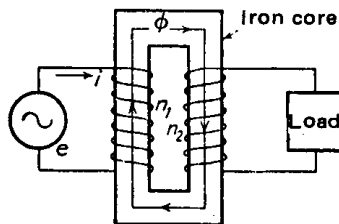
FIGURE 16.47   Transformer.

Generally speaking, in a powerhouse where electricity is generated, transformers are used to *step up* the voltage to a higher value for transmission purposes. At the other end of the transmission line, in substations and industrial plants, step-down transformers are used to reduce the voltage. Therefore, either the high-voltage winding or the low-voltage winding may be the primary or the secondary depending on the work the transformer is doing. *The power is supplied to the primary and given out by the secondary.*

70. Classes of Transformers. All transformers for power or distribution service are divided into two main classes: power transformers and distribution transformers. *Power transformers* are above 500 kVA and are used to step up the generated voltage to the transmission voltage; they are also used in substations to step down the transmission voltage to the distributing voltage or to a standard-service voltage. *Distribution transformers* are 500 kVA and under, being used to step down from a transmission voltage to a distribution voltage or to a standard-service voltage or from a distribution voltage to a service voltage.

A large percentage of all transformers are built *single-phase;* a small percentage of both power and distribution transformers are built *three-phase.* It is usually slightly cheaper to build one three-phase transformer than three single-phase transformers of equal total power. The reason for using single-phase transformers is flexibility. Where a system is supplied from a three-phase transformer, a short circuit or open circuit in this unit usually cuts off the whole system. In case of trouble with one single-phase transformer in a three-phase bank, this transformer can be cut out and the other two operated in open-delta, or a spare transformer can be substituted. With large sizes, the question of size and weight also must be considered on account of difficulties that may be encountered in shipping and handling.

Three single-phase transformers may be connected in either delta or Y, and these connections may be changed as desired. A three-phase transformer has the connections made permanently at the factory. It is the general practice in the United States to use single-phase transformers and in Europe to use three-phase transformers.

71. Cooling. Transformers may be *air-cooled,* by the natural circulation of the air around the core and coils, or they may be cooled by an *air blast,* the air being forced through the transformer. They may be oil-immersed, the cooling being obtained by natural circulation of oil within the tank and circulation of air around the tank, or *oil-immersed, forced oil-cooled,* where the oil is circulated by a pump through external coolers. They are *water-cooled, oil-immersed,* where the core and coils are immersed in oil but cooling water is forced through pipes within the transformer tank to carry off the heat. For the larger sizes, various devices—such as corrugated tanks, tanks with pipes welded to the outside, and radiators—are used to increase the radiating surface and improve the cooling process. Many large transformers are now being equipped with blowers and the necessary ducts to direct jets of air against the radiating surfaces, so that the cooling efficiency is greatly increased. Because the loss varies as the volume, and hence as the cube of the dimension, while the heat-radiating surface varies only as the square, the difficulty of cooling increases with the size of the transformer.

72. Mechanical Construction. There are several types of construction, each of which gives maximum service, best operating characteristics, and highest efficiency under certain conditions. A change in kilovoltampere capacity or operating voltage

emphasizes different heat-radiating requirements, mechanical forces, or voltage stresses which may be best handled in the design by different windings, insulation, or tank construction. Transformers which are operating 24 h per day, at low-load factors, should have small exciting and iron losses with comparatively large load losses, while large power transformers operating at a high constant load may have low load losses and high exciting and iron losses, provided the total loss is a minimum. In either case, a high all-day efficiency is obtained.

Mechanically, transformers are of the *core type* or *shell type,* as shown in Fig. 16.48. As the stresses set up by heavy currents tend to make the coil cirular, most transformer coils are now built in that shape. A large percentage of the small single-phase transformers are built as a modified core type (Fig. 16.48*b*), but may have three or four outside legs (Fig. 16.49). Three-phase transformers are usually of the core type shown in Fig. 16.48*a*.

The coils may be wound in concentric cylinders (Fig. 16.50). The low-voltage winding may be all next to the core (Fig. 16.50*a*), or both windings may be split and wound (Fig. 16.50*b*). The coils may be wound in pancake form (Fig. 16.51*a*), or may be a combination (Fig. 16.51*b*). The windings and core are so arranged that the oil or air may circulate in ducts through them and thus maintain uniform cooling.

The *tanks* are of welded-steel construction. Nearly all tanks are provided with lifting lugs and, with the exception of pole-type units, have jack bosses. Since the great majority of transformers are used outdoors, the covers and bushings must be built for this service.
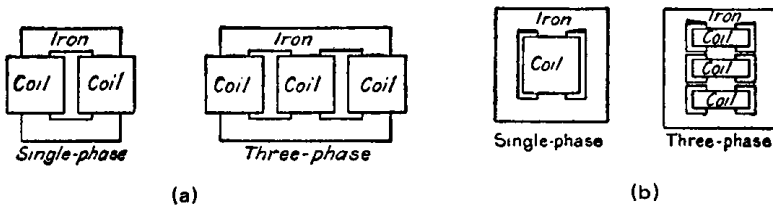
FIGURE 16.48   (*a*) Core-type transformers. (*b*) Shell-type transformers.
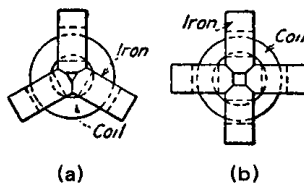
FIGURE 16.49
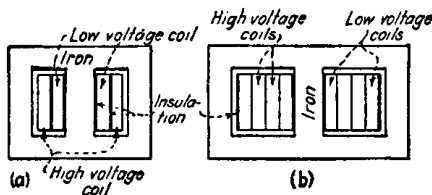
FIGURE 16.50
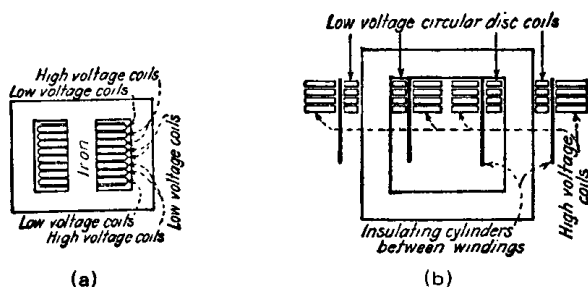
FIGURE 16.51

The *insulation* must be of a very high quality because of the high voltages involved and the compact arrangement of the transformers necessary for proper operating characteristics. High-class materials, such as mica, special thin paper built up to the required thickness and cemented together with synthetic resin under heat and pressure, and hard fiber, are used. All completed coils are vacuum-dried and treated with a moisture- and oil-resisting varnish. Oil is used as an insulating as well as a cooling medium. Oil introduces two difficulties: sludging or carbonization of the oil and the danger of explosion of oil vapor and air. Sludging, caused by oxygen coming into contact with hot oil, reduces the insulating and cooling value of the oil. Both these difficulties can largely be eliminated by providing some means of preventing the air from coming into contact with the hot oil. There are two methods in commercial use. Since the volume of oil changes with the temperature, it is not possible to seal the transformer tank. One method uses tanks with the space above the oil level filled with an inert gas to prevent air from coming into contact with the oil. Another method uses the transformers with oiltight covers and then adds a second, smaller tank slightly above to keep a constant head of oil on the transformer and to allow for expansion and contraction. As means are taken to ensure a very slow passage of oil in and out of this extra tank, its oil is cool, and contact with the air does not cause it to sludge. These methods also prevent explosions due to a mixture of oil vapor and air. Noncombustible insulating fluids are now being introduced as substitutes for oils.

73. Subtractive and Additive Polarity. Consider a single-phase transformer. Connect one high-voltage terminal to the adjacent low-voltage terminal and apply voltage across the two high-voltage terminals. Then if the voltage across the un-connected high-voltage and low-voltage terminals is less than the applied voltage, the polarity is *subtractive,* while if it is greater than the applied voltage, it is *additive* (see National Electrical Manufacturers' Association *Standards*). Additive polarity is standard for all single-phase transformers 200 kVA and smaller, whose high-voltage ratings are 2500 V and below. Subtractive polarity is standard for all others (see NEMA *Standards*).

74. Delta-Delta Connection. For three-phase circuits, transformers may have their primary or secondary windings connected in either delta or Y. Figure 16.52 shows a delta-delta connection.

Advantages. Any three similar single-phase transformers can be so connected. The bank can operate open delta if one unit is disabled (see Sec. 75). For low voltages and high currents, the delta connection gives a more economical design
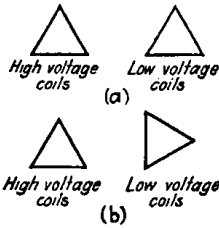
FIGURE 16.52 Delta-delta connection.
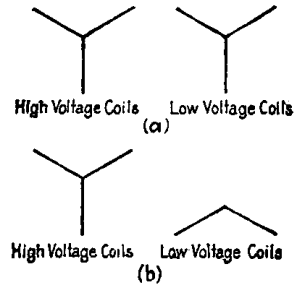
FIGURE 16.53 Y-Y connection.

than the Y connection. This connection is free from all third-harmonic troubles. The primary and secondary delta carry the third-harmonic magnetizing current which does not appear on the lines.

Disadvantages. The neutral cannot be derived. Differences in the voltage ratios cause a circulating current in both primary and secondary windings, limited only by their impedances. Differences in impedances cause unequal load division among the units. For very high voltages, the delta connection costs more than the Y connection. The connection of Fig. 16.52*a* cannot be connected in multiple with that of Fig. 16.52*b*.

75. Y-Y Connection. This is shown in Fig. 16.53.

Advantages. The neutral can be brought out for grounding. Differences in impedances and ratios of the units do not cause any circulating currents or appreciably unequal load division. For relatively high voltages and small currents, the Y connection is generally more economical than the delta. A short circuit in or on one unit does not cause a power short circuit, but does cause a very large magnetizing current because of the overexcitation of the remaining units at 1.73 times rated voltage.

Disadvantages. There is a third-harmonic voltage from line to neutral amounting to as much as 50 percent in some types of single-phase transformers. The neutral is unstable unless grounded. The units cannot be loaded, single-phase, line to neutral, unless the neutral of the primary is connected to that of the generator. If the neutral is grounded, it may cause telephone interference. The Y-Y bank cannot operate with two units when one is disabled. A short circuit in or on one unit raises the voltage on the other units to 1.73 times normal. Figure 16.53*a* cannot be connected in multiple with Fig. 16.53*b*.

76. Delta-Y Connection. This is generally considered to be the most satisfactory three-phase connection.

Advantages. The neutral can be brought out either for grounding or for loading. The neutral is stable, being locked by the delta. The connection is practically free from third-harmonic voltages, as third-harmonic magnetizing current circulates

through the delta winding. Differences of magnetizing current, voltage ratio, and impedance in the different units are adjusted by a small magnetizing current circulating in the delta. A short circuit in one leg of the Y does not affect the voltages of the secondary lines. A single-phase short circuit on the secondary lines causes a smaller short-circuit stress on a delta-Y-setup bank than on a delta-delta bank. Figure 16.54*a* can be multipled with Fig. 16.54*b* by properly selecting the leads.

Disadvantages. The delta-Y bank cannot operate temporarily with two units when one is disabled. A short circuit on or in one unit is extended to all three units. If the delta is on the primary side and is accidentally opened, the unexcited leg on the Y side may resonate with the line capacitance and cause damage.

77. Open-Delta or V-V Connection (see Fig. 16.55).
This connection requires only two units and is useful in an emergency; it is used in small control systems. Since the internal power factor is 0.866 (assuming a unity power-factor load), it can deliver only 0.866 of its rated kilovoltampere capacity or 58 percent of the capacity of three units. Load voltages become unbalanced under load, even with balanced three-phase load, the amount of unbalancing depending on the impedance of the units and the power factor of the load.

78. Three-Phase, Two-Phase Connections.
The *Scott connection* shown in Fig. 16.56 has the following *advantages:* It requires only two single-phase units or one two-phase unit. It can be used with either three-wire or four-wire two-phase service. Both two- and three-phase voltages can be obtained on the primary, with only four wires if the 86.6 percent tap of one winding is connected to the center of the other.

Disadvantages. If the two transformers are duplicates, there must be a 50 percent tap and an 86.6 percent tap on each unit. The three-phase side carries 15 percent more current than that corresponding to the two-phase side. It therefore requires 15 percent more copper, or the two-phase side must be operated at 86.6 percent of its
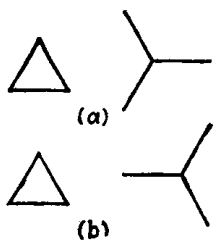
FIGURE 16.54   Delta-Y connection.
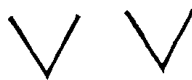
FIGURE 16.55   V-V connection.
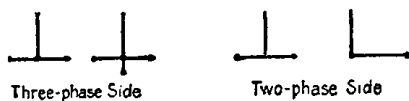
FIGURE 16.56   Three-phase two-phase transformer connection.

capacity. The decreased use of two-phase systems limits the application of the Scott connection. It is seldom, if ever, used for transformation of voltage without change in number of phases.

79. Autotransformers. An *autotransformer* differs from a regular transformer in that it has only one winding instead of a primary and secondary winding. The change in voltage is obtained by connecting the winding across the impressed voltage and tapping off another circuit with more or fewer turns. The connection for single-phase is shown in Fig. 16.57. Autotransformers may be wound single-phase but usually are three-phase or two-phase.

Advantages. For a given output, autotransformers are cheaper than transformers. This economy is greater the nearer the ratio comes to unity. Autotransformers also have better efficiency and regulation than transformers. This advantage also increases as the ratio approaches unity.

Disadvantages. There are some serious disadvantages in the use of autotransformers. The high- and low-voltage windings being continuous, the low-voltage circuit and connected apparatus are subjected to high voltage to ground and may be subjected to abnormal voltages due to disturbances and grounds on the high-voltage circuit. This is particularly true when there is a large difference between voltages. Short-circuit currents are larger with autotransformers, and this is worse as the voltage ratio approaches unity or the high and low voltages approach each other. Except for starting compensators, autotransformers are seldom used where the ratio exceeds 2:1.

Autotransformers are recommended for isolated systems and for grounded systems, provided the neutral of the autotransformer is also grounded. But they are used most extensively for supplying a low voltage for starting synchronous or induction motors.

The *Y connection* shown in Fig. 16.58 is the most economical and therefore the most used autotransformer connection. The ratio of rating to output is $(E_1 - E_2)/E_1$, where E_1 is the low voltage. The neutral may be grounded if the generator neutral is also grounded. The disadvantages are the same as given for Y-Y transformers (see Sec. 73).

The *V* or *open-delta connection* (Fig. 16.59*a*) is free from third-voltage harmonics, and the ratio of rating to output is 15 percent more than for the Y connection.

The *delta connection* (Fig. 16.59*b*) has characteristics similar to those for transformers connected delta-delta (Sec. 72). The rating to output is $(E_1^2 - E_2^2)/1.73 E_1 E_2$.

The *extended-delta connection* (Fig. 16.59*c*) gives a ratio of rating to output of $\sqrt{1 - 0.25(E_2/E_1)^2} - 0.8666 E_2/E_1$ or $1.73 E_x/E_1$, where E_x is the voltage of the extended portion. When the Y connection is undesirable, owing to the third har-
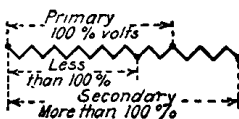
FIGURE 16.57

FIGURE 16.58 Y connection.
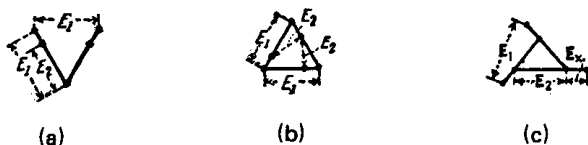
FIGURE 16.59 (*a*) Open-delta connection. (*b*) Delta connection. (*c*) Extended-delta connection.

monic, the V or extended-delta connection is more economical than the delta connection. When the low voltage is less than 92 percent of the high voltage, the extended-delta connection requires a smaller rating than either the straight delta or V connection. When this ratio is over 92 percent, the V connection requires the smallest rating of the three.

MOTORS

80. Introduction. Direct-current motors and ac synchronous motors are identical in construction with dc generators and ac generators, respectively. Motors and generators differ only in their functions. A *generator* converts mechanical energy into electric energy; a *motor* converts electric into mechanical energy. Like generators, motors are divided into two fundamental classes: ac motors and dc motors.

ALTERNATING-CURRENT MOTORS

81. Types. Alternating-current motors may be divided into *induction* and *synchronous* motors. A synchronous motor runs always at synchronous speed independently of the load; an induction motor runs at a speed below the synchronous speed (subsynchronous speed). The synchronous speed is determined by the following equation:

$$\text{rpm} = \frac{120f}{p}$$

where rpm is the number of revolutions per minute of the motor, f is the frequency of the supply source of emf, and p is the number of poles of the machine.

Both synchronous and induction motors may be single-phase or polyphase. Polyphase motors are, however, more used commercially because of their better operating characteristics. Single-phase motors are used where polyphase power is not available, as in homes. In all cases in which reference to cost or size is given, the squirrel-cage-type three-phase induction motor is used as a reference standard. Polyphase motors are usually three-phase. The different kinds of polyphase motors are described in the following paragraphs.

82. Squirrel-Cage Induction Motors. A *squirrel-cage induction motor* consists of a distributed wound stator and a short-circuited or "squirrel-cage" rotor. The power is applied to the stator winding only; this sets up a magnetic field which

revolves at synchronous speed. The rotor acts as the secondary of a transformer. With the rotor at standstill, the transformer action generates a low voltage in its rotor at the same frequency that is impressed on the primary or stator windings. As the rotor is short-circuited, this voltage forces a high current through it. The value of this current depends on the resistance and reactance of the rotor windings. The magnetic field set up by this current tends to travel with the primary revolving field and thus develops torque. This torque varies with the current. As the rotor increases in speed, its voltage, current, and frequency decrease, so that theoretically at synchronous speed they are zero and no torque is developed. The rotor must "slip" behind the revolving field enough to develop sufficient secondary voltage to force the required current through the rotor. For this reason, the slip varies directly with the load. This is the simplest and cheapest type of motor. It is used for any application requiring approximately constant-speed operation.

1. The *standard-type squirrel-cage motor* has low starting, high pullout torque, and draws a high starting current; these vary with the capacity and speed. Since the starting torque varies with the rotor resistance and reactance, there are several variations from the standard; these are known as *high-resistance squirrel-cage motors, high-reactance low-torque squirrel-cage motors,* and *high-reactance high-torque* or *double-winding squirrel-cage motors.*

2. The *high-resistance squirrel-cage motor* may have a very high starting torque and a high starting current. Since this motor has a large slip, it is especially suitable for applications where a flywheel is used, such as punch presses, and for frequent-starting duty. It is also used for elevator service. Its efficiency and power factor are lower than those for a standard squirrel-cage motor, and therefore, it is not often used for continuous operation.

3. The *high-reactance low-torque motor* has a starting torque about the same as that of the standard-type motor, but the starting current is lower. The efficiency and power factor are usually slightly lower than for a similar standard-type motor, but in some ratings they are being used instead of the standard type.

4. The *high-reactance high-torque motor* may be designed for high starting torque with a moderate starting current. This motor has still lower efficiency and power factor. Average speed-torque and current-speed, efficiency, and power-factor curves are given in Figs. 16.60 to 16.62.

5. *Starting.* Any of the squirrel-cage-type motors can be designed for across-the-line starting. It is standard practice to start motors at 5 hp (3.7 kW) and below in this way. In most applications, the high-reactance and high-resistance motors up
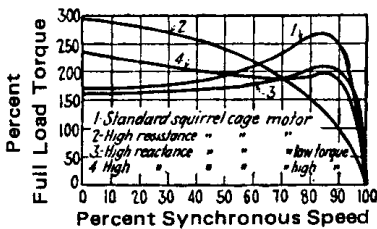
FIGURE 16.60 Torque characteristics of squirrel-cage motors.
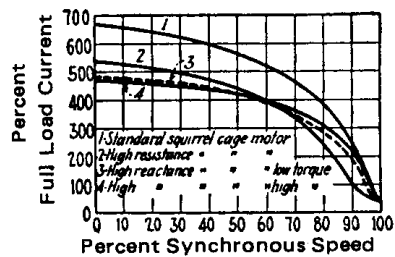
FIGURE 16.61 Current characteristics of squirrel-cage motors.
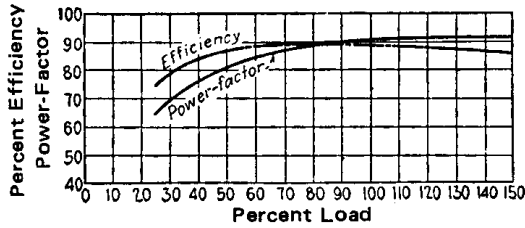
FIGURE 16.62

to 200 hp (150 kW) are thrown directly across the line. The limiting feature is the power system. For frequent starting or for applications where slow starting is desired, a reduced-voltage starter should be used.

83. Slip-Ring Induction Motor. This type has a stator winding similar to that of the squirrel-cage motor, but the rotor is not short-circuited; it is polar-wound, with the ends brought out to collector rings. There are usually three rings, although some special motors have six rings. The collector rings are connected, through brushes, to an external resistance. The speed at which any torque is developed may be controlled by varying the rotor or secondary resistance. That is, the higher the total rotor resistance, the greater must be the slip, below synchronism, to develop a given torque. A series of torque-speed curves for different values of rotor resistance is shown in Fig. 16.63, and the corresponding current curves are given in Fig. 16.64. The dotted curve *M* at the right in Fig. 16.64 gives the current of the motor with the brushes short-circuited. Curve 8 in Fig. 16.63 shows the actual slip with the resistance all cut out, but allows for the resistance of brushes, leads, etc. From these curves it may be seen that if a motor is operating at any speed and the power is reversed, it will develop torque in the opposite direction. This is called *countertorque* or *plugging* and is often used to stop the motor quickly. When a motor is plugged at normal speed, the secondary voltage is double the standstill voltage, and for values of torque up to 150 or 175 percent, it is nearly double what it would be at standstill with any given resistor value. In plugging, the rotor resistance also should be increased to correspond approximately to curve 3 in Fig. 16.63.

Regenerative braking can be obtained at speeds above synchronism when the load changes and drives the motor as a generator. With no secondary resistance,
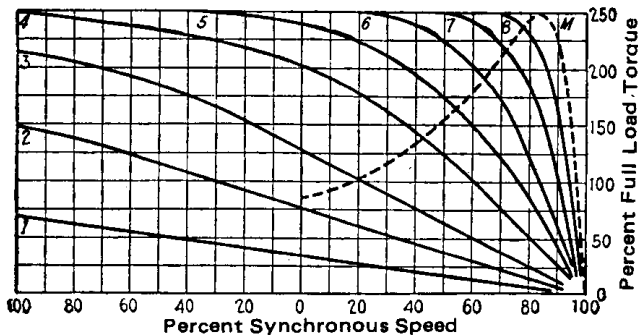
FIGURE 16.63 Torque-speed curves of slip-ring motors.
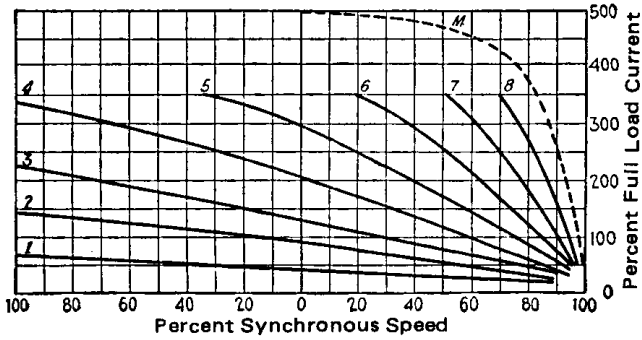
FIGURE 16.64 Current-speed curves of slip-ring motors.

the motor, operating as a generator, will develop normal torque with normal slip above synchronism. Cutting in resistance increases the speed at which normal torque will be developed.

Dynamic braking is obtained by disconnecting the ac source of power and exciting the stator from a dc source, usually two phases in series, the third being open. For any given torque requirement, the speed can be changed by changing the resistance in the secondary circuit. The more resistance, the higher the speed. Below about 15 or 20 percent of full-load speed, the motor becomes unstable and may speed up and run away. The maximum torque which a given motor can develop when acting as a dynamic brake is about 75 percent of its breakdown torque as a motor.

Slip-ring motors are used for any applications requiring high starting torque, low starting current, speed control, or reversing. With secondary resistance, full-load torque may be obtained with approximately full-load current at any speed. This motor is used on such heavy duty applications as mine hoists, steel-mill main-roll drives, etc. Its cost is about 15 to 50 percent more than that of a squirrel-cage motor, depending on the size. Small motors have the greater cost difference. For slip-ring motors it is always convenient to start with some resistance in the rotor circuit. This resistance may be cut out in one or more steps during accclcration.

84. Concatenated Motors. By connecting two similar slip-ring motors in series, half speed is obtained at normal current and normal full-load torque. The rotors are mounted on the same shaft and thus are rigidly connected together. The stator of one motor is connected to the line, and its rotor is connected to the stator of the second motor (see Fig. 16.65). The rotor of the second motor may be short-circuited or may have a resistance in series for starting or speed control. At half speed the secondary of the first motor develops one-half normal frequency. Thus the second motor has half normal frequency applied, and its primary acts as a resistance for the secondary of the first motor. For this connection the open-circuit slip-ring voltage must equal the primary voltage. By suitable switching the same two motors may be connected in multiple to give normal speed. This combination gives a 2:1 speed ratio at constant torque and therefore gives similar results to a pole-changing motor. Owing to the heating at normal-load current, at one-half speed, the motors must be larger than would otherwise be required. This same type of connection can be used with two dissimilar motors to get a lesser speed change. For instance, if a large motor of 24 poles were used and a smaller one of 4 poles, the speed of the
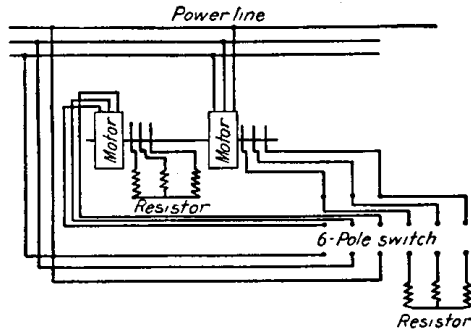
FIGURE 16.65 Concatenated motors.

main motor would be reduced in the ratio of (24 + 4):24, or for a 60-cycle circuit, from 300 to 257 rpm. If normal speed is desired, the smaller motor is simply cut out of the circuit. The smaller motor may be 4, 6, 8, or 10 poles, as required. These motors have been built up to 500-hp capacity for steel-mill main-roll drive.

85. Two- or Three-Speed Motors. Either the squirrel-cage or the slip-ring type of motor can be arranged with one or more sets of windings so as to have different synchronous speeds. This is known as *pole changing*. That is, a motor may be rated $^6/_{12}$ poles, $^{100}/_{50}$ hp ($^{75}/_{38}$ kW), $^{1200}/_{600}$ rpm. One method of stator connection is shown in Fig. 16.66. This connection is for variable torque, two-speed, three-phase. There are a number of other connections. This type of motor is considerably more expensive than a single-speed motor. Pole changing may be accomplished by a knife switch or by contactors.

86. Adjustable-Speed Brush-Shifting AC Motors. These motors may be of either the series or shunt speed characteristic type. They are used where an adjustable-speed ac motor is required and the slip-ring type does not suit or is of too low efticiency at low speeds.

Series Type. The motor with a series speed characteristic consists of a stator winding like that of a squirrel-cage motor. The rotor is wound like a dc armature with a commutator, except that it is wound for three or more phases. If the power line is three-phase, the rotor may be wound for an increased number of phases, the only limit being the complexity of brush gear and transformer. The number of brush-holder stud locations equals the number of pairs of poles times the rotor phases, but with a large number of phases, part of the studs for each phase are
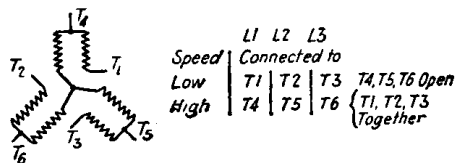
FIGURE 16.66 Connections for two-speed motors.

sometimes omitted. Power is applied to the primary or stator leads, each phase of the stator winding being in series with one phase of the primary of a rotor trans-former, the secondary of this transformer being in series with the rotor windings. Figure 16.67a shows this connection.

If the flux set up by the primary and that set up by the secondary winding have their axes in the same line, they oppose each other and no torque will be developed. For this reason, it is necessary to shift the brushes so that there will be a resultant component of these fluxes or magnetic field in a direction to produce torque. This torque is caused by the repulsion between the current in the rotor coils and the magnetic field. A set of speed-torque curves is shown in Fig. 16.67b, each curve being marked with the degrees of brush shift. In this motor, the maximum torque is developed with a brush shift of about 15°.

With the brushes set to give full-load starting torque, the starting current is about 150 to 175 percent. The efficiency at full load, full-load speed is slightly less than that of a slip-ring motor. At lower speeds, the efficiency is better than that of the slip-ring motor with secondary resistance. The cost is higher. The power factor is better than that of a slip-ring motor. These motors are to a great extent superseded by the shunt-type motors for constant-torque loads.

For small sizes, the motors may be thrown across the line, with the brushes in the low-speed position. For larger sizes, a Y-delta switch is sometimes required to give easier starting conditions. Therefore, the only controls besides the rotor trans-former are a line switch and possibly a Y-delta starting switch.

Shunt Type. The stator winding (the secondary) is constructed like the stator (primary) winding of an induction motor, except that phases are electrically inde-pendent and both ends of each phase are brought out for connection to the com-mutator brushes. The rotor has two windings. One rotor winding (primary) is identical in construction with the stator (primary) winding of a normal induction motor and is connected to the collector rings, to which the power is applied. The adjusting winding, which is in the top of the slots occupied by the primary winding, is connected to the commutator in the same manner as in a dc motor (Fig. 16.68).

The motor may be compared with a wound-rotor induction motor having its primary windings in the rotor and its secondary on the stator. The motor is provided with two brush-holder yokes arranged to shift in different directions. One end of each phase of the stator (secondary) winding is connected to brushes on one brush
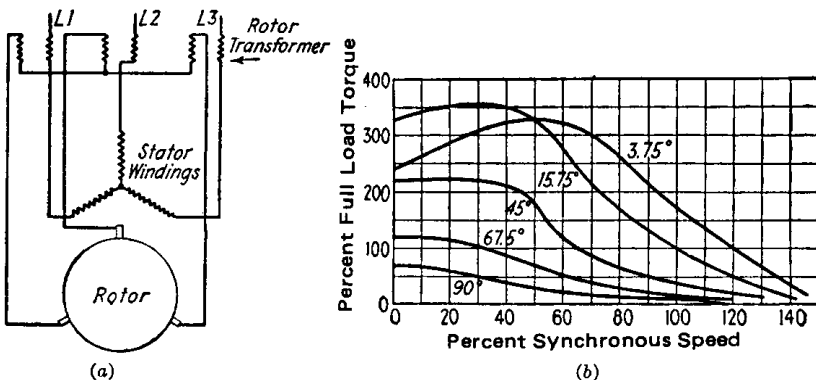
FIGURE 16.67 Connections (*a*) and speed-torque curves (*b*) of series-type motor.
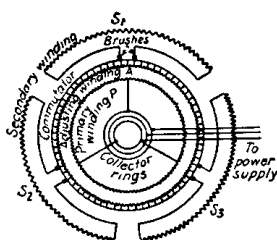
FIGURE 16.68 Connections for shunt-type adjustable-speed ac motor.

yoke, and the opposite ends are connected to brushes on the other yoke. When the brushes, to which each end of a secondary phase is connected, are on the same commutator segment, the adjusting winding is idle, the secondary winding is short-circuited, and the motor runs as an induction motor with speed corresponding to the number of poles and frequency of supply. As the brushes are moved apart, a section of the adjusting winding is included in series with the secondary winding, causing the secondary winding to generate a voltage to balance the voltage impressed upon it by the adjusting winding, so that the motor changes its speed. Moving the brush gear in one direction raises the speed, and moving it in the other direction reduces the speed. The motor operates both above and below the induction-motor synchronous speed.

With the brushes in the low-speed position, these motors give, according to the particular design, from 140 to 250 percent of normal torque at starting with 125 to 175 percent of full-speed line current. The maximum torque at low speeds is usually, for 3:1 speed-range motors, about the same as the starting torque, and increases for the high-speed position to from 300 to 400 percent of normal torque. The efficiency remains nearly constant over the greater part of the speed range but is somewhat lower at low speeds. The average efficiency is high as compared with that of wound-rotor induction motors with secondary resistance. The power factor is very high when the motor is running at high speeds. At synchronous speed, the power factor is approximately the same as that of an induction motor of similar rating. Many motors have a form of brush-shifting device such that the high power factor obtainable at top speed is maintained down to speeds somewhat below synchronous speed.

The decrease in speed from no load to full load is, at high speeds, from 5 to 10 percent, and at low speeds from 15 to 25 percent, according to the rating of the motor. Since the primary power is taken into the machine through collector rings, it is necessary to use 550 V or less, since no collector rings have yet been developed for higher voltages. This motor may be used for reversing service. The brushes are usually shifted by hand but may be shifted by a pilot motor. Figure 16.69 shows the speed-torque curves of a 7½-hp (5.6-kW) motor. It will be noted that the starting torque is very high.

Since these motors may be thrown across the line with the brushes in the slow-speed position, a line switch is all that is required for the control. For some special applications, the stator windings are opened and a resistor is inserted. This gives the same effect as the resistor of a slip-ring motor and requires similar control, although more than one resistance point is seldom needed, since the speed can still be regulated by brush shift.
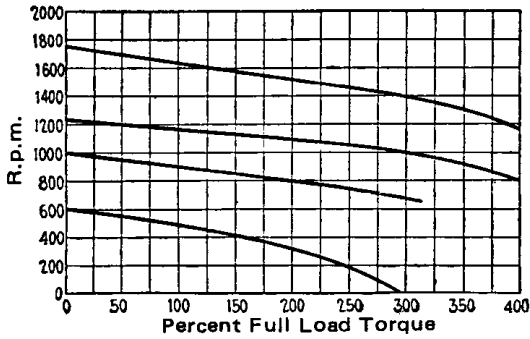
FIGURE 16.69

87. Single-Phase Motors.   Single-phase motors are built in a number of different types; they may be divided into those types having a *series* speed characteristic and those which have a *shunt* characteristic. Owing to their speed characteristics, series motors should not be belted, because if the belt should come off the motor would run away.

Universal Motor.   The *universal motor* is essentially a dc series motor. As first built, it had salient poles, but later types have a single-phase distributed stator winding like an induction motor. The rotor is a series-wound dc armature with a commutator. The stator winding and armature or rotor winding are in series (Fig. 16.70). The torque and speed characteristics are therefore like a series motor; i.e., the motor will develop high torque at low speeds, but at no load the speed is high. The speed may be varied by control of the line voltage (see Fig. 16.71). Applications are sewing machines, floor polishers, etc.; capacities, $\frac{1}{200}$ to 1 hp (3.7 to 746 W). Large motors of this fundamental type are used for railway motor service, but these usually have compensated or commutating windings and thus become more complicated. Since this is a very special application, any questions should be referred to a manufacturing company.

Repulsion Motor.   A *repulsion motor* is a single-phase, commutator-type motor with a single-phase stator winding. The rotor is a dc multiple-wound armature with commutator, with the brushes placed 180 electrical degrees apart and short-circuited. With the brushes on neutral, the magnetic axis of both stator and rotor
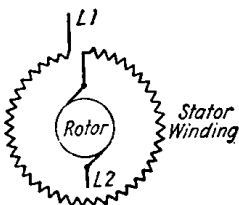
FIGURE 16.70   Universal motor.
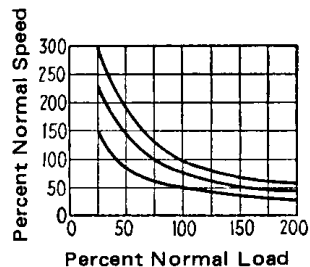
FIGURE 16.71   Speed-torque curves of repulsion motor.

is in the same line; therefore, no starting torque is developed. It is therefore necessary to shift the brushes about 20° out of phase with the primary winding to obtain the maximum starting torque. The brushes are fixed in position. This motor also has series speed-torque characteristics. Adjustable varying speed may be obtained by voltage control as shown in Fig. 16.71. The power factor of a 3-hp (2.2-kW) 1800-rpm motor would be about 0.8. Commutation at full load is very good.

Repulsion Brush-Shifting Motors. These are similar to the preceding motor but have arrangements for shifting the brushes. With the brushes set as before, the motor would have the same speed-torque curve. With any other brush position, the speed-torque curve would be similar. Thus a series of these curves may be obtained similar to those shown in Fig. 16.71 for change in voltage. Over a speed change of about 2.5:1, the motor gives constant torque. This motor can be used for reversing service by brush shifting.

Split-Phase Motors. The stator is wound as in a two-phase, squirrel-cage motor. The two phases are connected in multiple, with a resistance or reactance in series with one of them, to obtain starting torque (see Fig. 16.72). The starting winding is for intermittent service and is cut out by a centrifugal switch. Capacities are $\frac{1}{30}$ to $\frac{1}{4}$ hp (25 to 187 W). These motors are for constant-speed applications, such as for fans, blowers, washing machines, and other domestic appliances. The torque curve is somewhat like that of a squirrel-cage motor, starting at 75 to 100 percent torque at zero speed and reaching 200 to 250 percent at about 85 percent speed. When the starting winding is cut out, there is a drop in torque.

Capacitor or Condenser Motor. This motor is similar to the preceding motor but has a condenser instead of a resistor in series with one phase, this being left permanently in circuit. This motor has a low starting torque, 50 to 100 percent, but may be used on applications similar to those for the preceding split-phase motor. These motors have the advantage of not having either a centrifugal switch or a commutator. The power factor is good, approaching unity in some cases. These motors may be built up to 25 hp (18.7 kW), but are much more expensive than a three-phase motor of the same rating. They may be started by throwing directly on the line.

Another variation of the condenser motor has a second block of condensers in parallel with the starting condenser. This gives high starting torque, from 100 to 300 percent, depending on the condenser used. It requires a centrifugal switch to cut out the starting condenser, or an extra switch in the starter, or a normally closed contactor with its coil connected across the starting winding; this gives low voltage at starting, which increases with the speed. When the voltage reaches a certain value, the contactor opens. These motors are used where commutator motors are not desired.
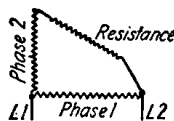
FIGURE 16.72 Connections for split-phase motor.

Repulsion-Induction Motors.   Here the stator winding is of the distributed single-phase type. The rotor is a combination of the repulsion type and the high-reactance squirrel-cage type, both windings being in the same slots. These windings are so arranged that during starting, the repulsion characteristics predominate, giving high starting torque. During running, the squirrel-cage winding predominates, giving nearly constant speed. Some builders use a switch to change from the starting to the running condition, while others make the change inherent in the machine. Owing to the combination winding, the running torque is very high. Figure 16.73 shows a torque-speed curve for a 3-hp (2.2-kW), 1800-rpm motor. The no-load speed is slightly above synchronism (about 3 percent), and the full-load speed is an equal amount below.

Since the motor operates, after starting, as a squirrel-cage motor, only a small number of brushes are required. Hence the motor is very quiet. The commutation is practically perfect. The efficiency should be between 75 and 80 percent. The power factor varies up to 95 percent. Speed control may be obtained by varying the line voltage.

Control.   Practically all single-phase motors are thrown directly on the line, usually by a simple snap switch. If speed control or reversing is required, the starter must be amplified accordingly.

88. Synchronous Motors.   A synchronous motor is built very nearly like an ac generator. It has a distributed stator winding and a rotor winding which is usually on salient poles. It differs from a generator in that it has a squirrel cage built into the rotor for starting and to damp out oscillations or hunting. The windings on the poles are the field and are excited from direct current. A few large and very high speed motors have a distributed rotor winding similar to that of a slip-ring induction motor. This is necessary for mechanical reasons.

When *starting,* the field or rotor circuit is open, and the motor operates as a squirrel-cage induction motor. Since the torque curve is similar to that of a squirrel-cage motor, it may be varied by changing the resistance and reactance of the squirrel cage. When the motor has reached about 95 percent of synchronous speed, the field is energized and the rotor pulls into step or synchronous speed. It then rotates at the same speed as the revolving field set up by the stator or primary winding.
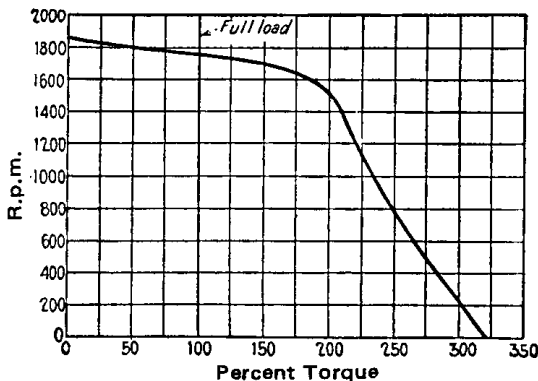
FIGURE 16.73   Speed-torque curve of repulsion-induction motor.

During running, the speed is constant and is fixed by the frequency and number of poles. The speed is not affected by change of load or field strength. Changing the field, however, does change the power factor. If load is applied beyond the breakdown point, the rotor falls out of step and stops, since it has no torque below synchronous speed.

Synchronous motors are being used for a greater variety of services each year because they improve the power factor. They must therefore be designed for their particular duty. A normal motor for industrial service has approximately 50 to 100 percent starting torque, 100 percent pull-in torque, and 175 percent pullout torque. *Pull-in torque* is that value developed by the squirrel-cage winding at a speed high enough to put on the field, usually about 95 percent. *Pull-out torque* is the maximum the motor will carry without falling out of step.

Motors with *high starting torque,* up to 175 percent, are sometimes necessary for certain applications, such as ball- or rod-mill drives. High starting torque means a motor of larger size to meet this particular requirement or a special type of motor. One method is to have the motor connected to the load through a *magnetic clutch.* Sometimes this clutch is built with the motor; sometimes it is separate. With this arrangement, the motor starts without load. After it is up to speed and the field is applied, the clutch is energized, bringing the load up to speed. This allows the motor to use its pull-out torque when starting the load.

A *supersynchronous* motor is sometimes used in which the rotor is solidly coupled to the load and the stator is placed in bearings. When power is applied to the motor, the stator revolves and comes up to synchronous speed. The field is applied and the motor is synchronized. A brake is then applied to the stator, bringing it to rest. The rotor, being held in synchronism with the stator, starts and comes up to speed. By this method, the pull-out torque is also available for starting the load.

A synchronous motor has the inherent characteristic of having its power factor dependent on the load and field strength. For a given load, an increase in field strength gives a leading power factor, and a decrease in field strength gives a lagging power factor. Figure 16.74 shows the relation between line current and field excitation and the resulting power factor.

In Fig. 16.74, curves 1, 2, and 3, line current is plotted against excitation. The low point of each of these *V*-curves gives unity power factor at that line current.
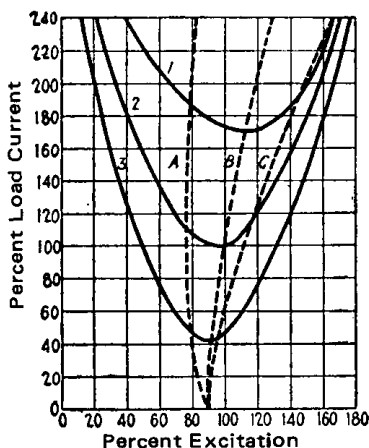
FIGURE 16.74 Excitation p.f. curves.

The field current necessary to give unity power factor at normal line current is the normal field current of the motor. Curves *A*, *B*, and *C* are curves of equal power factor. In order to hold the power factor at unity or leading, it is necessary to increase the field with load. For a motor with fixed excitation, the power factor varies with the load. With normal excitation and light load, the power factor is leading; as the load increases, the power factor reaches unity, and then becomes lagging.

As stated in Sec. 88, the impressed voltage is equaled by the induced voltage plus the *IR* drop. This statement is nearly true of synchronous motors, but there are other variables. Since the speed is fixed, the induced voltage varies with the field strength. The third member of the equation, the *IR* drop, is now superseded in importance by the *impedance drop ZI,* which is always nearly 90° out of phase with the current. Figure 16.75*a* shows these voltages, where *E* is the applied voltage and *e* is the back emf or induced voltage. Figure 16.75*a* is for unity power factor, and thus the current *I* is in phase with the applied voltage, and *IE* equals power or kilowatts. If when holding constant load the field is increased, then *e* is increased and the diagram becomes as shown in Fig. 16.75*b*. Since the power remains constant, the current in phase with the applied voltage remains the same. Therefore, the line representing the current, times cos *θ*, must equal the original current. *ZI* is always nearly 90° out of phase with *I* and must increase as *I* increases, to I_1 or $ZI_1 = ZI/\cos θ$. Completing the triangle gives e_1. The angle *θ* represents the power factor and in this case is leading. If the field is decreased, then *e* must decrease to e_2, which may be determined by the same reasoning. This, shown in Fig. 16.75*c*, gives a lagging power factor.

Synchronous motors are often used for power-factor correction. The motor usually does some mechanical work and is given a horsepower rating on this basis. It is then designed to carry a leading current so as to offset or neutralize the lagging current required by induction motors. The motor is rated 0.50 or other power factor to show that it can deliver leading current. Sometimes synchronous motors are used running idle on the line. All their capacity is then used to give leading current. They are then known as *synchronous condensers.*

89. Synchronous-Induction Motor. *Synchronous-induction motors* are, as the name implies, a combination. The stator is like that for a synchronous motor, and the rotor is usually wound as in a three-phase slip-ring motor. During starting, a three-phase secondary resistance is used, and the motor starts like a slip-ring motor. When full speed is reached, the rotor is excited by direct current like a synchronous motor. Since the rotor is wound three-phase, its winding when excited from direct current is not so efficient as the field winding of a synchronous motor, since one phase is either cut out of circuit or is connected in multiple with another phase and with two phases in series.

This motor is slightly more expensive than either a synchronous motor or a slip-ring motor, but the worst objection to it is that its excitation is at very low voltage
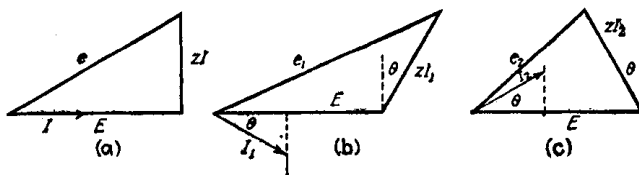
FIGURE 16.75   Synchronous-motor diagrams.

and the exciter is expensive. It has the advantages of slip-ring-motor starting characteristics and synchronous-motor operating characteristics. Owing to the cost of the motor and exciter, this type of motor is not widely used in this country.

DIRECT-CURRENT MOTORS

90. Types. A dc motor is built like a dc generator. A machine acting as a motor has a definite applied voltage. This voltage is partly used in overcoming *IR* drop through the armature coils and brushes, and the remainder is opposed by the back emf generated by the armature windings cutting through the magnetic field. Increasing the main field strength increases the magnetic field and causes the motor to develop the required back emf at a lower speed. In other words, strengthening the field lowers the speed, and weakening the field increases it.

A dc motor is in reality a type of ac motor with a commutator so as to use zero-frequency current. There are three types: *series,* having the fields in series with the armature; *shunt,* having the fields in multiple with the armature; and *compound,* having a combination of series and shunt fields.

91. Series Motors. The *series motors* give a drooping speed-torque curve; i.e., the greater the load, the less the speed. A typical curve is shown in Fig. 16.71. This motor develops a high torque per ampere during starting, because the field, being in series with the armature, is strengthened with increased load, and thus the magnetic field is stronger. The speed increases rapidly with decreasing load and would generally be sufficient to wreck the motor at no load. This type is especially desirable for frequent starting with heavy starting torque, as with hoists, auxiliary motors for steel-mill service, etc. In small sizes it may be thrown directly on the line, but for medium and large motors a starting rheostat is needed. There is no real limit to the size which can be built, but small or medium size is usual, although railway motors are almost always of this type and are built up to several hundred horsepower.

92. Shunt Motors. *Shunt* motors have constant field strength. Since the motor does not develop any back emf at standstill, it is necessary to use a starting resistance to limit the current at this point. As the speed and back emf increase, the resistance can be cut out, thus controlling the current which can pass through the armature or the torque to be developed. During normal operation this motor has nearly constant speed; the only change is due to the increased *IR* drop with increased load. This means slightly lower speed. If a flat speed curve is desired, a slight shift of the brushes, off the neutral, will give sufficient compounding effect. This type of motor is used for all ordinary work. It can be reversed easily and so is suitable for hoist work. It can be operated at different speeds by field control, within the limits of commutation (see Sec. 93) and also may be used at very low speeds with armature resistance or voltage control. It can be built in any size required. For operation by generator voltage control, the motor must be excited from a separate generator (see Sec. 119).

93. Compound-Wound Motors. A *compound-wound motor* is a combination of the two preceding types, having some of the characteristics of each. By varying the percentage of series field, the shape of the speed-torque curve can be varied to meet

any requirements. Generally speaking, the amount of series field used is very small. Sometimes the series field is in opposition to the shunt field, and the motor then is said to be *differentially wound.*

94. Braking.   A series motor cannot be used as a *regenerative brake,* as a study of Fig. 16.71 will show. Compound-wound motors can be so used but rarely are. Shunt-wound motors can be used to obtain regenerative braking with no change in connections. Any dc motor can be used with *dynamic braking.* A diagram is shown in Fig. 16.76 for a compound-wound motor. The line and dynamic-braking contactors *LE* and *DB* are mechanically interlocked. The dynamic-braking contactor coil is energized as soon as the line contactor closes. One accelerating contactor is shown; the interlock *A* may be controlled by a time relay or by other means.

95. Adjustable-Speed Motors.   An *adjustable-speed dc motor* is usually shunt wound, equipped with commutating fields, and sometimes with pole-face windings. The speed is changed by varying the shunt field. A motor for this service is usually built with a larger flux and fewer armature conductors than one for constant speed. Speed ranges of 3:1 and 4:1 are standard. If a greater range than 4:1 is desired, the lower range in speed must be obtained by cutting resistance into the armature circuit, or it may be necessary to use generator voltage control.

96. Table of Uses of Motors.   Quite often the user of motors is at a loss as to what type of motor to specify for a certain service. The decision is influenced by several factors, among which are

1. The *source of power available:* Is the power dc or ac, and if ac, is it single-phase or polyphase?

2. *The speed requirement:* Is the equipment to be run at constant or variable speed? If the speed is constant, must it be absolutely constant, or may a slight decrease with load be tolerated? If the speed is variable, what range of variation is required? Should the speed decrease considerably with the increase of load, as with a traction motor?

3. *Starting torque:* Is the motor to be subjected to heavy starting torque, or will the starting torque be light?

4. *Type of load:* Is the load steady or widely fluctuating? Is it a generally heavy load, or is it a light load?
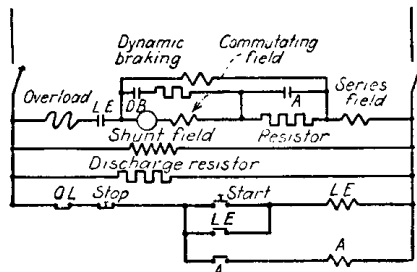
FIGURE 16.76 Connections for dynamic braking.

The balancing of these factors with the service and characteristics of the motor is not an easy task. Table 16.11 has been prepared to lighten the task of the motor buyer. It is well, however, when the purchase involves a large sum to engage the services of a consultant who will see to it that proper motors are specified for the particular types of service.

CONVERTERS

97. Introduction. In Secs. 53 to 66, descriptions were given of machines or apparatuses for obtaining two types of emf sources—the continuous or direct source of emf and the alternating source of emf. One source may be converted into the other by one of three general types of converters: (1) the rotary (or synchronous) converter, (2) motor-generator sets, and (3) the electronic converter. Such converters play an important role in electrical engineering practice, because practically all power is now generated as alternating current, and since there is a demand for dc power for certain services—traction, battery charging, three-wire dc system of lighting, etc.—it is necessary to have some method of changing ac power to dc.

THE SYNCHRONOUS CONVERTER

98. Definition. The *synchronous converter* is a combination synchronous motor and dc generator. It has only one set of armature conductors and one magnetic circuit. Being a dc generator, it is built with a revolving armature. The armature winding is the same as for a dc generator, and the collector rings are tapped off at equidistant points on the back end of the winding. On the front end is the usual commutator. Nearly all converters have commutating fields.

For any given machine with a fixed direction of rotation and fixed direction of the field, the armature current is in one direction when the machine is operating as a generator and in the opposite direction when operating as a motor. In a synchronous converter which operates as a synchronous motor taking power from the ac line and also as a de generator giving out power to the dc line, the currents tending to flow through the armature conductors are in opposition, and therefore, the actual amount flowing in these conductors is the difference.

99. Rating. Because of the preceding, a machine designed as a dc generator can operate at a higher rating when used as a synchronous converter. The amount of increase in the rating depends on the number of rings or phases on the ac end. If all losses are neglected, the theoretical ratios are

	Percent
dc generator	100
Single-phase single-circuit converter	85
Single-phase two-circuit converter	93
Three-phase converter (3 rings)	134
Quarter-phase converter (4 rings)	164
Six-phase converter (6 rings)	196
Twelve-phase converter (12 rings)	224

TABLE 16.11 Classification and Uses of Electric Motors

Source of power	Type of motor	Speed variation in per cent of rating	Torque, in per cent of normal — Starting torque	Torque, in per cent of normal — Maximum torque	Standard ratings, hp.	General types of service for which motor is adapted
D.c.	Shunt-wound	5 to 7.5	150	L.C.*	1/6 up	Steady loads
	Compound-wound	10 to 25	175 to 200	L.C.	1/6 up	Heavy starting and fluctuating loads
A.c. 1-phase	Split-phase induction	5 to 7.5	150 to 200	150 to 175	1/20 to 1/3	Infrequent starting steady load
	Repulsion	4 to 6	225 to 300	150 to 175	1/6 to 10	L.S.C., H.S.T.*
	Capacitor	4 to 5	200 to 225	125 to 150	1/6 to 10	Quiet operation with high p.f.
A.c. 2- and 3-phase	Std. squirrel-cage	3 to 6	150	200	1/6 up	General purposes
	Sq.-cage N.S.T., L.S.C.*	3 to 6	100 to 150	200	7½ to 200	General purposes (single control)
	Sq.-cage H.S.T., L.S.C.*	4 to 6	175 to 275	175 to 200	3 to 100	Heavy starting (single control)
	Sq.-cage L.S.T., L.S.C.*	4 to 6	50 to 60	125 to 150	40 to 100	Constant load
	Wound-rotor	D.L.*	200 to 250	175 to 250	1/4 up	Frequent and heavy starting L.S.C.*
	Synchronous	None	80 to 125	150 to 175	20 up	Infrequent starting, high p.f.
	Low-slip sq.-cage	4 to 6	175 to 200	175 to 200	½ to 100	Heavy starting (fluctuating load)
	H.T. sq-cage*	10 to 15	250	250	½ to 50	Intermittent service
D.c. Shunt-wound Multispeed	Field- and arm-control	D.L.	150	L.C.	½ up	Wide speed range, flexible control
	Voltage-control	D.L.	150	L.C.	1/2 up	Wide speed range, flexible control
A.c. Multispeed	Constant-power	D.L.	150	175 to 200	1/4 up	Speed is dependent on load
	Constant-torque	D.L.	150	175 to 200	1/4 up	Speed is dependent on load
	Variable-torque	D.L.	150	175 to 200	1/4 up	Speed is dependent on load
D.c. Compound-wound Variable-speed	Field- and arm-control	D.L.	150	L.C.	1/4 up	Wide speed range, flexible control
	Voltage-control	D.L.	150	L.C.	1/4 up	Wide speed range, flexible control
	Series-wound	D.L.	300 to 400	L.C.	1/4 up	Wide speed range, flexible control
	Wound-rotor	D.L.	200 to 250	175 to 250	3 to 50	Heavy starting duty, L.S.C.
A.c. Variable-speed	Brush-shifting	10 to 30	150 to 200	150 to 200	1/4 to 50	Wide speed range (high cost)
	High-slip sq.-cage	D.L.	250	250	40 to 100	Heavy starting (fluctuating load)

Uses — special types of service for which motor is adapted: Agitator, Baker (power), Ball mill, Blowers, Boring mill, Bridge (coal), Car puller, Compressor, Conveyor, Crane hoist, Crusher, Dough mixer, Driller, Dryer, Elevator (freight), Elevator (passenger), Fans, Flat ironer, Grinder, Hammer (power), Lathe, Laundry extractor, Laundry washer, Line shaft, Milling machine, Mine hoist, Molders, M.G. set a.c. to d.c., M.G. set d.c. to a.c., Planers (joiners), Planers (metal), Printing press (flat-bed), Printing press (job), Printing press (rotary and offset), Pumps (shears), Pumps (centrifugal), Pumps (displacement), Punches (shears), Sanders, Saw (band), Saw (circular), Saw (metal), Shapers, Shears (punches), Stokers, Tenoners, Turntables, Valves, Vehicles.

N.S.T. = normal starting torque, H.S.T. = high starting torque, L.S.T. = low starting torque, L.S.C. = low starting current, D.L. = depends on load, L.C. = limited by commutation, H.T. = high torque

Conversion factor: 1 hp = 0.75 kW.

16.86

These figures are based on unity power factor or on the load current being in phase with the voltage. At the present time, nearly all converters are wound for six-phase and are connected diametrically as shown in Fig. 16.77*a*, the circle representing the armature winding and the rings being tapped off 60° apart. One advantage of this winding is that the center points of the transformer coils can be connected together and the machine can be used on a three-wire system. Occasionally, the six rings are connected double-delta (Fig. 16.77*b*), and in smaller machines three-phase (Fig. 16.77*c*) is sometimes used.

For 600 V and above, converters are being largely superseded by polyphase rectifier units.

100. Voltage Ratio. No-load ratios of ac to dc emf's vary with the number of phases, being approximately proportional to the sides of polygons inscribed in a circle. Divergence from the sine wave and some other conditions will produce slight differences. The theoretical no-load values for 100 V dc are given in Table 16.12. These are the volts between adjacent rings, or the phase voltages. Note that the quarter-phase, or so-called two-phase, is really a four-phase and is properly represented by the inscribed square. The diametral voltage ratio of the four-, six-, and twelve-ring converters is always that of the single-phase, 70.7. Slightly higher ac voltage is necessary for the same dc voltage when a converter is loaded.

101. Frequency. Since a converter is a synchronous machine, its speed is determined by the number of poles and the frequency. It can be built for any commercial frequency. The lower the frequency, the more reliable can the converter be built. Converters as now built operate successfully on 60-cycle circuits.
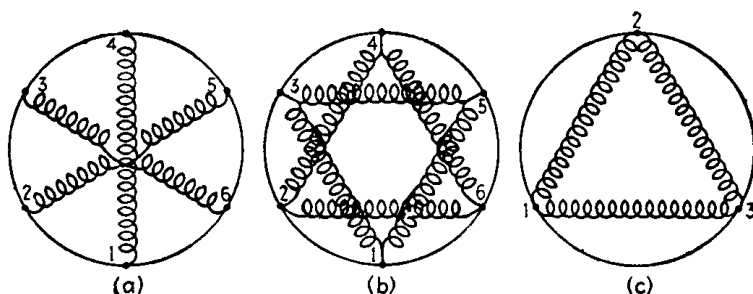
FIGURE 16.77 Connections for rotary converter armature windings and collector rings.

TABLE 16.12 Ratio of Alternating-Current Voltage to Direct-Current Voltage

	Two-ring	Three-ring	Four-ring	Six-ring	Twelve-ring
Phases	1-phase	3-phase	4-phase	6-phase	12-phase
No-load ac volts	70.7	61.2	50	35	18.0

MOTOR-GENERATOR SETS

102. Definition.   A *motor-generator set* is a device used to convert alternating current to direct current, and vice versa. It consists of two machines whose shafts are mechanically coupled either directly or through a belt or chain of gears. One of these machines is a dc machine and the other an ac machine. If alternating current is to be converted into direct current, the ac machine is used as a motor and the dc machine as a generator. If direct current is to be converted to alternating current, then the dc machine is used as a motor and the ac machine becomes a generator. Motor-generator sets are also used to convert from one dc voltage to another and from one ac frequency to another.

103. Synchronous Converters vs. Motor-Generator Sets.   Since synchronous converters are used to convert alternating current into direct current, and since motor-generator sets can be used for the same work, a comparison of the two methods is desirable. Synchronous motor-driven sets only will be considered. *Advantages of converters* are lower first cost, less floor space, higher efficiency, only one machine to keep up, and dc voltage independent of frequency. Since transformers are always necessary with converters, the cost and losses of the transformers should always be included with the converter. *Advantages of the motor-generator set* are the dc voltage is independent of the ac voltage, less liability of commutation troubles, power-factor correction, and up to 13,200 V, no step-down transformers are necessary.

The efficiency of large synchronous converters runs as high as 95 percent and the efficiency of transformers runs up to 99 or even 99½ percent, so that the combined efficiency may vary from 94 to 95 percent. Motor-generator sets of the same capacity have an efficiency up to 91 or 92 percent.

MERCURY-ARC RECTIFIERS (*CONVERTERS*)

104. Mercury-Arc Rectifiers.   *Mercury-arc rectifiers* are largely replacing synchronous converters and motor-generator sets for converting alternating current to direct current. A mercury-arc rectifier makes use of the fact that electricity can pass through a mercury arc in one direction only when operating in a high vacuum. They are common in sizes from 30 to 3000 kW. The ac power line, may be any voltage and any frequency. The dc voltage may be from 250 to 3000 V.

This rectifier consists essentially of a tank containing two or more terminals or anodes in the top or sides for the incoming ac power and one terminal or cathode in the bottom for the outgoing or dc power. The cathode is covered by liquid mercury, and the interior of the tank is a high vacuum. Figure 16.78*a* shows a simple diagram for a single-phase rectifier. From the diagram it can be seen that the center of the transformer secondary winding is brought out for a neutral and the negative dc line. Each end of the transformer secondary winding is connected to an anode of the rectifier. An arc is drawn between the cathode and the exciter terminal shown between the anodes. For one-half the cycle, the current passes through an arc from anode 1 to the cathode, and for the other half it passes from anode 2 to the cathode. This gives a wave as shown for single-phase in Fig. 16.78*b*.
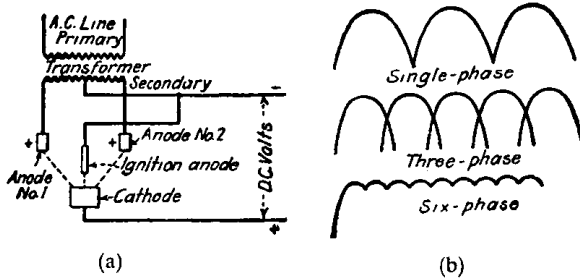
FIGURE 16.78   Simple mercury-arc rectifier.

Power rectifiers are built for 6 or 12 phases with all the electrodes in one large steel tank, or each anode is insulated in a separate tank with a mercury cathode in the bottom; these are connected together to form one terminal of the dc circuit.

105. Mercury-Arc vs. Rotary Converters.   The mercury-arc rectifier has several advantages over competitive equipment. Since there are no rotating parts, the operation is noiseless; this is a vital consideration in some locations. The overload capacity is high. Even a short circuit on the dc side has little effect on the mercury-arc rectifier itself, although it may impose a heavy load on the power system. It is not necessary to phase in a rectifier, since it has no rotating parts. Very little attention is required, since the operation of a rectifier is automatic.

CONTROL AND PROTECTIVE DEVICES AND SYSTEMS

106. Types of Protective Devices.   Protective devices may be grouped under headings as follows: overload or short-circuit protection, phase-failure or reversal protection, undervoltage protection or release devices, overspeed, bearing-temperature relays, limit switches, etc.

107. Overload Protection.   Overload protection for motors may be applied to running overloads and short circuits. For most industrial applications, the running overloads are taken care of by a magnetic switch with thermal overload relays, the short-circuit protection being taken care of by fuses or a circuit breaker.

The thermal relay characteristic is shown in Fig. 16.79. This is called *inverse-time* overload protection. That is, the greater the overload, the quicker the relay will function. Fuses give instantaneous protection, and circuit breakers may have relays for either inverse-time or instantaneous protection.

There are other types of installation where fire hazard is not a determining feature, where for small motors fuses only are used for overload protection. For larger motors, a circuit breaker with inverse-time relay is used. In these cases, the powerhouse and sectional circuit breakers are set for a higher rating so that the branch breaker will open first and thus relieve the system.
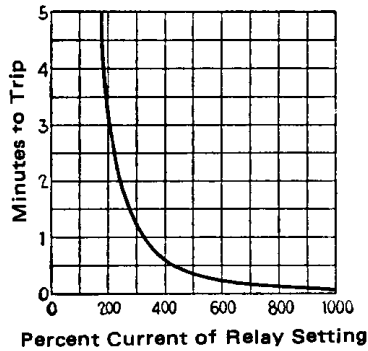
FIGURE 16.79   Thermal relay characteristics.

108. Phase-Failure and Reversal Protection.   Phase-failure and reversal protection are accomplished by use of relays with the coils so balanced that the tips remain closed when the three phases are normal; if any phase is open or reversed, the relay opens and thus opens the circuit breaker or line contactor.

109. Undervoltage Protection.   Undervoltage protection is obtained by the use of an undervoltage coil on the circuit breaker. The undervoltage coil and the overload coil may operate the same mechanical trip. Undervoltage protection also may be obtained by the use of a contactor whose tips carry the control circuit. After failure of power, the undervoltage relay or contactor has to be closed manually.

110. Overspeed Protection.   Overspeed protection is usually obtained by a centrifugal switch on the motor. This switch opens a control circuit, usually in the undervoltage circuit of the oil circuit breaker or contactor.

111. Bearing-Temperature Relays.   Bearing-temperature relays are used with automatic pumps, substations, etc. They usually consist of a bulb inserted in the metal of the bearing. The bulb and a connecting tube are filled with gas. An increase in bearing temperature causes the gas to expand and push a contact open and thus break the control circuit.

112. Limit Switches.   For some appliances such as cranes, hoists, machine tools, etc., it is necessary to limit the travel of the cage, hook, or table. Limit switches are used for this purpose. They may be divided in two general classes: track or hatchway and geared. *Track-type* switches consist of a cylinder or cam operated by a handle or rope wheel. When the traveling object hits the arm or handle, it turns the cylinder or cam and opens or closes one or more electric circuits. *Gear-type* switches consist of a set or train of gears, sometimes with a traveling-nut device, that opens or closes electric contacts after a certain distance has been traveled. Both types are made in a wide variety of mechanical designs and with from one to five or six circuits.

MOTOR CONTROL

113. Starting and Speed-Regulating. Starting and speed-regulating methods depend first on the type of motor and second on the work done. Today there is an endless variety of control arrangements.

114. Single-Phase Motors. Small single-phase motors are usually started by a snap switch or a knife switch, and if there is any overload protection, it is in the form of fuses.

115. Squirrel-Cage Motors. Squirrel-cage motors up to 5 hp (3.7 kW) are thrown across the line either by a knife switch or, more often, by a three-pole contactor, controlled by a start-stop push button, with a thermal-overload relay for protection. Motors up to 200 hp (148 kW) are sometimes thrown across the line by use of a contactor, as mentioned above, or a hand-operated switch. For a great many applications, reduced-voltage starting is required. This can be obtained by a resistance in the primary or by the use of ati autotransformer. The resistor-starter is suitable for small motors of 5 to 25 hp (3.7 to 18.7 kW). It is cheaper than the autotransformer or compensator type but has never been used to such an extent.

The autotransformer or compensator type of starter consists of an autotransformer to supply reduced voltage (this usually has taps at 50 to 65 and 80 percent), a five-pole starting switch for connecting the autotransformer to the line and the motor to the transformer taps, and a three-pole running switch. The two switches must be interlocked so that both cannot close at the same time. There are two types of these starters: the manually and the magnetically operated. For the manually operated compensator, the starting and running switches are usually of the knife-switch type. The magnetic type has contactors operated by a push button and with a time-delay relay to control the closing of the running contactor. In both types, the switches are oil-immersed for reasons of safety and small space, and also to make them good for 2200 V or more. Both types have overload and undervoltage devices. The magnetic starter may be remote-controlled by a push button, float switch, or any circuit-closing device.

116. Synchronous Motors. Synchronous-motor starters are the same types as mentioned above for squirrel-cage motors but have an additional device to close the motor field when near synchronous speed and one to short-circuit the field through a discharge resistor during starting and stopping. Larger motors are often started by using a five-pole-three-pole oil circuit breaker in place of contactors. Under certain conditions, these are cheaper.

117. The Korndorfer System. The Korndorfer system is often used for starting very large squirrel-cage or synchronous motors. Referring to Fig. 16.80, *A* is the line contactor, *B* and *C* are the starting contactors. *B* and *C* are closed first, connecting the autotransformer to the line and closing its *Y*-point. The motor now starts on reduced voltage. After the motor has nearly reached full speed, contactor *C* is opened, *A* is closed; then *B* is opened. The advantage of this system is that the motor is not disconnected from the line on the throw from the starting to the running
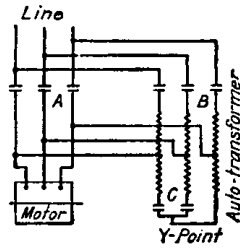
FIGURE 16.80 Korndorfer starter.

position. Contactor *A* simply short-circuits a section of the autotransformer while *B* remains closed.

118. Slip-Ring Motors. Slip-ring motors require control for the secondary or rotor circuit as well as for the primary circuit. This control may be of the manually operated, drum-controller type or of the magnetic type. The simplest equipment consists of a *drum controller* with segments carrying both the primary and secondary currents and a secondary resistance. This type should not be used with over 550 V. It gives no protection to the motor either during ordinary operation, for overloads, undervoltage, etc., or during starting, when the motor can easily be overloaded by cutting out resistance too quickly.

In Fig. 16.63 (Sec. 81) is shown a series of torque curves for a slip-ring motor. The location of each of these curves is fixed by the ohms in that step of the resistor. A resistor is usually expressed in ohms, 100 percent ohms being that value which will allow full-load current to pass through the rotor at standstill. The resistor is designed to give 50, 100, or 150 percent motor torque at standstill, whatever is required by the service. The resistor for the set of curves shown was for a mine hoist and had 300 percent ohms. The weight or mass of the resistor depends on the service. There are light-starting-duty, heavy-starting-duty, and regulating resistors, the resistor becoming progressively heavier for the heavier duty. If very lQw torque is required on the first step, this also increases the weight. The regulating type of resistor must have current capacity to carry the full current continuously.

119. Semimagnetic Control. To obtain *semimagnetic control,* a triple-pole primary contactor operated by the primary segments of the controller may be added to the above. This can be arranged to give overload and undervoltage protection, but still does not protect the motor during starting.

120. Full Magnetic Control. *Full magnetic control* consists of a primary contactor, usually to open all the motor lines; a number of secondary contactors, varying from three to eight; a number of relays; a rotor resistor; and a master switch or master controller which carries only the control circuits. This type of control can be expanded so as to accomplish any operating results. For instance, the master controller may be one, two, or more points to operate any number of contactors. If we consider a six-point master, then there will be six hand-controlled points— the other points controlled by relays. If the master switch is thrown quickly to the ''on'' position, the contactors will close in sequence, but controlled by the relays, thus protecting the motor. There are two systems of control by relays, current-limit and time-limit.

The *current-limit system* is the older and has operated successfully for many years. It consists of two or more current-limit relays connected so as to close one contactor after another. When the line current falls below a predetermined value, usually 150 percent, the relay operates, allowing the next contactor to close. Looking at Fig. 16.63 (Sec. 81), if a load were to be started requiring 100 percent torque, the first three contactors would close in sequence, the current always being less than 150 percent. The motor would start on the third step, but the fourth contactor would close at once and the motor would come up to about 33 percent speed when the current-limit relay would operate, thus closing the next contactor. When two relays are used, they alternate in closing the contactors. Sometimes a third relay is used to give low torque on the first point (see Fig. 16.81). With this control, a light load will accelerate faster than a heavy one. This system is rather complicated.

The *time-limit system* is similar to current limit except that the relays are definite-time relays. There are several types. Settings may be obtained from ½ up to 30 s each. The usual time for each relay is ½ to 5 s. While this system is independent of the current, it gives results that are comparable with current limit. This system is cheaper because the wiring is simpler and the relays require less space on the panel, although there is usually one for each secondary contactor (see Fig. 16.82 for diagram). Either system operates successfully. At the present writing, the current-limit system is more popular.

Magnetic control is especially adaptcd for use with automatic or semiautomatic schemes. For instance, a push button located in any convenient place can be used to start up one or any number of motors. If more than one motor is started, there is *sequence operation.* The final accelerating contactor on the first motor closes the first contactor of the second, and so on. The reverse may be true; if any motor of a group shuts down, owing to overload or any of its protective devices, the other motors (all or any number) can be shut down. This is desirable if a group of motors is driving conveyor sections, part of a long conveyor, or in any mill where the material is carried through several processes. This system is also used for the au-
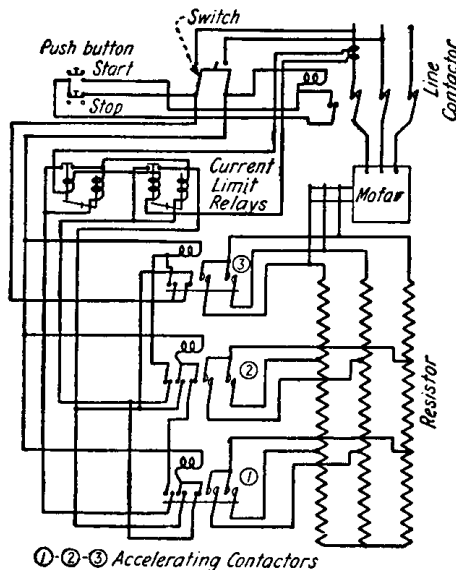
FIGURE 16.81 Current-limit control for slip-ring motors.
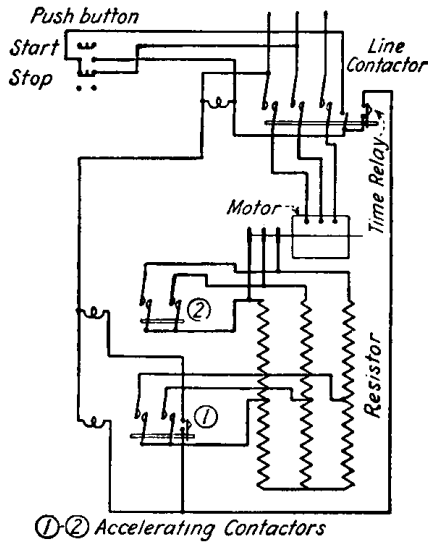
FIGURE 16.82 Time-limit control for slip-ring motors.

tomatic operation of centrifugal pumps, the control circuit being closed or opened by a float switch. When the water in a tank rises to a certain point, the float switch closes, starting up the motor and pump.

121. Speed-Regulating Control. *Speed-regulating control* of slip-ring motors may be obtained by the use of any of the before-mentioned systems. The only requirements are that the regulating steps must be hand-controlled, by either push button, controller, or master switch, and the resistor and contactors must be of sufficient capacity to carry the current continuously.

122. Reversing Control. *Reversing control* can be obtained with any of the above systems. The secondary control would not be changed, but the primary control must have reversing segments on the controller or master switch or have reversing primary contactors. Reversing contactors should be interlocked so that both cannot close at the same time.

123. Direct Current Motor. DC motors are controlled using a circuit similar to the ones shown in Figure 16.83. This circuit will allow control of the path DC current takes through the motor armature, allowing for control of the direction in which the motor turns. The NPN and PNP transistors must be matched and the values for "Ra" and "Rb" must be chosen while considering the Hfe of the transistors along with the operating currents of the motors. "Ra" and "Rb" must be chosen to ensure that the current passing through the NPN and PNP transistors does not exceed their rated values when the motor is stalled.

Motor-speed control of a series DC motor is accomplished by use of a PWM waveform used for "Forward Current" and Reverse Current control signals in Figure 16.83. The four diodes shown in Figure 16.83 are "Kickback suppressors" and will minimize the inductive voltage produced when the motors are turned on and off.
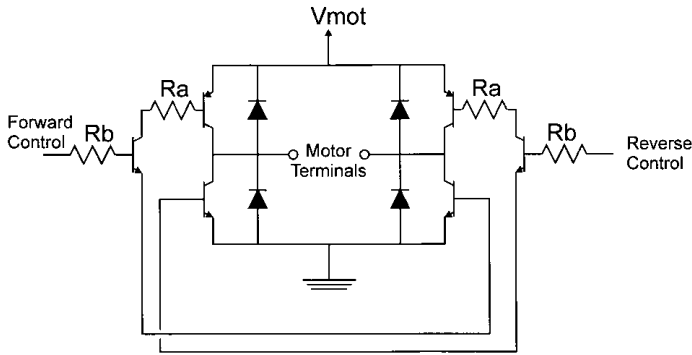
FIGURE 16.83 "Best" H-Bridge controller for small motors.

124. Generator-Voltage or Ward-Leonard Control. For certain classes of mine hoists, steel-mill appliances, ore bridges, electric shovels, etc., it is necessary to have high starting torque, accurate control, and constant-speed running. For these classes of service, a dc motor with generator-voltage control is the best arrangement obtainable. This requires a separate generator for each motor. The generator may be driven by any type of prime mover, but is usually driven by an induction or synchronous motor.

Figure 16.84 shows the connections for a sample generator-voltage control. The motor and generator are connected together, sometimes with a switch and circuit breaker or contactor in the circuit, sometimes without. Both generator and motor are usually separately excited. In Fig. 16.84 the strength and direction of the generator field are controlled by a drum controller. With this arrangement, the direction of rotation of the motor and its speed depend on the generator field. The torque developed by the motor varies with the armature current, since the field is constant. At standstill a system. very low voltage equal to the *IR* drop will force a large current through the motor circuit and thus develop high torque.

With connections as given, if the load tends to overhaul and drive the motor above the normal speed, the motor becomes a generator and pumps power back into the generator; the generator automatically becomes a motor, driving its prime mover. The power is dissipated either in an increase in speed or by being put back into the power system. With a given voltage, the motor speed is practically constant regardless of load.
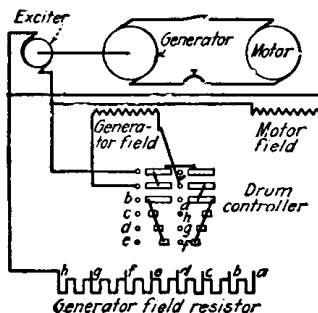
FIGURE 16.84 Generator-voltage control system.

This type of control lends itself to many refinements. During acceleration of the motor, the control may have from 3 to 15 or 20 points. It may be manually operated, by either drum controller or master switch with contactors having current or time relays; or the controller may be motor-operated. During slowdown, the controller may be turned off by mechanical cams or other means. By adding a resistance and one or two contactors across the motor armature, dynamic braking may be obtained when slowing down. A current relay may be connected with its coil in the generator-motor armature circuit and its trips in the generator-field circuit. At the beginning of slowdown the generator field is opened; the motor, driven by the inertia of the load, pumps back on the generator. As the current rises to the predetermined point, the relay trips close, putting field on the generator. This causes the pump-back current to fall. This cycle is repeated until the inertia of the system is dissipated and the motor is at rest.

The *Ilgner-Ward-Leonard system* is simply the addition of a flywheel to the motor-generator set. This is for use on equipments operating on a recurring cycle where the input must be kept constant or nearly so. The generator must be driven by a slip-ring type induction motor with some arrangement to obtain variable speed, either grid resistance cut in by a relay in one or two steps, or a liquid slip regulator controlled by a torque motor in the main induction-motor line. An increase in the line current causes the torque motor to rotate, cutting resistance into the secondary of the induction motor, causing the motor to slow down and the flywheel to give up its power to the generator. A decrease in line current allows the balancing weights to overcome the torque of the torque motor, causing the motor to rotate in the opposite direction, cutting resistance out of the induction-motor secondary, and allowing the motor to speed up and bring the flywheel up to normal speed.

CHAPTER 17

ELECTRONICS ENGINEERING*

The subject of electronics can be approached from the standpoint of the design of devices or the use of devices. For the practicing 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 normally have a series of four or five color bands printed on them as shown in Fig. 17.1. These bands are used to specify the resistance of the circuit using the formula:

$$\text{Resistance} = ((\text{Band1} \times 100) + (\text{Band2} \times 10) + (\text{Band3} \times 1)) \times 10 \ ** \ \text{Band4}$$

(17.1)

Table 17.1 lists the values for each band color and what they mean in the resistance specification formula.

Capacitors can come in a variety of different packages, as shown in Fig. 17.2. Disk capacitors may be stamped with a number like ''103'' indicating the value.

The value of the capacitor is determined using the formula:

$$\text{Capacitance} = (\text{First Two Digits}) \times 10 \ ** \ (\text{Third Digit}) \qquad (17.2)$$

For example, 330 pF would be stamped with 331, 3,300 pF (3.3 nF) would be stamped with 332 and so on. If the third digit is not present, then for the purposes of the formula above, assume that it is zero.

A *rectifier,* or *diode,* is an electronic device which offers unequal resistance to forward and reverse current flow. Figure 17.3 shows the schematic symbol for a diode. The arrow beside the diode shows the direction of current flow, and the bar at the point of the arrow corresponds to the band around the physical diode indicating the direction of current flow through the component. Current flow is taken to be the flow of positive charges, i.e., the arrow is counter to electron flow. Figure

---

*This section is based on material from *Marks' Standard Handbook for Mechanical Engineers,* 9th ed., by E. A. Avallone and T. Baumeister III, Copyright 1987. Used by Permission of McGraw-Hill, Inc. Updated and extended by Myke Predko in 2002.

17.1
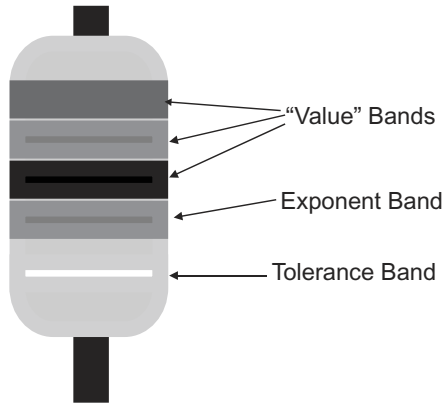
FIGURE 17.1   Resistor appearance and marking.

TABLE 17.1   Resistor Value Color Bands

Number	Color	Band1	Band2	Band3	Exponent Band	Tolerance Band
0	Black	N/A	0	0	$10 ** 0$	N/A
1	Brown	1	1	1	$10 ** 1$	1% Tolerance
2	Red	2	2	2	$10 ** 2$	2% Tolerance
3	Orange	3	3	3	$10 ** 3$	N/A
4	Yellow	4	4	4	$10 ** 4$	N/A
5	Green	5	5	5	$10 ** 5$	0.5% Tolerance
6	Blue	6	6	6	$10 ** 6$	0.25% Tolerance
7	Violet	7	7	7	$10 ** 7$	0.1% Tolerance
8	Gray	8	8	8	$10 ** 8$	0.05% Tolerance
9	White	9	9	9	$10 ** 9$	N/A
N/A	Gold	N/A	N/A	N/A	$10 ** -1$	5% Tolerance
N/A	Silver	N/A	N/A	N/A	$10 ** -2$	10% Tolerance
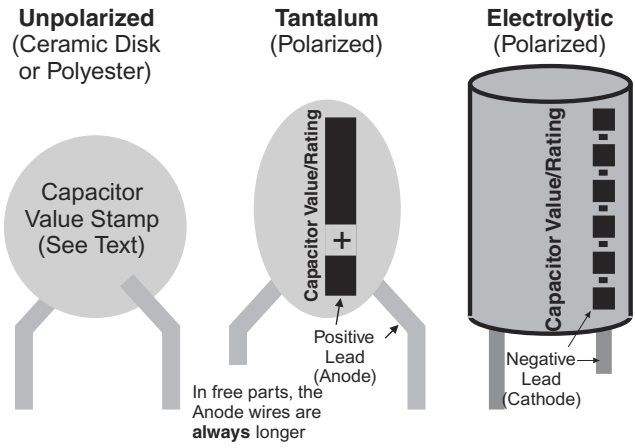
FIGURE 17.2   Capacitor appearance and markings.
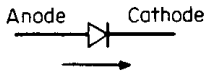
FIGURE 17.3   Diode schematic symbol.
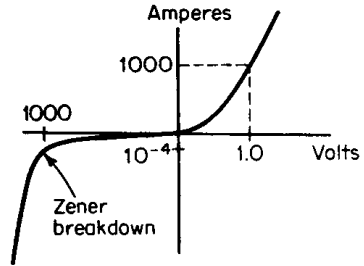
FIGURE 17.4   Diode forward-reverse characteristic.

17.4 shows typical forward and reverse voltampere 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 1000 A. Reverse voltage breakdown varies from 50 V to more than 2500 V. At rated forward current, the forward voltage drop 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. 17.5 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.

A very popular method of indicating circuit status or driving out information is by the use of "Light Emitting Diodes." The most common LED is the 5mm pin-through-hole package shown in Fig. 17.6. These devices behave like standard diodes but output light (red, green, yellow, blue and white are available) when current is passing through them.

Single LEDs are often designed to light with 5–20 mA passing through them. When calculating series current-limiting resistors to be used with LEDs, assume
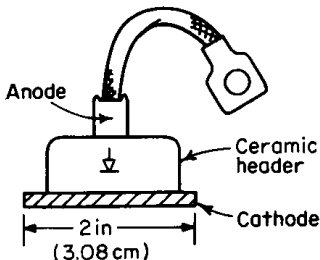
FIGURE 17.5   Physical diode package.
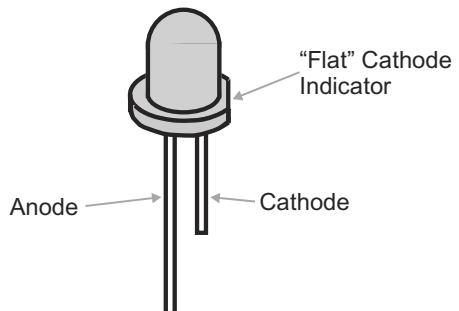
FIGURE 17.6   Light emitting diode appearance.

that the voltage drop across the LED is 2.0 volts—this is different from a silicon diode, which has a nominal voltage drop of 0.7 volts.

Multiple LED display packages are available for displaying non-binary data. The most common multiple LED display is the ''7 Segment'' display (shown in Fig. 17.7) and is used in a variety of different applications. The anodes or cathodes of all the LEDs in the display are tied together to simplify application wiring. Multiple displays can be wired together using a single driver, sequencing through each digit by controlling the operation of the current display common cathode or anode as shown in Fig. 17.8.

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 semicon-
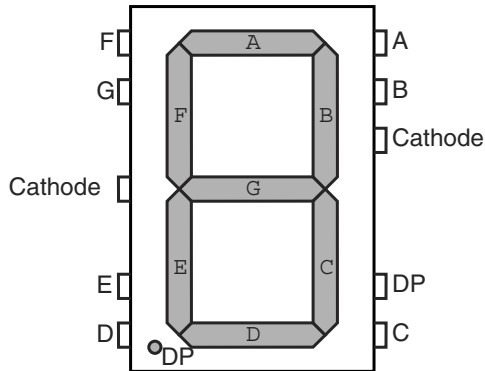
FIGURE 17.7   7 segment ''common cathode'' LED display with pinout.
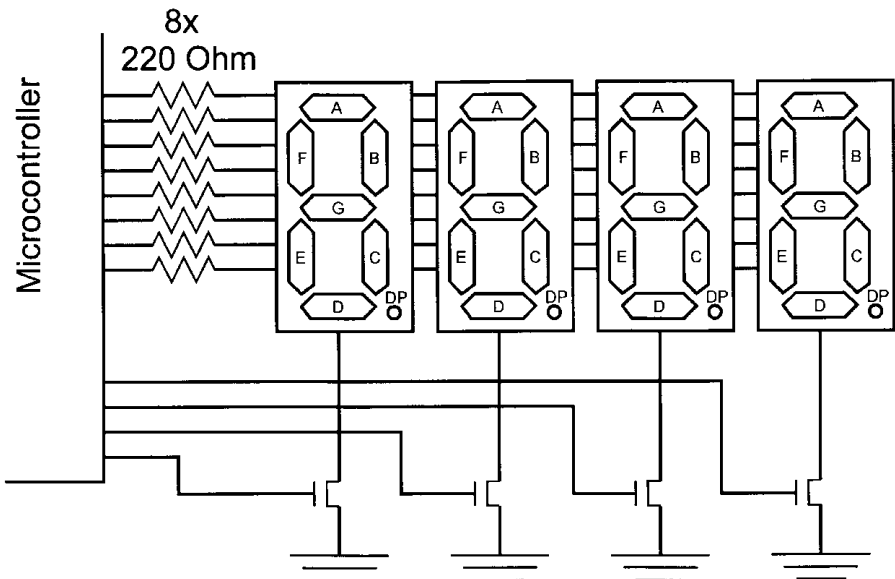
FIGURE 17.8   Wiring four 7 segment LED displays.

ductor 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. 17.9. 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.

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
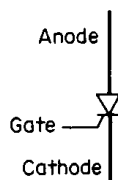
FIGURE 17.9 Thyristor schematic symbol.

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 17.2 lists typical thyristor characteristics.

A *triac* is a bilateral SCR. It blocks current in either direction 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 bipolar transistor is shown in Fig. 17.10. There are two types of bipolar transistors, *p-n-p*

TABLE 17.2 Typical Thyristor Characteristics

| Voltage | Current, A | | $I_2^2 t$, A · S | 1-Cycle surge, A | di/dt, A/s | dv/dt, V/s | Turn-off time, s |
	rms	avg					
400	35	20	165	180	100	200	10
1200	35	20	75	150	100	200	10
400	110	70	4,000	1000	100	200	40
1200	110	70	4,000	1000	100	200	40
400	235	160	32,000	3500	100	200	80
1200	235	160	32,000	3500	75	200	80
400	470	300	120,000	5500	50	100	150
1200	470	300	120,000	5500	50	100	150
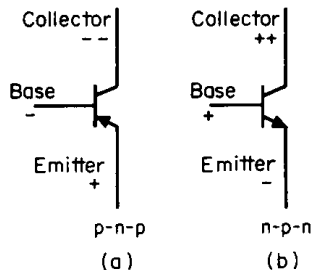
FIGURE 17.10 Transistor schematic symbol.

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 bipolar 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 bipolar transistor occurs because of two facts: First, 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. Second, 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 17.3 shows basic ratings for some typical transistors and Fig. 17.11 shows the pinout for 2N3904/2N3906 transistors in TO-92 packages. There is a great profusion of transistor types so that the choice of type depends on availability and cost as well as operating characteristics.

The gain of a bipolar 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 in Fig. 17.12. 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

TABLE 17.3 Typical Transistor Characteristics

Part 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		†
BC547	*n-p-n*	80	500 mW	200 mA	290
BC557	*p-n-p*	45	500 mW	200 mA	290
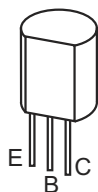
FIGURE 17.11 TO-92 bipolar transistor pinout.
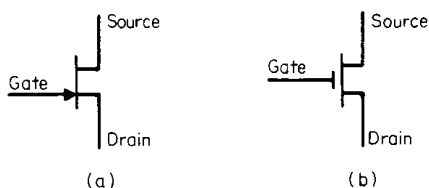
FIGURE 17.12 Field effect transistor: (*a*) bipolar junction type (JFET); (*b*) metal oxide semiconductor type (MOSFET).

designed for high-power applications. The two basic constructions used for FETs are *bipolar junctions* and *metal oxide semiconductors.* The schematic symbols for each of these are shown Fig. 17.12*a* and *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 that is 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. 17.13. 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.

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 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
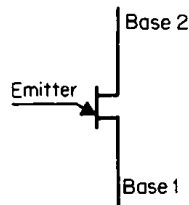
FIGURE 17.13   Unijunction.

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 electric energy when a fault occurs in a circuit.

DISCRETE-COMPONENT CIRCUITS

Several common rectifier circuits are shown in Fig. 17.14. 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. 17.14. 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:

$$I = \frac{E_{dc}PCT}{200\pi fNL} \tag{17.3}$$

where I is the rms ripple current, E_{dc} is the dc load voltage, PCT is percent ripple from Fig. 17.14, f is line frequency, N is number of cycles of ripple frequency per cycle of line frequency, and L is equivalent series inductance in load. Equation (17.1) 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. 17.15. 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. 17.14. 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

Type	Circuit	Output voltage waveform	E_{dc} (avg)	Ripple fundamental frequency	% ripple	Peak inverse voltage
Half-wave 1φ		e_{load} E_M e_{ac}	$0.318\,E_M$ $0.45\,E_{ac}$	F	121	$3.14\,E_{dc}$
Full-wave 1φ		e_{load} E_M e_{ac}	$0.636\,E_M$ $0.9\,E_{ac}$	2F	48	$3.14\,E_{dc}$
Bridge 1φ		e_{load} E_M e_{ac}	$0.636\,E_M$ $0.9\,E_{ac}$	2F	48	$1.57\,E_{dc}$
Half-wave 3φ		e_{load} E_M e_{ac}	$0.827\,E_M$ $1.17\,E_{ac}$	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}

FIGURE 17.14 Comparison of rectifier circuits.

be possible to build up current in an inductive load. The gate-control circuitry is not shown in Fig. 17.15 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. 17.15 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
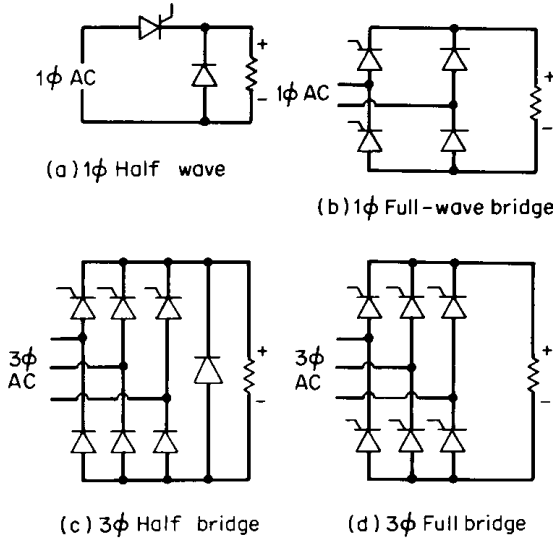
FIGURE 17.15 Basic thyristor circuits.

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. 17.15 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. 17.16. 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.

Transistor amplifiers can take many different forms. The circuits described here illustrate basic principles. A basic *single-stage amplifier* is shown in Fig. 17.17.
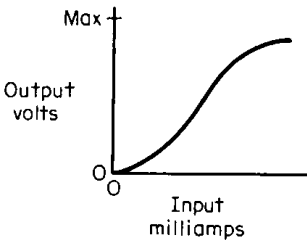
FIGURE 17.16 Thyristor control characteristic.
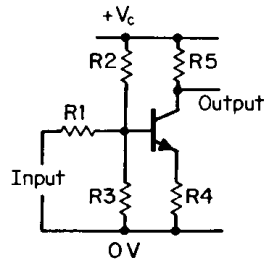
FIGURE 17.17 Single-stage amplifier.

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 17.18 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. 17.17. 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 17.19 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 2500. This high level of gain is not very stable if the ambient temperature changes, but in many cases this drift is tolerable.

Figure 17.20 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 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 amplification 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.

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
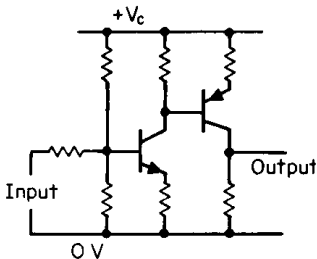
FIGURE 17.18   Two-stage amplifier.
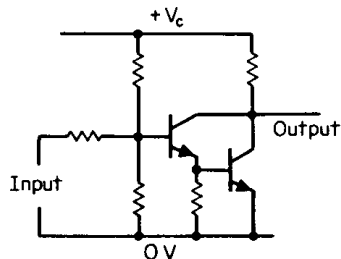
FIGURE 17.19   Darlington connection.
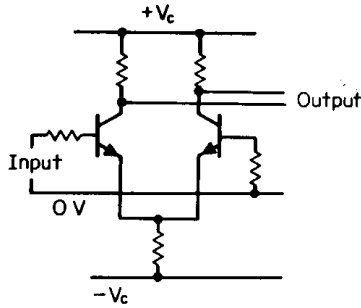
FIGURE 17.20   Differential amplifier.

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

$$f = \tfrac{1}{2}\pi\sqrt{CL} \tag{17.4}$$

where f is frequency (Hz), C is capacitance (F), and L is inductance (H). A 1-MHz oscillator might typically be designed with 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. 17.21 differs from the *Hartley oscillator* shown in Fig. 17.22 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* shown in Fig. 17.23 has much greater frequency stability than the circuits in Figs. 17.22 and 17.23. 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.

As the desired operating frequency becomes 500 MHz and greater, *resonant cavities* are used as tank circuits instead of discrete capacitors and inductors. A
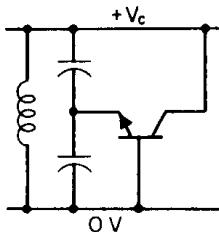
FIGURE 17.21   Colpitts oscillator.
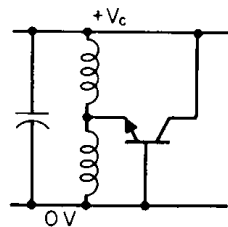
FIGURE 17.22   Hartley oscillator.
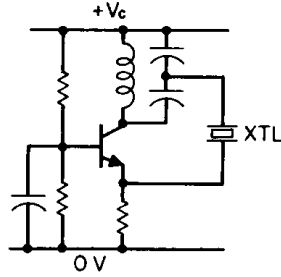
FIGURE 17.23 Crystal-controlled oscillator.

rough guide to the relationship between frequency and resonant-cavity size is the wavelength of the frequency

$$\lambda = 300 \times 10^6/f \qquad (17.5)$$

where λ is wavelength (m), 300×10^6 is the speed of light (m/s), and f is frequency (Hz). The resonant cavities will be smaller than indicated by Eq. (17.5) 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.

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.

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 sim-

ple amplitude modulator is shown in Fig. 17.24. 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. 17.25. The carrier must be changed in frequency in 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. 17.26. 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 output contains only the information signal. An FM *demodulator* is shown in Fig. 17.27. 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. 17.26.

The waveform of the basic *electronic timing circuit* is shown in Fig. 17.28 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 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
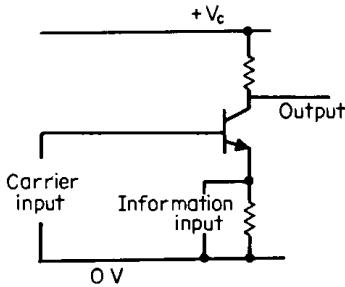
FIGURE 17.24   AM modulator.
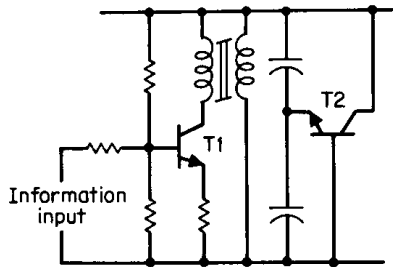
FIGURE 17.25   FM modulator.
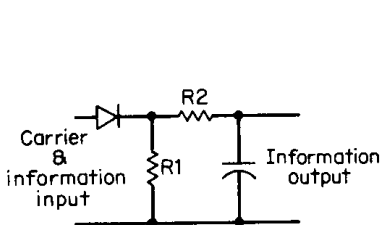
FIGURE 17.26   AM demodulator.
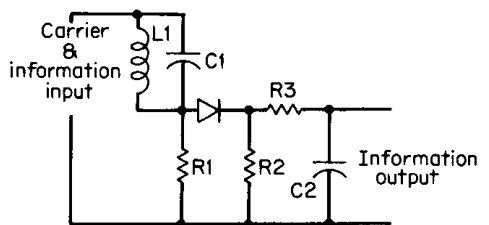
FIGURE 17.27   FM discriminator.
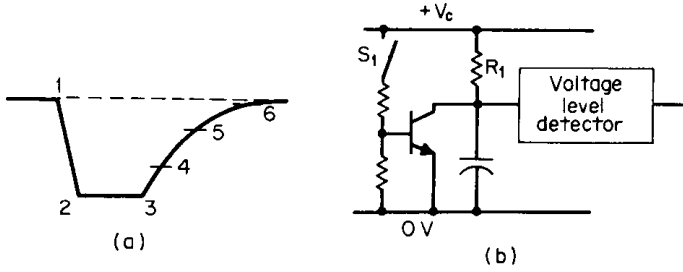
FIGURE 17.28 Basic timing circuit.

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 which 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

$$t = RC \qquad (17.6)$$

where t is time (s), R is resistance (Ω), and C is 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.

   An improved timing circuit is shown in Fig. 17.29. 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 17.4 lists some of the more common physical packages for discrete component and integrated semiconductor devices. Although discrete components are still
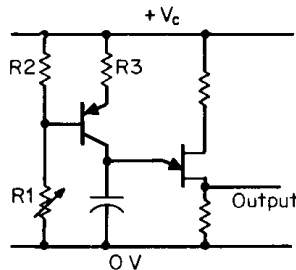
FIGURE 17.29 Improved timing circuit.

TABLE 17.4 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. 17.30)
Flat pack	
Chip carrier (50-mil centers)	

used for electronic design, *integrated circuits* (*ICs*) are predominant in almost all types of electronic equipment. Dimensions of common dual inline pin (DIP) integrated-circuit devices are shown in Fig. 17.30. An IC costs far less than circuits made 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 on 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

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
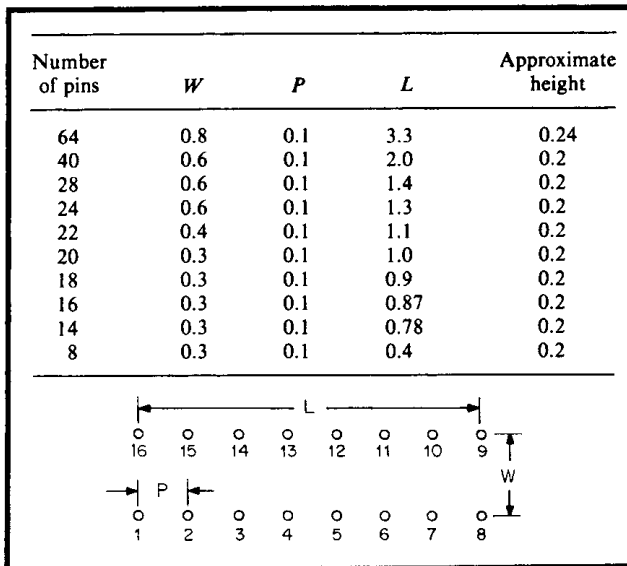
FIGURE 17.30 Approximate physical dimensions of dual in-line pin (DIP) integrated circuits. All dimensions are in inches. Dual in-line packages are made in three different constructions: molded plastic, cerdip, and ceramic.

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 0.13 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 17.5 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 17.31 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. 17.31.

Table 17.6 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.

TABLE 17.5   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
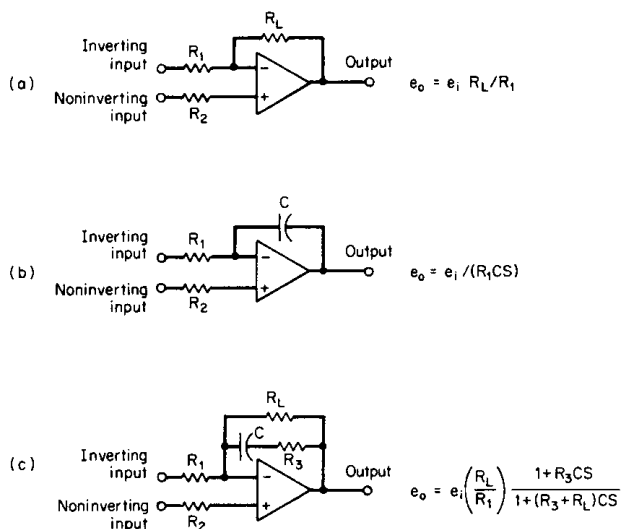
FIGURE 17.31 Operational amplifier circuits.

TABLE 17.6 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

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.

Table 17.7 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 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.

TABLE 17.7 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 17.8 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 dcsign of digital-switched telephone circuits.

A very useful chip that is often overlooked by modern application developers is the NE 555 timer (pinout and internal circuitry is shown in Fig. 17.32). This versatile chip comes in a variety of different packages, formats and technologies. The basic "555" chip uses bipolar transistors and is capable of utilizing anywhere from 5 to 15 volts "Vcc". CMOS versions can run from 2.5 to 15 volts and do not cause the large ground "bounce" that is characteristic of the bipolar technology versions. The 556 chip consists of two 555 timers, and there are other chips integrating more than two 555 timers on a single chip.

The basic function of the 555 timer is the oscillator shown in Fig. 17.33. By varying the two resistor values, a PWM output, with varying duty cycles can be

TABLE 17.8   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)
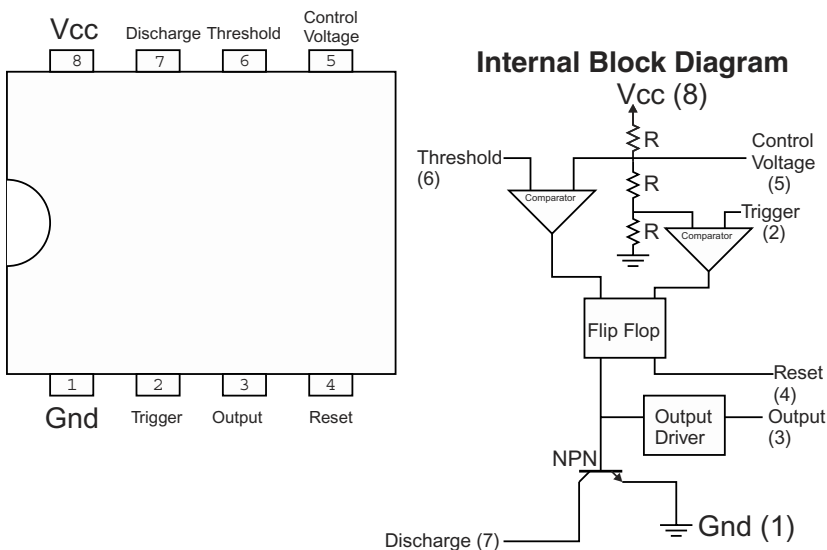
FIGURE 17.32   NE 555 timer chip.
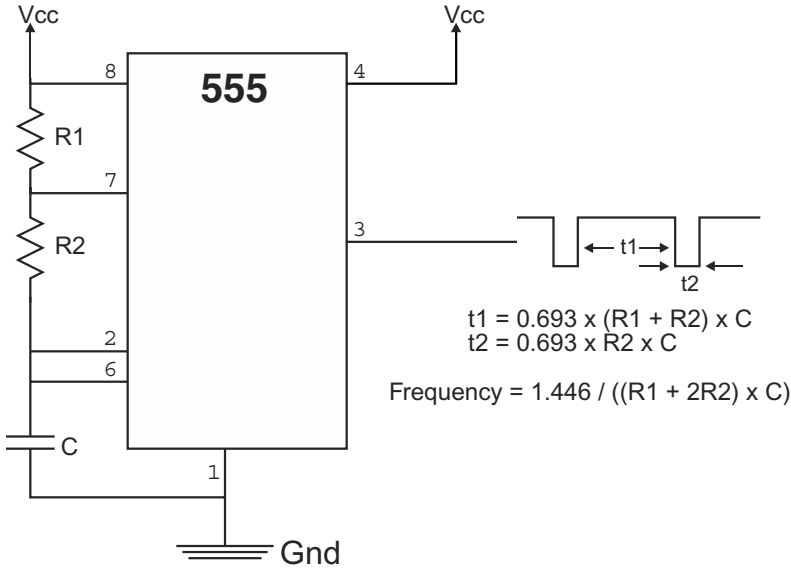
FIGURE 17.33   555 oscillator circuit.

output. A popular use for the 555 timer chip is to produce a timed pulse from a negative pulse input as shown in Fig. 17.34—the 555 is often used to debounced momentary button input as shown in Fig. 17.35. Other applications that can be implemented using the 555 timer chip include varying tone sirens and missing pulse detectors.

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 opera-
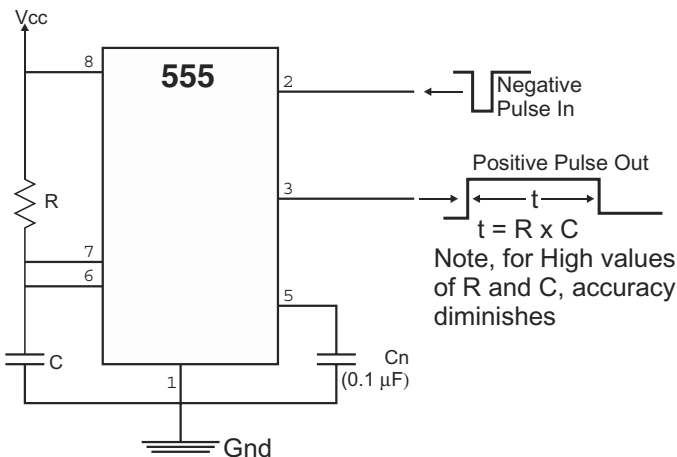
FIGURE 17.34   555 monostable circuit.
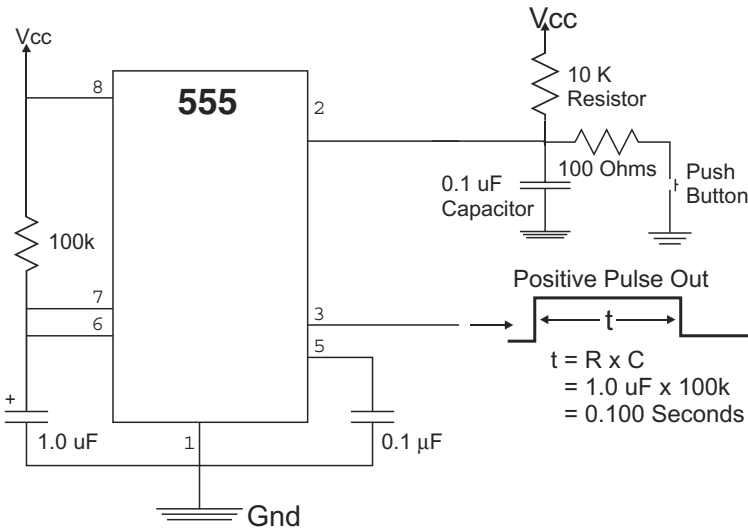
FIGURE 17.35 555 button debounce circuit.

tional 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-transistor-logic* (TTL) *gates* are listed in Table 17.9, which references diagrams with the pinouts of the different chips.

The TTL part number is in the format "74xx##(#)" where the "xx" is the logic technology specifier. There are several different kinds of TTL technology for each of the chips listed in Table 17.9. The input parameters of these technologies are listed in Table 17.10, and the output parameters are listed in Table 17.11. Note that for "true" TTL, the "threshold" voltages are based on the amount of current drawn from the chip inputs.

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 on 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 17.12 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-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 17.12 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.

TABLE 17.9 Digital Integrated-Circuit Devices

Type (74xx##)	No circuits per device	No inputs per device	Function	Pinout
74xx00	4	2	NAND	Fig. 17.36
74xx02	4	2	NOR	Fig. 17.37
74xx04	6	1	NOT	Fig. 17.38
74xx06	6	1	Buffer	Fig. 17.39
74xx08	4	2	AND	Fig. 17.40
74xx32	4	2	OR	Fig. 17.41
74xx74	2	4	D-Flip Flop	Fig. 17.42
74xx83	1	9	Adder	Fig. 17.43
74xx86	4	2	XOR	Fig. 17.44
74xx125	4	2	Tri-Sate	Fig. 17.45
74xx138	1	6	3-8 Decoder	Fig. 17.46
74xx139	2	3	2-4 Decoder	Fig. 17.47
74xx174	6	1+ Common Clock and Common Clear	D-Flip Flop	Fig. 17.48
74xx244	8	1+ Common Enable	Tri-State	Fig. 17.49
74xx245	8	1+ Common Enable	Bi-directional Tri-State	Fig. 17.50
74xx373	8	1+ Common Clock and Common Output Enable	D-Flip Flop	Fig. 17.51
74xx374	8	1+ Common Clock and Common Output Enable	D-Flip Flop	Fig. 17.52
74xx573	8	1+ Common Clock and Common Output Enable	D-Flip Flop	Fig. 17.53
74xx574	8	1+ Common Clock and Common Output Enable	D-Flip Flop	Fig. 17.54
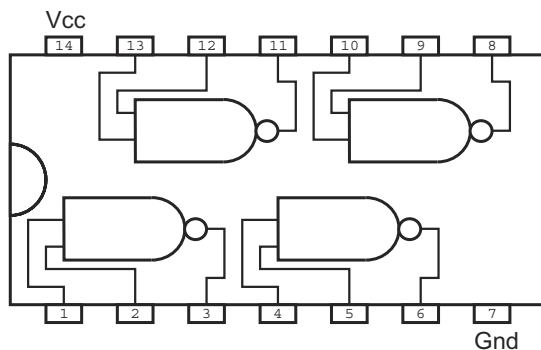
FIGURE 17.36 Quad two input "NAND" gate TTL chip (74xx00).
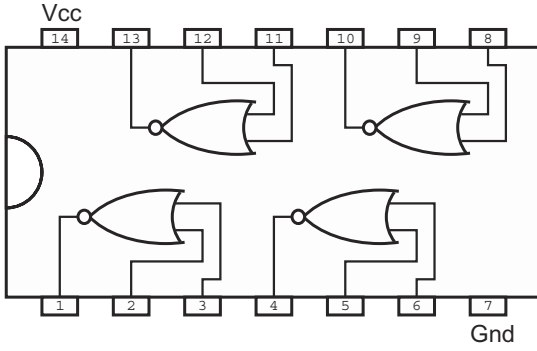
FIGURE 17.37 Quad two input "NOR" gate TTL chip (74xx02).
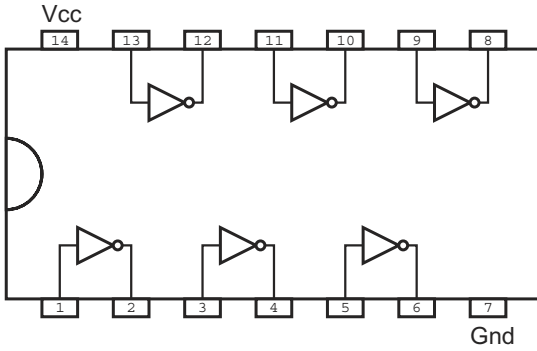
FIGURE 17.38 Hex inverters with totem pole outputs (74xx04).
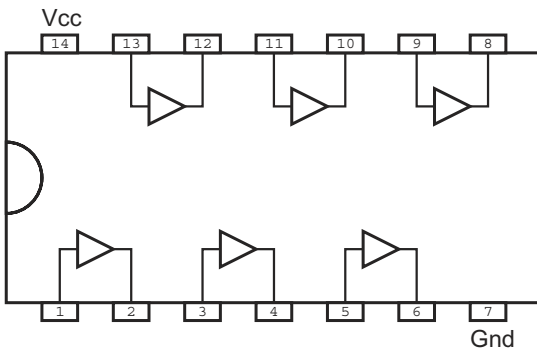
FIGURE 17.39 Hex buffers with totem pole outputs (74xx06).
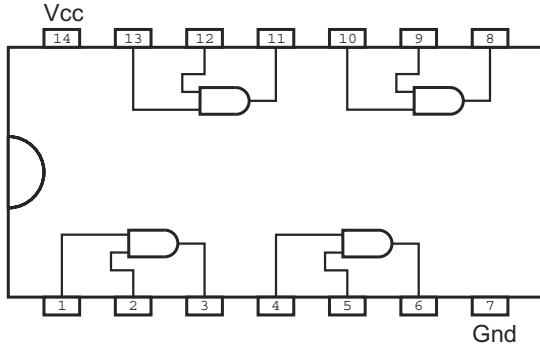
FIGURE 17.40 Quad two input ''AND'' gate TTL chip (74xx08).
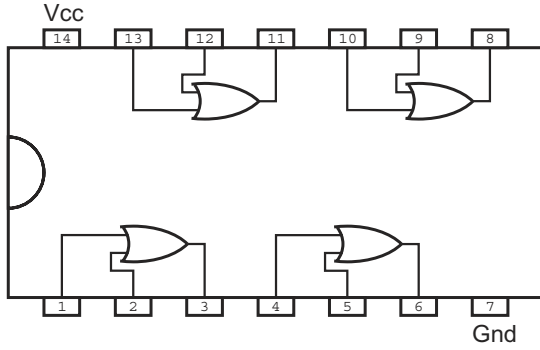
FIGURE 17.41 Quad two input ''OR'' gate TTL chip (74xx32).
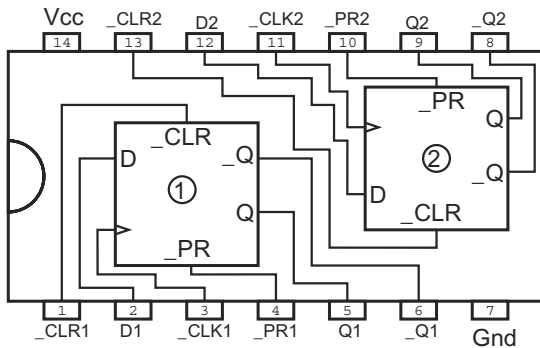
FIGURE 17.42 Dual ''D-type'' flip flops (74xx74).
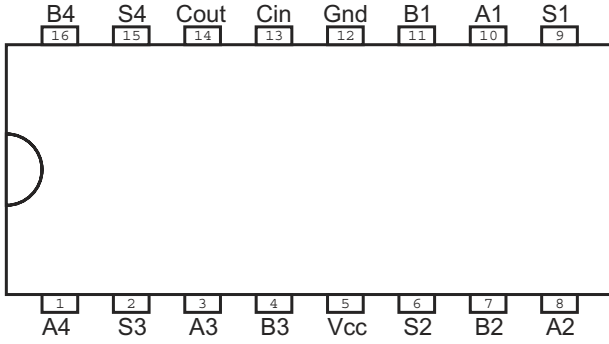
7483

FIGURE 17.43 Four bit full adder with fast carry (74xx83).
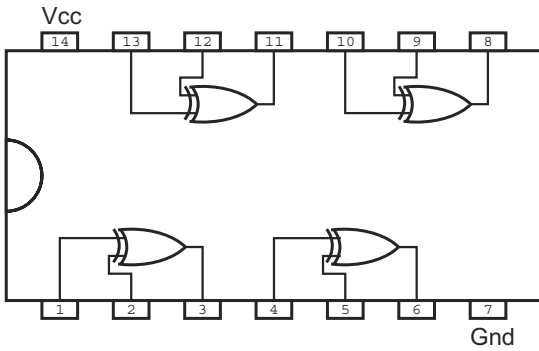
FIGURE 17.44 Quad two input "XOR" gate TTL chip (74xx86).
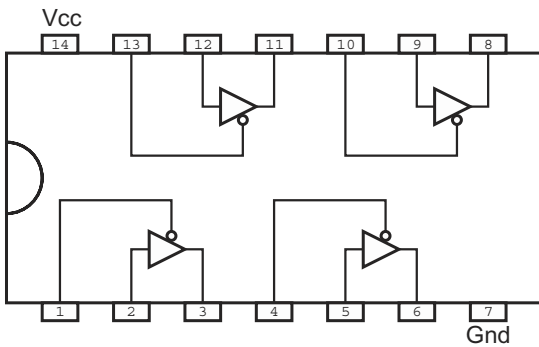
FIGURE 17.45 Quad tri-state buffers with negative active enable (74xx125).

Vcc Y0 Y1 Y2 Y3 Y4 Y5 Y6
16 15 14 13 12 11 10 9

A B C _G2A _G2B G1 Y7 Gnd
1 2 3 4 5 6 7 8
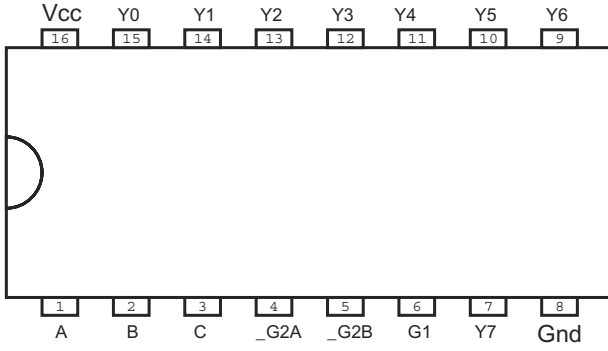
Y# is Active Low When G1/G2 are Set Correctly
= (C x 4) + (B x 2) + (A x 1)

FIGURE 17.46 Three to eight decoder (74xx138).

Vcc _G2 A2 B2 Y20 Y21 Y22 Y23
16 15 14 13 12 11 10 9

_G1 A1 B1 Y10 Y11 Y12 Y13 Gnd
1 2 3 4 5 6 7 8
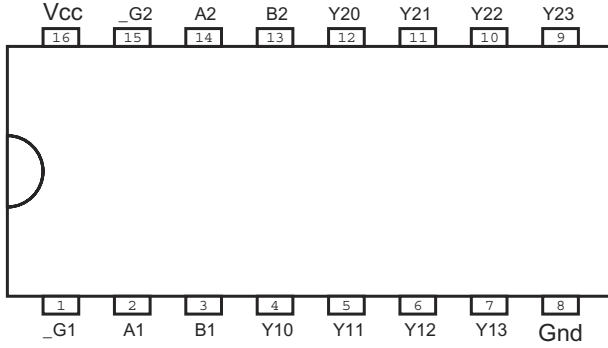
Y# is Active Low When _G# Pulled Low
= (B x 2) + (A x 1)

FIGURE 17.47 Dual two to four decoder (74xx139).

Vcc Q6 D6 D5 Q5 D4 Q4 CLK
16 15 14 13 12 11 10 9

_CLR Q1 D1 D2 Q2 3D Q3 Gnd
1 2 3 4 5 6 7 8
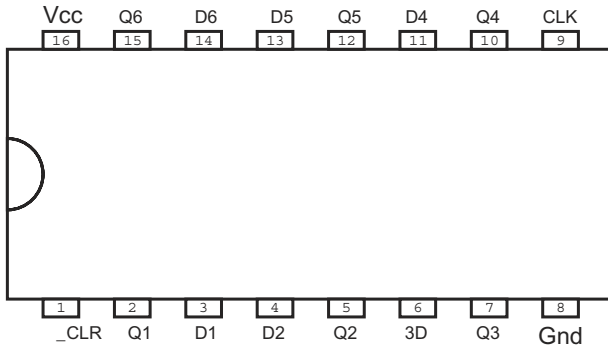
Data Latched on Rising Edge of ''CLK''

FIGURE 17.48 Hex ''D-type'' flip flops (74xx174).
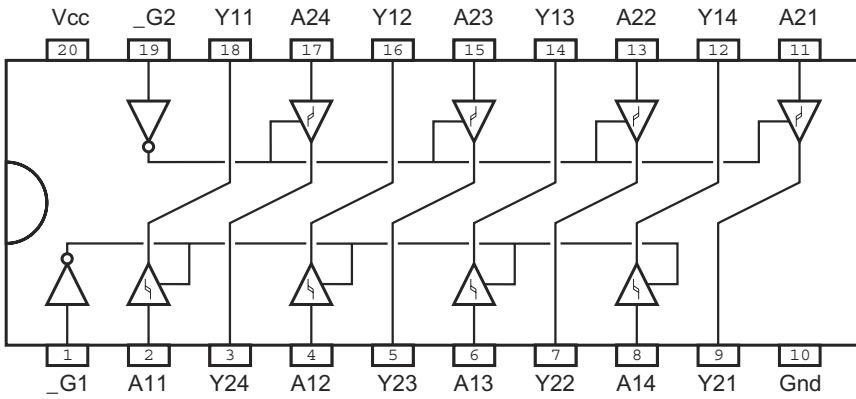
FIGURE 17.49   Eight bit tri-state driver (74xx244).

DIR = 1, Data From A to B
DIR = 0, Data From B to A
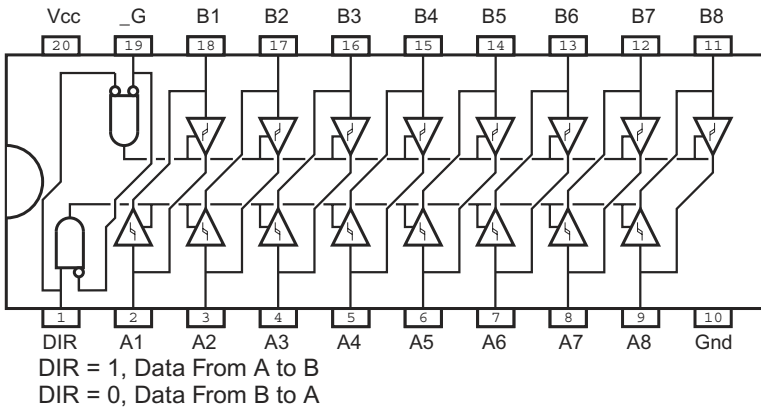
FIGURE 17.50   Bi-directional eight bit tri-state driver (74xx245).
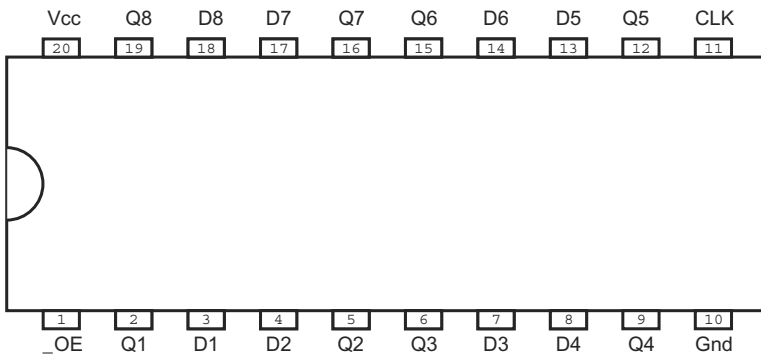
Latches are Transparent when "CLK" is High.
Data Latched in on negative Edge of "CLK"

FIGURE 17.51   Eight bit latch with tri-state output driver (74xx373).

Vcc Q8 D8 D7 Q7 Q6 D6 D5 Q5 CLK
20 19 18 17 16 15 14 13 12 11

1 2 3 4 5 6 7 8 9 10
_OE Q1 D1 D2 Q2 Q3 D3 D4 Q4 Gnd
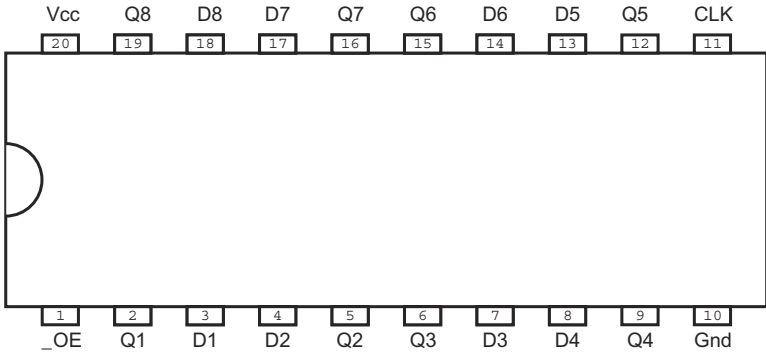
Output Changes After Latch Operation.
Data Latched in on Positive Edge of "CLK"

FIGURE 17.52 Eight bit latch with tri-state output driver (74xx374).

Vcc Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 CLK
20 19 18 17 16 15 14 13 12 11

1 2 3 4 5 6 7 8 9 10
_OE D1 D2 D3 D4 D5 D6 D7 D8 Gnd
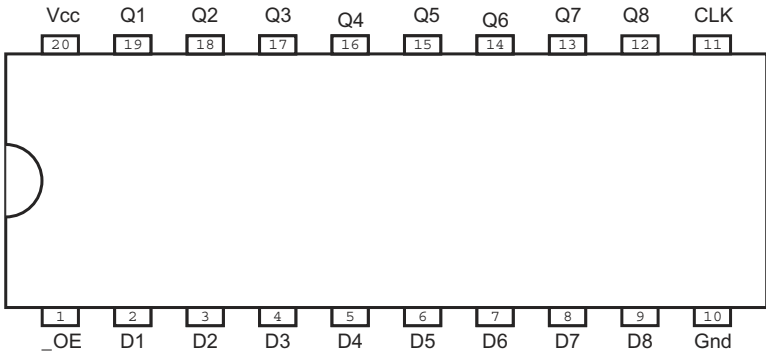
Latches are Transparent when "CLK" is High.
Data Latched in on negative Edge of "CLK"

FIGURE 17.53 Eight-bit latch with tri-state output driver (74xx573).

Vcc Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 CLK
20 19 18 17 16 15 14 13 12 11

1 2 3 4 5 6 7 8 9 10
_OE D1 D2 D3 D4 D5 D6 D7 D8 Gnd
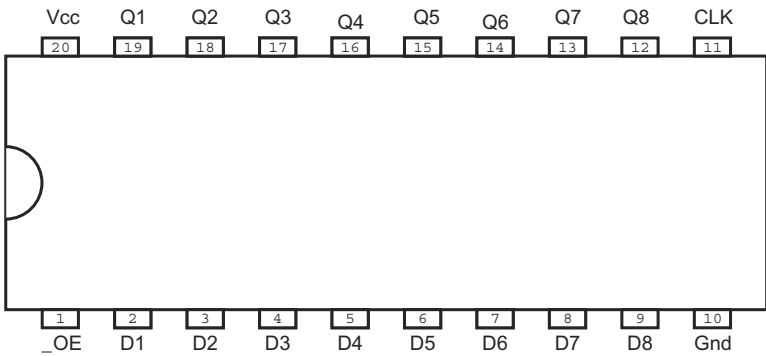
Output Changes After Latch Operation.
Data Latched in on Positive Edge of "CLK"

FIGURE 17.54 Eight bit latch with tri-state output driver (74xx574).

TABLE 17.10 5 Volt Logic Technologies Input Parameters

Type	Threshold	"0" Output	"1" Output
TTL	1.4 volts	0.3 volts	3.3 volts
HC	2.4 volts	0.1 volts	4.9 volts
HCT	1.4 volts	0.1 volts	4.9 volts
CMOS	2.5 volts	0.1 volts	4.9 volts

TABLE 17.11 Logic Technologies Output Parameters

Family	Transition Time	Maximum Current Sink
Straight TTL (No Letter)	8 nsec	12 mA
"L" TTL	15 nsec	15 mA
"LS" TTL	10 nsec	8 mA
"S" TTL	5 nsec	40 mA
"AS" TTL	2 nsec	20 mA
"ALS" TTL	4 nsec	8 mA
"F" TTL	3.5 nsec	20 mA
"C" CMOS	50 nsec	1.3 mA
"HC"/HCT CMOS	9 nsec	8 mA
"4000" CMOS	30 nsec	0.5 mA
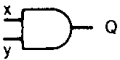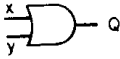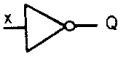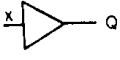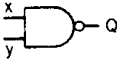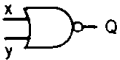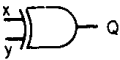
TABLE 17.12 Flip-Flop Sequential Devices

Name		Graphic symbol	Algebraic function	Truth table
JK flip-flop	Clock		$Q(t+1) = JQ'(t) + K'Q(t)$	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	Clock		$Q(t+1) = TQ'(t) + T'Q(t)$	T $Q(t)$ | $Q(t+1)$ 0 0 | 0 0 1 | 1 1 0 | 1 1 1 | 0
D flip-flop	Clock		$Q(t+1) = D$	D $Q(t)$ | $Q(t+1)$ 0 0 | 0 0 1 | 0 1 0 | 1 1 1 | 1

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 17.9.

Various types of gates are shown in Table 17.13. 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

TABLE 17.13 Combinational Gate Logic

Name	Graphic symbol	Algebraic function	Truth table
AND		$Q = xy$	$x\ y\ \vert\ Q$ $0\ 0\ \vert\ 0$ $0\ 1\ \vert\ 0$ $1\ 0\ \vert\ 0$ $1\ 1\ \vert\ 1$
OR		$Q = x + y$	$x\ y\ \vert\ Q$ $0\ 0\ \vert\ 0$ $0\ 1\ \vert\ 1$ $1\ 0\ \vert\ 1$ $1\ 1\ \vert\ 1$
Inverter		$Q = x'$	$x\ \vert\ Q$ $0\ \vert\ 1$ $1\ \vert\ 0$
Buffer		$Q = x$	$x\ \vert\ Q$ $0\ \vert\ 0$ $1\ \vert\ 1$
NAND		$Q = (xy)'$	$x\ y\ \vert\ Q$ $0\ 0\ \vert\ 1$ $0\ 1\ \vert\ 1$ $1\ 0\ \vert\ 1$ $1\ 1\ \vert\ 0$
NOR		$Q = (x + y)'$	$x\ y\ \vert\ Q$ $0\ 0\ \vert\ 1$ $0\ 1\ \vert\ 0$ $1\ 0\ \vert\ 0$ $1\ 1\ \vert\ 0$
EXCLUSIVE-OR		$Q = xy' + x'y$ $= x + y$	$x\ y\ \vert\ Q$ $0\ 0\ \vert\ 0$ $0\ 1\ \vert\ 1$ $1\ 0\ \vert\ 1$ $1\ 1\ \vert\ 0$

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 17.13 have only two inputs. Standard TTL chips are available with gates that can 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: ·, 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 17.14.

To facilitate the analysis of digital circuits and to aid in the application of the rules given in Table 17.14, *Karnaugh maps* are used. Typical two-variable and four-variable Karnaugh maps are shown in Fig. 17.55 along with the algebraic expressions represented by each map.

TABLE 17.14   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:

$(X + Y)' = X'{\cdot}Y'$	$(X{\cdot}Y)' = X' + Y'$
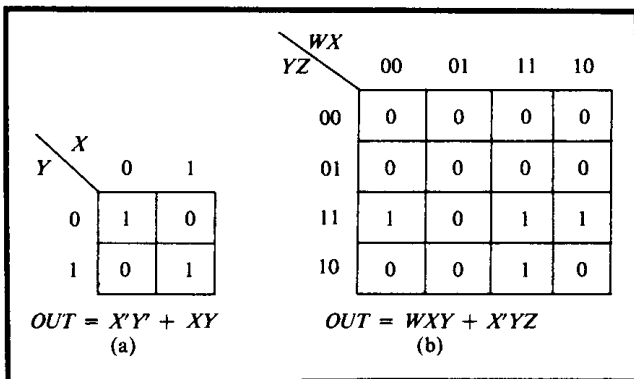
FIGURE 17.55   Karnaugh maps of typical logic functions.

TTL chips have had a number of part numbers given the 74xx##(#) designation which really do not seem to fit in the family of logic, flip flops and other digital logic circuits. One example of this is the 74C922 switch matrix keyboard controller (Fig. 17.56). This chip allows for the easy addition of a four-by-four switch matrix keypad to a digital electronic application as shown in (Fig. 17.57). These parts can both simplify an application as well as allow for "straight" logic chips to be used for applications that may have required the use of a microcontroller in the past.

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 17.15 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
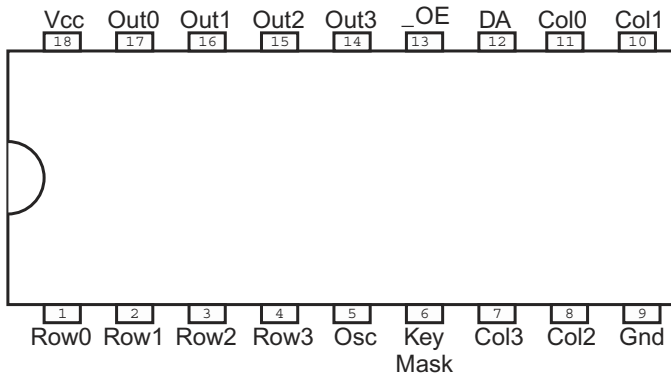
FIGURE 17.56   16 key switch matrix keypad controller (74C922).
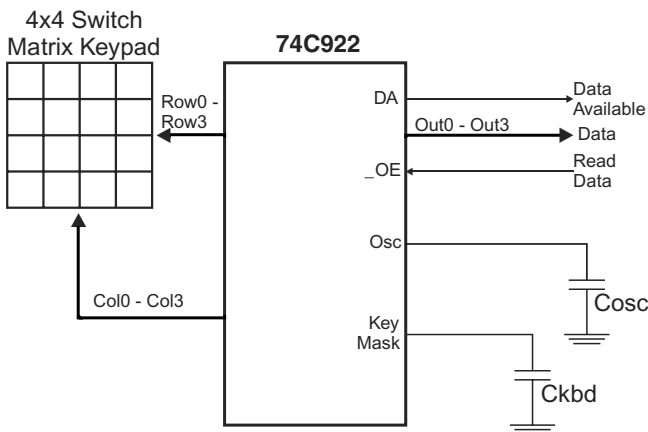
FIGURE 17.57   Connecting the 74C922 keypad controller.

TABLE 17.15 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

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 1000 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 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 32-bit registers. New designs use 16-bit microprocessors. For scientific computing or high-resolution graphics, 64-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 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 memories is lost when power is removed. For this reason, IC memory is 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. PROMs have not been used in electronic systems since the early 1980s because of the availability of erasable and reprogrammable devices as well as their poor reliability. Modern erasable PROMs (known as *EPROMs*) or *Flash* can be used for different applications, and their costs are competitive with ROMs in many products.

In addition to the CPU and memory ICs, other IC devices are required for computer support. These are listed in Table 17.16. 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.

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

TABLE 17.16 Digital Computer Integrated-Circuit Devices

Microprocessors	Network interface
Static random-access memory	Video device controller
Tristate buffer	Hard drive controller
Programmable timer	CD-ROM/DVD controller
Analog-to-digital converter	Programmable read-only memory
Floppy disk controller	Dynamic random-access memory
Universal asynchronous transmitter-receiver (UART)	Tristate receiver
	Parallel interface adapter
	Digital-to-analog converter
(Synchronous) dynamic random-access memory	IEEE 488 bus interface

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 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.

COMPUTER PROGRAMMING

Numbering Systems   Normally, numbers are represented in "Base" or "Radix" 10 with each digit multiplied by the appropriate power of 10. For example, "123" is one hundreds (10 to the power 2), two tens (10 to the power 1) and three ones (10 to the power 0) and could be written as:

$$123 = (1 * 100) + (2 * 10) + (3 * 1)$$
$$= (1 * (10**2)) + (2 * (10**1)) + (3 * (10**0)) \qquad (17.7)$$

In a computer, numbers are represented by a series of bits, where a bit (or "digit") can have the value "0" or "1". Binary values are represented as a series of ones or zeros with each digit representing a power of two. Using this system, B′101100′ can be converted to decimal using the binary exponent of each digit.

$$B'101100' = (1 * (2**5)) + (0 * (2**4)) + (1 * (2**3)) + (1 * (2**2))$$
$$+ (0 * (2**1)) + (0 * (2**1)) \text{ (Decimal)}$$
$$= 32 + 8 + 4 \text{ (Decimal)}$$
$$= 44 \text{ (Decimal)} \qquad (17.8)$$

For convenience, binary numbers are normally written as "hexadecimal" digits instead of individual bits. Hexadecimal digits combine four bits into a single value. The numbers "0" to "9" and "A" to "F" are used to represent each hex digit. For multi-digit hex numbers, each digit is multiplied by sixteen to the appropriate power:

$$0x0123 \text{ (hex)} = (1 * (16**2)) + (2 * (16**1)) + (3 * (16**0)) \text{ (Decimal)}$$

$$= 256 + 32 + 3 \text{ (Decimal)}$$

$$= 291 \text{ (Decimal)} \tag{17.9}$$

Table 17.17 lists common prefixes used to indicate number bases in programming languages/environments.

Hexadecimal digits greater than "9" are normally given the letters "A" through "F" as shown in Table 17.18.

The negation of a value to its two's complement value is accomplished by the formula:

$$\text{Negative} = (\text{Positive} \wedge 0x0FF) + 1 \tag{17.10}$$

The most significant bit of the two's complement value is known as the "sign bit" and, when set, indicates a negative number and, when reset, indicates a positive number.

It is recommended that computer programs be written with base 10 as the default with other values explicitly specified. This will make reading of applications easier.

Table 17.19 lists different values for the different bases (and two's complement) for eight bits.

TABLE 17.17 Explicit Number Base Specimens

Prefix	Number Type	Base	Comments
0b0...	Binary	2	
B'...'	Binary	2	
%...	Binary	2	
\...	Octal	8	This is the "C" Convention
No Prefix	Decimal	10	
.	Decimal	10	
0x0 ...	Hexadecimal	16	
H'...'	Hexadecimal	16	
H...	Hexadecimal	16	Note that in this case, the first digit must be '0' to '9' or the compiler will assume that the character string represents a variable and not a constant
$ ...	Hexadecimal	16	
0...H	Hexadecimal	16	Note that in this case, the first digit must be '0' to '9' or the compiler will assume that the character string represents a variable and not a constant

TABLE 17.18 Hexadecimal Number Representations

Decimal Value	Hexadecimal Value	Phonetic Name
0	0	0
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	A	"Able"
11	B	"Baker"
12	C	"Charlie"
13	D	"Dog"
14	E	"Easy"
15	F	"Fox"

There is one other numbering system that was popular with IBM "mainframe" computers—"Binary Coded Decimal" ("BCD"). In this numbering system, each base-10 digit of a number is represented as a single nybble. Using BCD, one byte can represent the decimal numbers "00" to "99". Some processors provide "Decimal Adjust for Addition" ("DAA") instructions for converting the result of a BCD addition operation from a hexadecimal result to a BCD result. Other operations are rarely supported using this type of instruction.

Fractional or "Real" data values can also be represented with binary memory. To store Real data values, a format that is very analogous to the "Scientific notation" is used. This type of data is often known as "Floating Point" or "Real" in programming languages.

The most popular floating point data format is known as the "IEEE Standard 754" format, which requires four or eight bits to store a floating point number. The IEEE 754 "Double Precision" floating point number is defined as:

$$s \ mmmm \times (2^{**}(e - bias)) \tag{17.11}$$

where "s" is the sign of the number, "mmmm" is known as the "mantissa" or "precision" of the value, "e" is the 2's complement exponent that denotes the order of magnitude of the mantissa. The exponent has a "bias" value subtracted from it to compensate for the number of mantissa bits in the number.

The 64 bit "Double Precision" data format is:

Bits	Purpose
63	Sign Bit
62–52	Exponent
51–0	Mantissa

TABLE 17.19 8 Bit Conversion Table

Binary	Octal	Decimal	Hexadecimal	Two's Complement
0000 0000	000	0	00	0
0000 0001	001	1	01	1
0000 0010	002	2	02	2
0000 0011	003	3	03	3
0000 0100	004	4	04	4
0000 0101	005	5	05	5
0000 0110	006	6	06	6
0000 0111	007	7	07	7
0000 1000	010	8	08	8
0000 1001	011	9	09	9
0000 1010	012	10	0A	10
0000 1011	013	11	0B	11
0000 1100	014	12	0C	12
0000 1101	015	13	0D	13
0000 1110	016	14	0E	14
0000 1111	017	15	0F	15
0001 0000	020	16	10	16
0001 0001	021	17	11	17
0001 0010	022	18	12	18
0001 0011	023	19	13	19
0001 0100	024	20	14	20
0001 0101	025	21	15	21
0001 0110	026	22	16	22
0001 0111	027	23	17	23
0001 1000	030	24	18	24
0001 1001	031	25	19	25
0001 1010	032	26	1A	26
0001 1011	033	27	1B	27
0001 1100	034	28	1C	28
0001 1101	035	29	1D	29
0010 1110	036	30	1E	30
0010 1111	037	31	1F	31
0010 0000	040	32	20	32
0010 0001	041	33	21	33
0010 0010	042	34	22	34
0010 0011	043	35	23	35
0010 0100	044	36	24	36
0010 0101	045	37	25	37
0010 0110	046	38	26	38
0010 0111	047	39	27	39
0010 1000	050	40	28	40
0010 1001	051	41	29	41
0010 1010	052	42	2A	42
0010 1011	053	43	2B	43
0010 1100	054	44	2C	44
0010 1101	055	45	2D	45
0010 1110	056	46	2E	46
0010 1111	057	47	2F	47
0011 0000	060	48	30	48
0011 0001	061	49	31	49
0011 0010	062	50	32	50

TABLE 17.19 8 Bit Conversion Table (*Continued*)

Binary	Octal	Decimal	Hexadecimal	Two's Complement
0011 0011	063	51	33	51
0011 0100	064	52	34	52
0011 0101	065	53	35	53
0011 0110	066	54	36	54
0011 0111	067	55	37	55
0011 1000	070	56	38	56
0011 1001	071	57	39	57
0011 1010	072	58	3A	58
0011 1011	073	59	3B	59
0011 1100	074	60	3C	60
0011 1101	075	61	3D	61
0011 1110	076	62	3E	62
0011 1111	077	63	3F	63
0100 0000	100	64	40	64
0100 0001	101	65	41	65
0100 0010	102	66	42	66
0100 0011	103	67	43	67
0100 0100	104	68	44	68
0100 0101	105	69	45	69
0100 0110	106	70	46	70
0100 0111	107	71	47	71
0100 1000	110	72	48	72
0100 1001	111	73	49	73
0100 1010	112	74	4A	74
0100 1011	113	75	4B	75
0100 1100	114	76	4C	76
0100 1101	115	77	4D	77
0100 1110	116	78	4E	78
0100 1111	117	79	4F	79
0101 0000	120	80	50	80
0101 0001	121	81	51	81
0101 0010	122	82	52	82
0101 0011	123	83	53	83
0101 0100	124	84	54	84
0101 0101	125	85	55	85
0101 0110	126	86	56	86
0101 0111	127	87	57	87
0101 1000	130	88	58	88
0101 1001	131	89	59	89
0101 1010	132	90	5A	90
0101 1011	133	91	5B	91
0101 1100	134	92	5C	92
0101 1101	135	93	5D	93
0110 1110	136	94	5E	94
0110 1111	137	95	5F	95
0110 0000	140	96	60	96
0110 0001	141	97	61	97
0110 0010	142	98	62	98
0110 0011	143	99	63	99
0110 0100	144	100	64	100
0110 0101	145	101	65	101

TABLE 17.19 8 Bit Conversion Table (*Continued*)

Binary	Octal	Decimal	Hexadecimal	Two's Complement
0110 0110	146	102	66	102
0110 0111	147	103	67	103
0110 1000	150	104	68	104
0110 1001	151	105	69	105
0110 1010	152	106	6A	106
0110 1011	153	107	6B	107
0110 1100	154	108	6C	108
0110 1101	155	109	6D	109
0110 1110	156	110	6E	110
0110 1111	157	111	6F	111
0111 0000	160	112	70	112
0111 0001	161	113	71	113
0111 0010	162	114	72	114
0111 0011	163	115	73	115
0111 0100	164	116	74	116
0111 0101	165	117	75	117
0111 0110	166	118	76	118
0111 0111	167	119	77	119
0111 1000	170	120	78	120
0111 1001	171	121	79	121
0111 1010	172	122	7A	122
0111 1011	173	123	7B	123
0111 1100	174	124	7C	124
0111 1101	175	125	7D	125
0111 1110	176	126	7E	126
0111 1111	177	127	7F	127
1000 0000	200	128	80	−128
1000 0001	201	129	81	−127
1000 0010	202	130	82	−126
1000 0011	203	131	83	−125
1000 0100	204	132	84	−124
1000 0101	205	133	85	−123
1000 0110	206	134	86	−122
1000 0111	207	135	87	−121
1000 1000	210	136	88	−120
1000 1001	211	137	89	−119
1000 1010	212	138	8A	−118
1000 1011	213	139	8B	−117
1000 1100	214	140	8C	−116
1000 1101	215	141	8D	−115
1000 1110	216	142	8E	−114
1000 1111	217	143	8F	−113
1001 0000	220	144	90	−112
1001 0001	221	145	91	−111
1001 0010	222	146	92	−110
1001 0011	223	147	93	−109
1001 0100	224	148	94	−108
1001 0101	225	149	95	−107
1001 0110	226	150	96	−106
1001 0111	227	151	97	−105
1001 1000	230	152	98	−104

TABLE 17.19 8 Bit Conversion Table (*Continued*)

Binary	Octal	Decimal	Hexadecimal	Two's Complement
1001 1001	231	153	99	−103
1001 1010	232	154	9A	−102
1001 1011	233	155	9B	−101
1001 1100	234	156	9C	−100
1001 1101	235	157	9D	−99
1010 1110	236	158	9E	−98
1010 1111	237	159	9F	−97
1010 0000	240	160	A0	−96
1010 0001	241	161	A1	−95
1010 0010	242	162	A2	−94
1010 0011	243	163	A3	−93
1010 0100	244	164	A4	−92
1010 0101	245	165	A5	−91
1010 0110	246	166	A6	−90
1010 0111	247	167	A7	−89
1010 1000	250	168	A8	−88
1010 1001	251	169	A9	−87
1010 1010	252	170	AA	−86
1010 1011	253	171	AB	−85
1010 1100	254	172	AC	−84
1010 1101	255	173	AD	−83
1010 1110	256	174	AE	−82
1010 1111	257	175	AF	−81
1011 0000	260	176	B0	−80
1011 0001	261	177	B1	−79
1011 0010	262	178	B2	−78
1011 0011	263	179	B3	−77
1011 0100	264	180	B4	−76
1011 0101	265	181	B5	−75
1011 0110	266	182	B6	−74
1011 0111	267	183	B7	−73
1011 1000	270	184	B8	−72
1011 1001	271	185	B9	−71
1011 1010	272	186	BA	−70
1011 1011	273	187	BB	−69
1011 1100	274	188	BC	−68
1011 1101	275	189	BD	−67
1011 1110	276	190	BE	−66
1011 1111	277	191	BF	−65
1100 0000	300	192	C0	−64
1100 0001	301	193	C1	−63
1100 0010	302	194	C2	−62
1100 0011	303	195	C3	−61
1100 0100	304	196	C4	−60
1100 0101	305	197	C5	−59
1100 0110	306	198	C6	−58
1100 0111	307	199	C7	−57
1100 1000	310	200	C8	−56
1100 1001	311	201	C9	−55
1100 1010	312	202	CA	−54
1100 1011	313	203	CB	−53
1100 1100	314	204	CC	−52

TABLE 17.19 8 Bit Conversion Table (*Continued*)

Binary	Octal	Decimal	Hexadecimal	Two's Complement
1100 1101	315	205	CD	−51
1100 1110	316	206	CE	−50
1100 1111	317	207	CF	−49
1101 0000	320	208	D0	−48
1101 0001	321	209	D1	−47
1101 0010	322	210	D2	−46
1101 0011	323	211	D3	−45
1101 0100	324	212	D4	−44
1101 0101	325	213	D5	−43
1101 0110	326	214	D6	−42
1101 0111	327	215	D7	−41
1101 1000	330	216	D8	−40
1101 1001	331	217	D9	−39
1101 1010	332	218	DA	−38
1101 1011	333	219	DB	−37
1101 1100	334	220	DC	−36
1101 1101	335	221	DD	−35
1110 1110	336	222	DE	−34
1110 1111	337	223	DF	−33
1110 0000	340	224	E0	−32
1110 0001	341	225	E1	−31
1110 0010	342	226	E2	−30
1110 0011	343	227	E3	−29
1110 0100	344	228	E4	−28
1110 0101	345	229	E5	−27
1110 0110	346	230	E6	−26
1110 0111	347	231	E7	−25
1110 1000	350	232	E8	−24
1110 1001	351	233	E9	−23
1110 1010	352	234	EA	−22
1110 1011	353	235	EB	−21
1110 1100	354	236	EC	−20
1110 1101	355	237	ED	−19
1110 1110	356	238	EE	−18
1110 1111	357	239	EF	−17
1111 0000	360	240	F0	−16
1111 0001	361	241	F1	−15
1111 0010	362	242	F2	−14
1111 0011	363	243	F3	−13
1111 0100	364	244	F4	−12
1111 0101	365	245	F5	−11
1111 0110	366	246	F6	−10
1111 0111	367	247	F7	−9
1111 1000	370	248	F8	−8
1111 1001	371	249	F9	−7
1111 1010	372	250	FA	−6
1111 1011	373	251	FB	−5
1111 1100	374	252	FC	−4
1111 1101	375	253	FD	−3
1111 1110	376	254	FE	−2
1111 1111	377	255	FF	−1

TABLE 17.20 Basic Data Type Specifier Suffixes

Suffix	Data Type
$	String Data
%	Integer
&	Long Integer (32 Bits)—Microsoft BASIC Extension
!	Single Precision (32 Bits)—Microsoft BASIC Extension
#	Double Precision (64 Bits)—Microsoft BASIC Extension

TABLE 17.21 Basic Built in Functions

Statement	Function
BASE	Starting Array Element
DATA	Data Block Header
DIM	Dimension Array Declaration
OPTION	Starting Array Element
LET	Assignment Statement (Not Mandatory)
RANDOMIZE	Reset Random Number "Seed"
INPUT [Prompt ,] Variables	Get Terminal Input
PRINT	Output to a Terminal
?	Output to a Terminal
READ	Get "Data" Information
GOTO	Jump to Line Number/Label
GOSUB	Call Subroutine at Line Number/Label
RETURN	Return to Caller from Subroutine
IF Condition [THEN] Statement	Conditionally Execute the "Statement"
FOR Variable = Init TO Last [STEP Inc] . . . NEXT [Variable]	Loop Specified Number of Times
ON Event GOTO	On an Event, Jump to Line Number/Label
RESTORE	Restore the "DATA" Pointer
STOP	Stop Program Execution
END	End Program Execution
'	Comment—Everything to the Right is Ignored
REM	Comment—Everything to the Right is Ignored
ABS	Get Absolute Value of a Number
SGN	Return the Sign of a Number
COS	Return Cosine of an Angle (input usually in Radians)
SIN	Return Sine of an Angle (input usually in Radians)
SIN	Return Tangent of an Angle (input usually in Radians)
ATN	Return the Arc Tangent of a Ratio
INT	Convert Real Number to Integer
SQR	Return the Square Root of a Number
EXP	Return the Power of e for the input
LOG	Return the Natural Logarithm for the Input
RND	Return a Random Number
TAB	Set Tab Columns on Printer

Before the value can be stored, the exponent has to be calculated with respect to the "Bias Value." For the Double Precision numbers, this value is 1023. To find the correct exponent, the following formula is used:

$$\text{Required Exponent} = \text{Saved Exponent} - \text{Bias Value} \qquad (17.12)$$

The IEEE 745 double precision representation of "6.5" is:

$$\text{0x03FE000000000000D}$$

The BASIC Language   BASIC variables do not have to be declared except in specialized cases. The variable name itself follows normal conventions of a letter or "_" character as the first character, followed by alphanumeric characters and "_" for variable names. Variable (and Address "Label") names may be case sensitive, depending on the version.

To specify data types, a "suffix" character is added to the end of the variable name (see Table 17.20).

In Microsoft BASICs, the "DIM" statement can be used to specify a variable type and whether it is an array:

DIM Variable[([Low TO] High[, [Low TO] High . . .])]   [AS Type]

There are a number of built in statements that are used to provide specific functions to applications as shown in Table 17.21.

For assignment and "if" statements, Table 17.22 lists operators that are available in BASIC. Table 17.23 shows that BASIC's order of operations is quite standard for programming languages. Table 17.24 lists functions that are available in Microsoft Versions of BASIC for the PC.

TABLE 17.22   Basic Operations

Operator	Operation
+	Addition
−	Subtraction
*	Multiplication
/	Division
^	Exponentiation
"	Start/End of Text String
,	Separator
;	Print Concatenation
=	Assignment/Equals To Test
<	Less than
<=	Less than or Equals To
>	Greater than
>=	Greater than or Equals To
<>	Not Equals

TABLE 17.23 Basic Order of Operations

Operators	Priority	Type
Functions		Expression Evaluation
= <> < <= > >=	Highest	Conditional Tests
^		Exponentiation
* /		Multiplication/Division
+ −	Lowest	Addition/Subtraction

TABLE 17.24 Microsoft BASIC Language Enhancement Functions

Function	Operation
AND	AND Logical Results
OR	OR Logical Results
XOR	XOR Logical Results
EQV	Test Equivalence of Logical Results
IMP	Test Implication of Logical Results
MOD	Get the Modulus (remainder) of an Integer Division
FIX	Convert a Floating Point Number to Integer
DEFSTR Variable	Define the Variable as a String (instead of the ''DIM'' Statement)
DEFINT Variable	Define the Variable as an Integer (instead of the ''DIM'' Statement)
DEFLNG Variable	Define the Variable as a ''long'' Integer (instead of the ''DIM'' Statement)
DEFSNG Variable	Define the Variable as a Single Precision Floating Point Number (instead of the ''DIM'' Statement)
DEFDBL Variable	Define the Variable as a Double Precision Floating Point Number (without using the ''DIM'' Statement)
REDIM Variable([low TO] High[, [low TO] High . . .]) [AS Type]	Redefine a Variable
ERASE	Erase an Array Variable from Memory
LBOUND	Return the First Index of an Array Variable
UBOUND	Return the Last Index of an Array Variable
CONST Variable = Value	Define a Constant Value
DECLARE Function | Subroutine	Declare a Subroutine/Function Prototype at Program Start
DEF FNFunction(Arg[, Arg . . .])	Define a Function (''FNFunction'') that returns a Value. If a Single Line, then ''END DEF'' is not required
END DEF	End the Function Definition
FUNCTION Function(Arg[, Arg . . .])	Define a Function. Same Operation, Different Syntax as ''DEF FNFunction''
END FUNCTION	End a Function Declaration
SUB Subroutine(Arg[, Arg . . .])	Define a ''Subroutine'' that does not return a Value. If a Single Line, then ''END DEF'' is not required
END SUB	End the Subroutine Definition

TABLE 17.24 Microsoft BASIC Language Enhancement Functions (*Continued*)

Function	Operation
DATA Value[, Value . . .]	Specify File Data
READ Variable[, Variable . . .]	Read from the ''Data'' File Data
IF Condition THEN Statements ELSE Statements END IF	Perform a Structured If/Else/Endif
ELSEIF Condition THEN . . .	Perform a Condition Test/Structured If/Else/ Endif instead of simply ''Else''
ON ERROR GOTO Label	On Error Condition, Jump to Handler
RESUME [Label]	Executed at the End of an Error Handler. Can either return to current location, 0 (Start of Application) or a specific label
ERR	Return the Current Error Number
ERL	Return the Line the Error Occurred at
ERROR #	Execute an Application-Specific Error (Number ''#'')
DO WHILE Condition Statements LOOP	Execute ''Statements'' while ''Condition'' is True
DO Statements LOOP WHILE Condition	Execute ''Statements'' while ''Condition'' is True
DO Statements LOOP UNTIL Condition	Execute ''Statements'' until ''Condition'' is True
EXIT	Exit Executing ''FOR'', ''WHILE'' and ''UNTIL'' Loops without executing Check
SELECT Variable	Execute based on ''Value'' ''CASE'' Statements used to Test the Value and Execute Conditionally
CASE Value	Execute within a ''SELECT'' Statement if the ''Variable'' Equals ''Value''. ''CASE ELSE'' is the Default Case
END SELECT	End the ''SELECT'' Statement
LINE INPUT	Get Formatted Input from the User
INPUT$(#)	Get the Specified Number (''#'') of Characters from the User
INKEY$	Check Keyboard and Return Pending Characters or Zero
ASC	Convert the Character into an Integer ASCII Code
CHR$	Convert the Integer ASCII Code into a Character
VAR	Convert the String into an Integer Number
STR$	Convert the Integer Number into a String
LEFT$(String, #)	Return the Specified Number (''#'') of left most Characters in ''String''
RIGHT$(String, #)	Return the Specified Number (''#'') of right most Characters in ''String''
MID$(String, Start, #)	Return/Overwrite the Specified Number (''#'') of Characters at Position ''Start'' in ''String''
SPACE$(#)	Returns a String of the Specified Number (''#'') of ASCII Blanks
LTRIM$	Remove the Leading Blanks from a String
RTRIM$	Remove the Trailing Blanks from a String
INSTR(String, SubString)	Return the Position of ''SubString'' in ''String''

TABLE 17.24 Microsoft BASIC Language Enhancement Functions (*Continued*)

Function	Operation
UCASE$	Convert all the Lower Case Characters in a String to Upper Case
LCASE$	Convert all the Upper Case Characters in a String to Upper Case
LEN	Return the Length of a String
CLS	Clear the Screen
CSRLIN	Return the Current Line that the Cursor is on
POS	Return the Current Column that the Cursor is on
LOCATE X, Y	Specify the Row/Column of the Cursor (top left is 1,1)
SPC	Move the Display the Specified Number of Spaces
PRINT USING "Format"	Print the Value in the Specified Format. "+", "#", ".", "^" Characters are used for number formats
SCREEN mode[,[Color][,[Page][,Visual]	Set the Screen Mode. "Color" is 0 to display on a "Color" display, 1 to display on a "Monochrome." "Page" is the Page that receives I/O and "Visual" is the Page that is currently active.
COLOR [foreground][,[background][,border]]	Specify the Currently Active Colors
PALETTE [attribute, color]	Change Color Assignments.
VIEW [[SCREEN] (x1,y1) − (x2,y2)[,[color]][,border]]]	Create a small Graphics Window known as a "Viewport"
WINDOW [[SCREEN] (x1,y1) − (x2,y2)]	Specify the Viewport's logical location on the Display
PSET (x,y)[,color]	Put a Point on the Display
PRESET (x,y)	Return the Point to the Background Color
LINE (x1,y1) − (x2,y2)[,[Color][,[B|BF] [,style]]]	Draw a Line between the two specified points. If "B" or "BF" specified, Draw a Box ("BF" is "Filled)
CIRCLE (x,y),radius[,[color][,[start] [,end][,aspect]]]	Draw the Circle at center location and with the specified "radius". "start" and "end" are starting and ending angles (in radians). "aspect" is the circle's aspect for drawing ellipses
DRAW CommandString	Draw an arbitrary Graphics Figure. There should be spaces between the commands
	Commands:
	U# – Moves Cursor up # Pixels
	D# – Moves Cursor down # Pixels
	E# – Moves Cursor up and to the right # Pixels
	F# – Moves Cursor down and to the right # Pixels
	G# – Moves Cursor down and to the left # Pixels
	H# – Moves Cursor up and to the left # Pixels
	L# – Moves Cursor left # Pixels
	R# – Moves Cursor right # Pixels

TABLE 17.24 Microsoft BASIC Language Enhancement Functions (*Continued*)

Function	Operation
	Mxy – Move the Cursor to the Specified x,y Position
	B – Turn Off Pixel Drawing
	N – Turn On Cursor and Move to Original Position
	A# – Rotate Shape in 90-Degree Increments
	C# – Set the Drawing Color
	P#Color#Border – Set the Shape Fill and Border Colors
	S# – Set the Drawing Scale
	T# – Rotates # Degrees
LPRINT	Send Output to the Printer
BEEP	"Beep" the Speaker
SOUND Frequency, Duration	Make the Specified Sound on the PC's Speaker
PLAY NoteString	Output the Specified String of "Notes" to the PC's Speaker
DATE$	Return the Current Date
TIME$	Return the Current Time
TIMER	Return the Number of Seconds since Midnight
NAME FileName AS NewFileName	Change the Name of a File
KILL FileName	Delete the File
FILES [FileName.Ext]	List the File (MS-DOS "dir"). "FileName.Ext" can contain "Wild Cards"
OPEN FileName [FOR Access] AS #Handle	Open the File as the Specified Handle (Starting with the "#" Character).
	Access:
	I – Open for Text Input
	O – Open for Text Output
	A – Open to Append Text
	B – File is Opened to Access Single Bytes
	R – Open to Read and Write Structured Variables
CLOSE #Handle	Close the Specified File
RESET	Close all Open Files
EOF	Returns "True" if at the End of a File
READ #Handle, Variable	Read Data from the File
GET #Handle, Variable	Read a Variable from the File
INPUT #Handle, Variable	Read Formatted Data from the File using "INPUT," "INPUT USING" and "INPUT$" Formats
WRITE #Handle, Variable	Write Data to the File
PUT #Handle, Variable	Write a Variable to a File
PRINT #Handle, Output	Write Data to the File using the "PRINT" and "PRINT USING" Formats
SEEK #Handle, Offset	Move the File Pointer to the Specified Offset within the File

The "C" Language

"C" Declarations

Constant Declaration:

    const int Label = Value;

Variable Declaration:

    type Label [= Value];

"Value" is an optional Initialization Constant. Where "type" can be:

    char
    int
    unsigned int
    float

Note that "int" is defined as the "word size" of the processor/operating system. For PCs, an "int" can be a Word (16 Bits) or a Double Word (32 Bits).
There may also be other basic types defined in the language implementation. Single dimensional arrays are declared using the form:

    type Label[Size] [= { Initialization Values. .}];

Note that the array "Size" is enclosed within square brackets ("[" and "]") and should not be confused with the optional "Dimension Size Value."
Strings are defined as single dimensional ASCIIZ arrays:

    char String[17] = "This is a String";

where the last character is an ASCII "NUL".
Strings can also be defined as pointers to characters:

    char *String = "This is a String";

although this implementation requires the text "This is a String" to be stored in two locations (in code and data space).
Multidimensional arrays are defined with each dimension separately identified within square brackets ("[" and "]"):

    int ThreeDSpace[32][32][32];

Array dimensions must be specified unless the Variable is a pointer to a Single Dimensional Array.
Pointers are declared with the "*" character after the "type"

    char   * String = "This is a String";

Accessing the address of the Pointer in Memory is accomplished using the "&" character:

    StringAddr = &String;

Accessing the address of a specific element in a String is accomplished using the "&" character and a String Array Element:

    StringStart = &String[n];
        In PC MS-DOS and Win32 programming absolute offset:segment addresses within the PC memory space to avoid problems with varying segments.

    The Variable's "Type" can be "overridden" by placing the new type in front of the variable in brackets:

    (long) StringAddr = 0x0123450000;

"C" Statements

Application start:

```
main(envp)
  char *envp;
{   //   Application Code
    :              //   Application Code
}   //   End Application
```

Function Format:

```
Return_Type Function( Type Parameter [, Type Parameter..])
{   //   Function Start
    :              //   Function Code
  return value;
}   //   End Function
```

Function Prototype:

    Return_Type Function(Type Parameter [, Type Parameter..]);

Expression:

    [(. . .] Variable | Constant [Operator [(. . .] Variable | Constant][)..]]

Assignment Statement:

    Variable = Expression;

"C" Conditional Statements consist of "if", "?", "while", "do", "for" and "switch".
    The "if" statement is defined as:

```
      if ( Statement )
; | { Assignment Statement | Conditional Statement . . . } | Assignment Statement
| Conditional Statement
      [else ;| { Assignment Statement | Conditional Statement . . .} | Assignment
Statement | Conditional Statement ]
```

The "? :" statement evaluates the statement (normally a comparison) and if note equal to zero, execute the first statement, else execute the statement after the ":".

```
      Statement ? Assignment Statement | Conditional Statement : Assignment
Statement | Conditional Statement
```

The "while" statement is added to the application following the definition below:

```
      while ( Statement ) ; | { Assignment Statement | Conditional Statement. . . } |
Assignment Statement | Conditional Statement
```

The "for" statement is defined as:

```
      for ( initialization (Assignment) Statement[,   . . .]; Conditional Statement;
Loop Expression (Increment) Statement[,   . . .])
; | { Assignment Statement | Conditional Statement. . . } | Assignment Statement
| Conditional Statement
```

To jump out of a currently executing loop, "break" statement

```
      break;
```

is used.

The "continue" statement skips over remaining code in a loop and jumps directly to the loop condition (for use with "while", "for" and "do/while" Loops). The format of the statement is:

```
      continue;
```

For looping until a condition is true, the "do/while" statement is used:

```
      do
         Assignment Statement | Conditional Statement. .
      while ( Expression );
```

To conditionally execute according to a value, the "switch" statement is used:

```
      switch( Expression ) {
         case Value:                // Execute if "Statement" == "Value"
            [Assignment Statement | Conditional Statement. .]
            [break;]
         default:                   // If no "case" Statements are True
            [Assignment Statement | Conditional Statement. .]
      }   //   End switch
```

Finally, the "goto Label" statement is used to jump to a specific address:

```
      goto Label;
      Label:
```

To return a value from a function, the "return" statement is used:

return Statement;

"C" Operators   The Statement Operators ("C" Operators) and their respective operations can be found in Tables 17.25 and 17.26. Table 17.27 lists the ordered operations for the "C" language.

"C" Directives   Refer to Table 17.28 for a list of "C" Directives. Note that all Directives start with "#".

TABLE 17.25   "C" Operators

Operator	Operation
!	Logical Negation
^	Bitwise Negation
&&	Logical AND
&	Bitwise AND, Address
||	Logical OR
|	Bitwise OR
^	Bitwise XOR
+	Addition
++	Increment
−	Subtraction, Negation
−−	Decrement
*	Multiplication, Indirection
/	Division
%	Modulus
==	Equals
!=	Not Equals
<	Less Than
<=	Less Than or Equals To
<<	Shift Left
>	Greater Than
>=	Greater Than or Equals To
>>	Shift Right

TABLE 17.26   Compound Assignment Operators

Operator	Operation
&=	AND with the Variable and Store Result in the Variable
|=	OR with the Variable and Store Result in the Variable
^=	XOR with the Variable and Store Result in the Variable
+=	Add to the Variable
−=	Subtract from the Variable
*=	Multiply to the Variable
/=	Divide from the Variable
%=	Get the Modulus and Store in the Variable
<<=	Shift Left and Store in the Variable
>>=	Shift Right and Store in the Variable

TABLE 17.27 Order of Operations

Operators	Priority	Type
( ) [ ] . −>	Highest	Expression Evaluation
− ~ ! & * ++ −−		Unary Operators
* / %		Multiplicative
+ −		Additive
<< >>		Shifting
< <= >= >		Comparison
== !=		Comparison
&		Bitwise AND
^		Bitwise XOR
|		Bitwise OR
&&		Logical AND
||		Logical OR
?:		Conditional Execution
= &= |= ^= +=		Assignments
−= *= /= %=		
>>= << =		
,	Lowest	Sequential Evaluation

"C" Reserved Words. The following words cannot be used in "C" applications as labels:

break
case
continue
default
do
else
for
goto
if
return
switch
while

"C" "Backslash" Characters See Table 17.29 for "C" backslash characters.

Common "C" Functions A list of common "C" functions, with their respective operations, is shown in Table 17.30.

TABLE 17.28 "C" Directives

Directive	Function
#define Label[(Parameters)] Text	Define a Label that will be replaced with "Text" when it is found in the code. If "Parameters" are specified, then replace them in the code, similar to a macro.
#undefine Label	Erase the defined Label and Text in Memory.
#include "File" | <File>	Load the Specified File in Line to the Text. When "<" ">" encloses the Filename, then the file is found using the "INCLUDE" Environment Path Variable. If """" """" encloses the Filename, then the file in the current directory is searched before checking the "INCLUDE" Path.
#error Text	Force the Error listed in "Text"
#if Condition	If the "Condition" is True, then Compile the following code to "#elif", "#else" or "#endif". If the "Condition" is False, then ignore the following code to "#elif", "#else" or "#endif".
#ifdef Label	If the "#define" Label exists, then Compile the Following Code. "#elif", "#else" and "#endif" work as expected with "#if".
#ifndef Label	If the "#define" Label does NOT exist, then Compile the Following Code. "#elif", "#else" and "#endif" work as expected with "#if".
#elif Condition	This Directive works as an "#else #if" to avoid lengthy nested "#if"'s. If the previous condition was False, checks the Condition.
#else	Placed after "#if" or "#elif" and toggles the current Compile Condition. If the Current Compile Condition was False, after "#else", it will be True. If the Current Compile Condition was True, after "#else", it will be False.
#endif	Used to End an "#if", "#elif", "#else", "#ifdef" or "#ifndef" directive.
#pragma String	This is a Compiler dependant Directive with different "Strings" required for different products.

All Directives start with "#" and are executed before the code is compiled.

TABLE 17.29 "C" "Backslash" Characters

String	ASCII	Character
\r	0x00D	Carriage Return ("CR")
\n	0x00A	Line Feed ("LF")
\f	0x00C	Form Feed ("FF")
\b	0x008	Backspace ("BS")
\t	0x009	Horizontal Tab ("HT")
\v	0x00B	Vertical Tab ("VT")
\a	0x007	Bell ("BEL")
\'	0x027	Single Quote (" ' ")
\"	0x022	Double Quote (" " ")
\\	0x05C	Backslash ("\")
\ddd	N/A	Octal Number
\xddd	0x0dd	Hexadecimal Character

TABLE 17.30 Common "C" Functions
As defined by Kernighan and Ritchie

Function	Operation
int getchar(void)	Get one Character from "Standard Input" (the Keyboard). If no Character available, then wait for it.
int putchar(int)	Output one Character to the "Standard Output" (the Screen).
int printf(char *Const[, arg . . .])	Output the "Const" String Text. "Escape Sequence" Characters for Output are embedded in the "Const" String Text. Different Data Outputs are defined using the "Conversion Characters":
	%d, %i—Decimal Integer
	%o—Octal Integer
	%x, %X—Hex Integer (with upper or lower case values). No leading "0x"character String Output
	%u—Unsigned Integer
	%c—Single ASCII Character
	%s—ASCIIZ String
	%f—Floating Point
	%#e, %#E—Floating Point with the precision specified by "#"
	%g, %G—Floating Point
	%p—Pointer
	%%—Print "%" Character
	Different C Implementations will have different "printf" parameters.
int scanf(char *Const, arg [, *arg . . .])	Provide Formatted Input from the user. The "Const" ASCIIZ String is used as a "Prompt" for the user. Note that the input parameters are always pointers. "Conversion Characters" aresimilar to "printf":
	%d—Decimal Integer
	%i—Integer. In Octal if leading "0" or hex if leading "0x" or "0X"
	%o—Octal Integer (Leading "0" Not Required)
	%x—Hex Integer (Leading "0x" or "0X" Not Required)
	%c—Single Character
	%s—ASCIIZ String of Characters. When Saved, a NULL character is put at the end of the String
	%e, %f, %g—Floating Point Value with optional sign, decimal point and exponent
	%%—Display "%" character in prompt
handle fopen(char *FileName, char *mode)	Open File and Return Handle (or NULL for Error). "mode" is a String consisting of the optional characters:
	r—Open File for Reading
	w—Open File for Writing
	a—Open File for Appending to Existing Files
	Some systems handle "Text" and "Binary" files. A "Text" file has the CR/LF characters represented as a single CR. A "Binary" file does not delete any characters.
int fclose(handle)	Close the File.
int getc(handle)	Receive data from a file one character at a time. If at the end of an input file, then "EOF" is returned.
int putc(handle, char)	Output data to a file one character at a time. Error is indicated by "EOF" returned.
int fprintf(handle, char *Const[, arg . . .])	Output String of Information to a File. The same "Conversion Characters" and arguments as "printf" are used.

TABLE 17.30 Common ''C'' Functions (*Continued*)

Function	Operation
int fscanf(handle, char *Const, arg[, arg . . .])	Input and Process String of Information from a File. The same ''Conversion Characters'' and arguments as ''scanf'' are used.
int fgets(char *Line, int LineLength, handle)	Get the next ASCIIZ String from the file.
int fputs(char *line, handle)	Output an ASCIIZ String to a file.
strcat(Old, Append)	Put ASCIIZ ''Append'' String on the end of the ''Old'' ASCIIZ String.
strncat(Old, Append, #)	Put ''#'' of characters from ''Append'' on the end of the ''Old'' ASCIIZ String.
int strcmp(String1, String2)	Compare two ASCIIZ Strings. Zero is returned for match, negative for ''String1'' < ''String2'' and positive for ''String1'' > ''String2''.
int strncmp(String1, String2, #)	Compare two ASCIIZ Strings for ''#'' characters. Zero is returned for match, negative for ''String1'' < ''String2'' and positive for ''String1'' > ''String2''.
strcpy(String1, String2)	Copy the Contents of ASCIIZ ''String2'' into ''String1''.
strncpy(String1, Strint2, #)	Copy ''#'' Characters from ''String2'' into ''String1''.
strlen(String)	Return the length of ASCIIZ Character ''String''.
int strchr(String, char)	Return the Position of the first ''char'' in the ASCIIZ ''String''.
int strrchr(String, char)	Return the Position of the last ''char'' in the ASCIIZ ''String''.
system(String)	Executes the System Command ''String''.
*malloc(size)	Allocate the Specified Number of Bytes of Memory. If insufficient space available, return NUL.
*calloc(#, size)	Allocate Memory for the specified ''#'' of data elements of ''size''.
free(*)	Free the Memory.
float sin(angle)	Find the ''Sine'' of the ''angle'' (which in Radians).
float cos(angle)	Find the ''Cosine'' of the ''angle'' (which in Radians).
float atan2(y, x)	Find the ''Arctangent'' of the ''X'' and ''Y'' in Radians.
float exp(x)	Calculate the natural exponent.
float log(x)	Calculate the natural logarithm.
float log10(x)	Calculate the base 10 logarithm.
float pow(x, y)	Calculate ''x'' to the power ''y''.
float sqrt(x)	Calculate the Square Root of ''x''.
float fabs(x)	Calculate the Absolute Value of ''x''.
float frand( )	Get a Random Number.
int isalpha(char)	Return Non-Zero if Character is ''a''-''z'' or ''A''-''Z''.
int isupper(char)	Return Non-Zero if Character is ''A''-''Z''.
int islower(char)	Return Non-Zero if Character is ''a''-''z''.
int isdigit(char)	Return Non-Zero if Character is ''0''-''9''.
int isalnum(char)	Return Non-Zero if Character is ''a''-''z'', ''A''-''Z'' or ''0''-''9''.
int isspace(char)	Return Non-Zero if Character is '' '', HT, LF, CR, FF or VT.
int toupper(char)	Convert the Character to Upper Case.
int tolower(char)	Convert the Character to Lower Case.

Intel 8051 Microcontroller Instruction Set   For description of the Intel 8051 Microcontroller Instruction Set, the operations, and Op codes, see Table 17.31.

HTML Reference   In describing, ''Hyper Text Markup Language'' (''HTML''), the primary tags (language/object specifier) are listed with the tags that can be used within them (see Table 17.32). Tags are case insensitive and are not subject to any formatting rules.

The basic format for an HTML page is:

```
<html>
   <head>
      <title>Optional: HTML Page Title</title>
   </head>
   <body>
<!   Comment – can be anywhere within ''html'' and ''/html'' tags>
<!   Content of HTML Page placed between ''body'' and ''/body'' tags>
   </body>
</html>
```

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 17.33. Low-speed communications can be accomplished by asynchronous links using a 20-mA current loop (RS-422 or RS-485) or EIA RS-232C Standard.

A fundamental relationship exists between digital and analog information called the *Nyquist criterion.* The required digital pulse rate depends on the highest-frequency component contained in the analog information. Equation (17.13) shows this relationship. The minimum pulse rate (pulses per second) must be at least twice the highest frequency (hertz).

$$pr = 2f \qquad\qquad (17.13)$$

This is a bilateral relationship. The frequency bandwidth of a transmission system must be equal to at least half the pulse rate:

$$BW = \frac{pr}{2} \qquad\qquad (17.14)$$

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 5000 Hz, so the max-

TABLE 17.31   Intel 8051 Microcontroller Instruction Set

Description	Instruction	Operation	Op Code
Close Call to Subroutine (11 Bit Addr)	ACALL Address	SP = SP + 2 @(SP) = PC PC = PC + Address	AAA1 0001 (Adr & 0x0FF)
Close Jump (11 Bit Addr)	AJMP Address	PC = PC + Address	AAA0 0001 (Adr & 0x0FF)
Add to Accumulator	ADD A, Rn ADD A, Direct ADD A, @R0/1 ADD A, Imm	A = A + Rn A = A + Direct A = A + @(Rn) A = A + Imm	0010 1RRR 0010 0101 Direct 0010 0101/0111 0010 0100 Immediate
Add to Accumulator with Carry	ADDC A, Rn ADDC A, Direct ADDC A, @R0/1 ADDC A, #Imm	A = A + Rn + Cy A = A + Direct + Cy A = A + @(Rn) + Cy A = A + Imm + Cy	0011 1RRR 0011 0101 Direct 0011 0101/0111 0011 0100 Immediate
Bitwise And	ANL A, Rn ANL A, Direct ANL A, @R0/1 ANL A, #Imm ANL Direct, A ANL Direct, #Imm	A = A & Rn A = A & Direct A = A & @(Rn) A = A & Imm Direct = Direct & A Direct = Direct & Imm	0101 1RRR 0101 0101 Direct 0101 0110/0111 0101 0100 Immediate 0101 0010 Direct 0101 0011 Direct, Imm
And Bit with Carry	CPL C, Bit	Bit = Bit & Cy	1000 0010 Bit
And NOT Bit with Carry	CPL C, /Bit	Bit = Cy & (Bit $^\wedge$ 1)	1011 0000 Bit
If A is NOT Equal to Direct, Jump	CJNE A, Dir, Adr	if (A != Dir) PC = PC + Adr	1011 0101 Dir Adr
If A is NOT Equal to Immediate, Jump	CJNE A, #Imm, Adr	if (A != Imm) PC = PC + Adr	1011 0100 Imm Adr
If Rn is NOT Equal to Immediate, Jump	CJNE Rn,#Imm,Adr	if (Rn != Imm) PC = PC + Adr	1011 1RRR Imm Adr
If @Rn is NOT Equal to Immediate Jump	CJNE @R0/ 1,#Imm,Adr	if (@Rn != Imm) PC = PC + Adr	1011 0110/0111 Dir Adr
Clear Accumulator	CLR A	A = 0	1110 0100
Clear Bit	CLR Bit	Bit = 0	1100 0010 Bit
Complement Accumulator	CPL A	A = A $^\wedge$ 0x0FF	1111 0100
Complement Bit	CPL Bit	Bit = Bit $^\wedge$ 1	1011 0010 Bit
Complement Carry	CPL C	Cy = Cy $^\wedge$ 1	1011 0011

TABLE 17.31 Intel 8051 Microcontroller Instruction Set (*Continued*)

Description	Instruction	Operation	Op Code
Decimal Adjust for Addition	DA A	if ((A & 15) > 0) | (AC = 1) A = A + 6 if ((A & 0x0F0) > 0) | (Cy = 1) A = A + 0x060	1101 0100
Decrement Accumulator or Register	DEC A DEC Rn DEC Direct DEC @R0/1	A = A − 1 Rn = Rn − 1 Direct = Direct − 1 @Rn = @Rn − 1	0001 0100 0001 1RRR 0001 0100 Direct 0001 0110/0111
Divide	DIV AB	A = A/B B = A // B	1000 0100
Decrement Rn and Continue if Rn = 0	DJNZ Rn, Adr	Rn = Rn − 1 if (Rn != 0) PC = PC + Adr	1101 1RRR Adr
Decrement Direct and Continue if = 0	DJNZ Direct, Adr	Direct = Direct − 1 if (Direct !=0) PC = PC + Adr	1101 0101 Adr
Increment Accumulator or Register	INC A INC Rn INC Direct INC @R0/1	A = A + 1 Rn = Rn + 1 Direct = Direct + 1 @Rn = @Rn + 1	0000 0100 0000 1RRR 0000 0100 Direct 0000 0110/0111
Inc DPTR	INC DPTR	DPTR = DPTR + 1	1010 0011
Indexed Jump	JMP @(A + DPTR)	PC = A + DPTR	0111 0011
If Bit is Set, Jump	JB Bit, Address	if (Bit == 1) PC = PC + Addr	0010 0000 Bit Address
If Bit is Set, Reset it and Jump	JBC Bit, Address	if (Bit == 1) Bit = 0 PC = PC + Addr	0011 0000 Bit Address
If Carry is Set, Jump	JC Address	if (C == 1) PC = PC + Addr	0100 0000 Address
If Bit is Reset, Jump	JnB Bit, Address	if (Bit == 0) PC = PC + Addr	0010 0000 Bit Address
If Carry is Reset, Jump	JNC Address	if (C == 0) PC = PC + Addr	0101 0000 Address
If Accumulator is NOT equal to Zero, Jump	JNZ Address	if (A != 0) PC = PC + Addr	0111 0000 Address
If Accumulator is equal to Zero, Jump	JZ Address	if (A == 0) PC = PC + Addr	0110 0000 Address
Long Call to Subroutine (16 Bit Addr)	LCALL Address	SP = SP + 2 @(SP) = PC PC = Address	0001 0010 Adrlo AdrHi
Long Jump (16 Bit Addr)	LJMP Address	PC = Address	0000 0010 Adrlo AdrHi

TABLE 17.31 Intel 8051 Microcontroller Instruction Set (*Continued*)

Description	Instruction	Operation	Op Code
Move Byte of Data	MOV A, Rn	A = Rn	1110 1RRR
	MOV A, Direct	A = Direct	1110 0101 Direct
	MOV A, @R0/1	A = @(Rn)	1110 0110/0111
	MOV A, #Imm	A = Imm	0111 0100 Immediate
	MOV Rn, A	Rn = A	1111 1RRR
	MOV Rn, Direct	Rn = Direct	1010 1RRR Direct
	MOV Rn, #Imm	Rn = Imm	0111 1RRR Immediate
	MOV Direct, A	Direct = A	1111 0101 Direct
	MOV Direct, Rn	Direct = Rn	1000 1RRR Direct
	MOV Direct, Dir2	Direct = Dir2	1000 0101 Direct, Dir2
	MOV Direct, @R0/1	Direct = @(Rn)	1000 0110/0111 Direct
	MOV Direct, #Imm	Direct = Imm	0111 0101 Direct Imm
	MOV @R0/1, A	@Rn = A	1111 0110/0111
	MOV @R0/1, Direct	@Rn = Direct	1010 0110/0111 Direct
	MOV @R0/1, #Imm	@Rn = Imm	0111 0110/0111 Imm
	MOV DPTR, #Imm	DPTR = Imm	1001 0000 Immlo ImmHi
Move Bit into Carry	MOV C, Bit	Cy = Bit	1010 0010 Bit
Move Carry into Bit	MOV Bit, C	Bit = Cy	1001 0010 Bit
Move from Code Space	MOVC A, @A + DPTR	A = @(A + DPTR)	1001 0011
	MOVC A, @(A + PC)	A = @(A + PC)	1000 0011
Move from External Memory	MOVX A, @R0/1	A = @(Rn + P2)	1111 0010
	MOVX A, @DPTR	A = @(DPTR)	1110 0000
	MOVX @R0/1, A	@(Rn + P2) = A	1111 0010/0011
	MOVX @ DPTR, A	@(DPTR) = A	1111 0000
Multiply	MUL AB	B:A = A * B	1010 0100
Do Nothing	NOP		0000 0000
Bitwise Or	ORL A, Rn	A = A | Rn	0100 1RRR
	ORL A, Direct	A = A | Direct	0100 0101 Direct
	ORL A, @R0/1	A = A | @(Rn)	0100 0110/0111
	ORL A, #Imm	A = A | Imm	0100 0100 Immediate
	ORL Direct, A	Direct = Direct | A	0100 0010 Direct
	ORL Direct, #Imm	Direct = Direct | Imm	0100 0011 Direct, Imm
Or Bit with Carry	CPL C, Bit	Bit = Bit | Cy	0111 0010 Bit
Or NOT Bit with Carry	CPL C, /Bit	Bit = Cy | (Bit $^\wedge$ 1)	1010 0000 Bit
Pop Data from Stack	POP Direct	Direct = @(SP) SP = SP − 1	1101 0000 Direct

TABLE 17.31   Intel 8051 Microcontroller Instruction Set (*Continued*)

Description	Instruction	Operation	Op Code
Push Data onto Stack	PUSH Direct	SP = SP + 1 @(SP) = Direct	1100 0000 Direct
Rotate Accumulator Left	RL A	A = (A << 1) + ((A & 0x080) >> 7)	0010 0011
Rotate Accumulator Left through Carry	RLC A	A = (A << 1) + Cy Cy = (A >> 7)	0011 0011
Rotate Accumulator Right	RR A	A = (A >> 1) + ((A & 0x001) << 7)	0000 0011
Rotate Accumulator Right through Carry	RRC A	A = (A >> 1) + (Cy << 7) Cy = (A & 1)	0001 0011
Return from Subroutine	RET	PC = @(SP) SP = SP − 2	0010 0010
Return from Interrupt	RETI	PC = @(SP) SP = SP − 2	0011 0010
Set Bit	SETB Bit	Bit = 1	1101 0010 Bit
Set Carry	SETB C	Cy = 1	1101 0011
Short Jump (8 Bit Addr)	SJMP Address	PC = PC + Address	1000 0001 Address
Subtract from Accumulator with Carry	SUBB A, Rn	A = A + Rn − Cy	1001 1RRR
	SUBB A, Direct	A = A + Direct − Cy	1001 0101 Direct
	SUBB A, @R0/1	A = A + @(Rn) − Cy	1001 0101/0111
	SUBB A, #Imm	A = A + Imm − Cy	1001 0100 Immediate
Swap Accumulator	SWAP A	A = ((A & 0x00F) << 4) + ((A & 0x0F−) >> 4)	1100 0100
Exchange A With Register	XCH A, Rn	A ⇔ Rn	1100 1RRR
	XCH A, Direct	A ⇔ Direct	1100 0101 Direct
	XCH A, @R0/1	A ⇔ @(Rn)	1100 0110/0111
	XCHD A, @R0/1	A & 15 ⇔ (@Rn & 15)	1101 0110/0111
Bitwise Xor	XRL A, Rn	A = A $^\wedge$ Rn	0100 1RRR
	XRL A, Direct	A = A $^\wedge$ Direct	0110 0101 Direct
	XRL A, @R0/1	A = A $^\wedge$ @(Rn)	0110 0110/0111
	XRL A, #Imm	A = A $^\wedge$ imm	0110 0100 Immediate
	XRL Direct, A	Direct = Direct $^\wedge$ A	0110 0010 Direct
	XRL Direct, #Imm	Direct = Direct $^\wedge$ Imm	0110 0011 Direct, Imm

Notes used in defining the instructions.
A—Accumulator
Direct—Internal 128 RAM, I/O Ports, Control and Status Registers
Rn—Working Register (R0 – R7)

TABLE 17.32 Hyper Text Markup Language ("HTML")

Primary Tag	Tag	Comments
`<html> ... </html>`	`<! String>`	Identifies body of HTML page Comment. Can be anywhere within "html" and "/html" tags
`<html> ... </html>`		Specifies Initial Page Information
`<head> ... </head>`	`<head> ... </head>` `<title>String</title>`	String identifying contents of page—displayed at top of Browser window
`<head> ... </head>`	`<base url=" ... ">String</base>`	Used to record the original address of HTML page for tracking when page copied
`<head> ... </head>`	`<isindex>`	Mark the document as searchable
`<head> ... </head>`	`<style>`	Indicate that style sheets follow
`<head> ... </head>`	`<script>`	Indicate that script code follows
`<head> ... </head>`	`<link url=" ..." [name=" rev"| "rel"] [rel=" made | ..."] [rev=" made | ..."] [urn=" ..."] [title=" ..."] [methods=" ..."]>`	Define relationship between this html page and other objects or documents
`<head> ... </head>`	`<meta [Http-equiv=" ..."] [Http-equiv=" refresh" Content=" n"] [Http-equiv=" refresh" Content=" n; Url= ..."] [Http-equiv=" expires" Content=" string-date"] [Name=" ..."] [Content=" ..."]>`	Specify Meta Information about HTML page. Http-equiv=" ...". – Specify protocol response in server Http-equiv=" refresh" Content=" n" – Reload page after " n" seconds [Http-equiv=" refresh" Content=" n; Url= ..."] – Load new page after " n" seconds [Name=" ..."] – "Meta" name for search engines [Content=" ..."] – "Meta" information tag for search engines
`<head> ... </head>`	`<nextid n=" ..." >`	Specify next element to be used by text generators
`<html> ... </html>`	`<body [Background = "URL"] [BGColor=" color"] [Text=" color"] [Link=" color"] [VLink=" color"] [ALink=" color"]> ... </body>`	Specify start of body of HTML document along with operational parameters
`<body> ... </body>`	`<a href=" URL" [class=" ..."] [target=" ..."]> ... </a>`	Specify a "clickable" link on HTML page
`<body> ... </body>`	`<acronym> ... </acronym>`	Identify acronym in page
`<body> ... </body>`	`<address> ... </address>`	List page author information

17.63

TABLE 17.32 Hyper Text Markup Language ("HTML") (*Continued*)

Primary Tag	Tag	Comments
\<map\> . . . \</map\>	\<area shape="rect" coords="left,top,right,bottom" href=". . ."|nohref\> . . . \</area\>	Define a rectangle in an image map
\<map\> . . . \</map\>	\<area shape="poly" coords="x1,y1,x2,y2,x3,y3, . . ." href=". . ."|nohref\> . . . \</area\>	Define a polygon in an image map
\<map\> . . . \</map\>	\<area shape="circle" coords="x,y,r" href=". . ."|nohref\> . . . \</area\>	Define a circle in an image map
\<body\> . . . \</body\>	\<au\> . . . \</au\>	Identify the name of the author
\<body\> . . . \</body\>	\<blockquote\> . . . \</blockquote\>	Identify extended quoted material
\<body\> . . . \</body\>	\<b\> . . . \</b\>	Make text bold
\<body\> . . . \</body\>	\<basefont [size="n"|"+"|"−"] [color=". . ."\> . . . \</basefont\>	Change the page's base font
\<body\> . . . \</body\>	\<big\> . . . \</big\>	Print text larger than current font
\<body\> . . . \</body\>	\<blink\> . . . \</blink\>	Blink enclosed text
\<body\> . . . \</body\>	\<br\>	Force a Line Break
\<table\> . . . \</table\>	\<caption [align="top"|"bottom"|"left"|"right"]\> . . . \</caption\>	Specify a Caption for a Table or Figure
\<body\> . . . \</body\>	\<center\> . . . \</center\>	Center Text in middle of page
\<body\> . . . \</body\>	\<cite\> . . . \</cite\>	Delimites a Citation
\<body\> . . . \</body\>	\<code\> . . . \</code\>	Delimites programming source code
\<body\> . . . \</body\>	\<del\> . . . \</del\>	Delete text when page amended
\<body\> . . . \</body\>	\<dfn\> . . . \</dfn\>	Indicates Defining instance of a term
\<body\> . . . \</body\>	\<dir\>\<li\> . . . [\<li\> . . .]\</dir\>	Display Text as an unordered list. No element in the list can be longer than 20 characters
\<body\> . . . \</body\>	\<div [align="left"|"right"|"middle"\> . . . \</div\>	Define block of text
\<body\> . . . \</body\>	\<dl [compact]\>\<dt\> . . . \<dd\> . . .[\<dt\> . . . \<dd\> . . .]\< /dl\>	Create a definitions list with "\<dt\>" indicating the term and "\<dd\>" providing definition text
\<body\> . . . \</body\>	\<em\> . . . \</em\>	Indicates emphasized text
\<body\> . . . \</body\>	\<fig\> . . . \</fig\>	Define a figure on the page

`<body>...</body>`	`<font [size="n"|"+"|"-"] [color="..."]>...</font>`	Change the enclosed string's font
`<frameset>...</frameset>`	`<frames [src="..."] [name="..."] [marginwidth="n"] [marginheight="n"] [scrolling="yes"|"no"|"auto"] [Noresize]>...</frame>`	Define Frame Element
`<body>...</body>`	`<frameset rows="*|n[,*|n..." cols="*|n[,*|n...">...</frameset>`	Define a frame for use in the page
`<body>...</body>`	`<h1>...</h1>`	Level 1 Header
`<body>...</body>`	`<h2>...</h2>`	Level 2 Header
`<body>...</body>`	`<h3>...</h3>`	Level 3 Header
`<body>...</body>`	`<h4>...</h4>`	Level 4 Header
`<body>...</body>`	`<h5>...</h5>`	Level 5 Header
`<body>...</body>`	`<h6>...</h6>`	Level 6 Header
`<body>...</body>`	`<hr [size="n"][noshade] [width="n"|"n%"] [align="left"|"right"|"center"]>...</hr>`	Put ruled line to separate text
`<body>...</body>`	`<i>...</i>`	Italicize text
`<body>...</body>`	`<img href="..." [alt="..."] [align="top"|"middle"|"bottom"|"left"|"right"] [width="n"] [height="n"] [vspace="n"] [hspace="n"] [ismap usemap="...">`	Insert Image into a page
`<body>...</body>`	`<ins>...</ins>`	Insert text when page amended
`<body>...</body>`	`<kbd>...</kbd>`	Indicates text to be entered in via keyboard
`<body>...</body>`	`<lang>...</lang>`	Specifies the language currently defined
`<body>...</body>`	`<map name="...">...</map>`	Define a client side Image Map
`<body>...</body>`	`<menu><li>...[<li>...]</menu>`	Display Text as an unordered menu
`<frameset>...</frameset>`	`<noframes>...</noframes>`	Display text if frames support is not present in the browser
`<body>...</body>`	`<ol [type="1"|"a"|"A"|"i"|"I"][start="..."]><li>...[<li>...]</ol>`	Display Text as an unordered list with either a number or the specified data format. List can start at specified value
`<body>...</body>`	`<p [align="left"|"right"|"center"]>...`	Define the start of a paragraph
`<body>...</body>`	`<person>...</person>`	Identify the name of a person
`<body>...</body>`	`<pre>...</pre>`	Display text with monospace font

17.65

TABLE 17.32 Hyper Text Markup Language ("HTML") (*Continued*)

Primary Tag	Tag	Comments
`<body> ... </body>`	`<q> ... </q>`	Indicates Quoted String
`<body> ... </body>`	`<samp> ... </samp>`	Indicates text that should be displayed "AS IS"
`<body> ... </body>`	`<small> ... </small>`	Print text smaller than current font
`<body> ... </body>`	`<strong> ... </strong>`	Indicates emphasized text
`<body> ... </body>`	`<sub> ... </sub>`	Text is a subscript
`<body> ... </body>`	`<sup> ... </sup>`	Text is a superscript
`<body> ... </body>`	`<table [align="bleedleft"|"left"|` `center"|"right"|"bleedright"|"justify"]` `[valign="top"|"middle"|"bottom"|"baseline"]` `[Border|Border="n"] [cellspacing="n"]` `[cellpadding="n"] [width=" n"|"n%"] [class="string"]>` `... </table>`	Define a table to be displayed on Page
`<tr> ... </tr>`	`<td [align="left"|"center"|"right"|` `justify"|"decimal"]` `[valign="top"|"middle"|"bottom"|"baseline"]` `[colspan="n"] [rowspan="n"] [nowrap]> ... </td>`	Specify a Table Element
`<tr> ... </tr>`	`<th [align="left"|"center"|"right"|` `justify"|"decimal"][valign="top"|"` `middle"|"bottom"|"baseline"] [colspan="n"]` `[rowspan="n"] [nowrap]> ... </th>`	Specify a Table Heading Element
`<table> ... </table>`	`<tr> ... </tr>`	Indicate the contents of a table row
`<body> ... </body>`	`<tt> ... </tt>`	Indicates Typewrite font for text
`<body> ... </body>`	`<u> ... </u>`	Indicates Underlined text
`<body> ... </body>`	`<ul [type="disc"|"circle"|"square"]><li> ...[<li> ...]<` `/ul>`	Display Text as an unordered list. Specified "bullet" type at the start of each element
`<body> ... </body>`	`<var> ... </var>`	Indicates variable text name

"Color" is specified in the format "#rrggbb," where "rr" is a two hexadecimal digit value for "red" (0x000 no red, 0x0FF full red). "gg" is a two hexadecimal digit value for "green," and "bb" is a two hexadecimal digit value for "blue".

TABLE 17.33 Organizations Which Provide Communication Standards

CCITT	Comité Consultatif Internationale de Telegraphic 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

imum pulse rate for these lines is about 10,000 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. 17.58. 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.
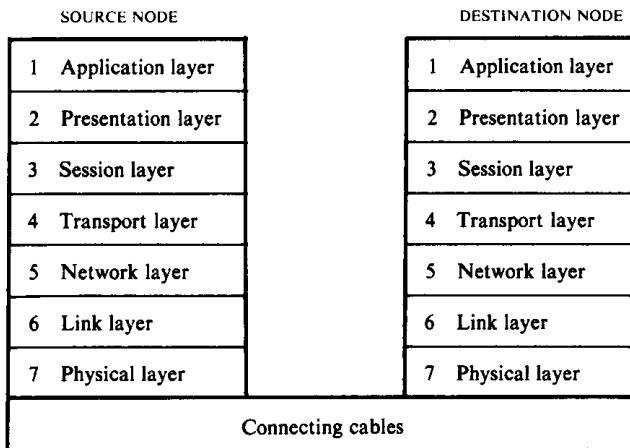
FIGURE 17.58 ISO-layered model for open system interconnection.

For general purpose communications, the Electronics Industries Association has established the RS-232 Communications model shown in Fig. 17.59.

In RS-232, different equipment is wired according to the functions they perform (refer to Table 17.34).

The standard RS-232 connector is either a male 25-pin or male 9-pin connector and is available on the back of the DCE for each serial port as shown in Fig. 17.60.
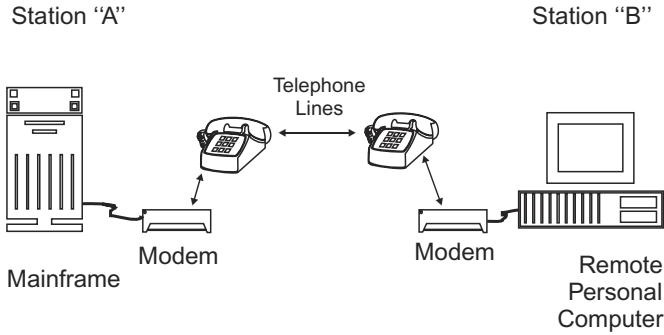
FIGURE 17.59   2 computer communication via modem.

TABLE 17.34

Modem device acronym	Modem device	Function
DTE	Data Terminal Equipment	Computers/Other devices connected to RS-232 interfaces
DCE	Data Communications Equipment	Equipment that allow computers to communicate with one another (i.e., modems)


```
Pin Name     25 Pin     9 Pin      I/O Direction
TxD             2          3        Output ("O")
RxD             3          2        Input ("I")
Gnd             7          5
RTS             4          7        O
CTS             5          8        I
DTR            20          4        O
DSR             6          6        I
RI             22          9        I
DCD             8          1        I
```
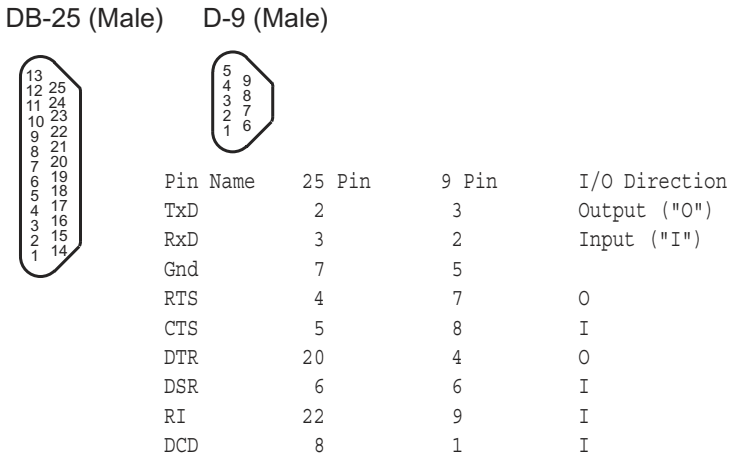
FIGURE 17.60   IBM PC DB-25 and D-9 Pin RS-232 connectors.

The "9-pin" standard was originally developed for the PC/AT because the serial port was put on the same adapter card as the printer port and there wasn't enough room for the serial port <I>and</I> parallel port to both use 25-pin D-Shell connectors.

RS-232 "Mark" ("1") is −5 to −15 volts and a "Space" ("0") is +5 to +15 V. The voltage range of −3 volts to +3 volts is marked as the "switching region" and should be considered indeterminate to allow for noise protection. The RS-232 voltage logic levels are shown in Fig. 17.61.

The six additional lines (which are at the same logic levels as the transmit/receive lines and shown in Fig. 17.61) are used to interface between devices and control the flow of information between computers and are known as "Handshaking" lines (Table 17.35).

There is a common ground connection between the DCE and DTE devices. This connection is critical for the RS-232 level converters to determine the actual incoming voltages. The ground pin should never be connected to a chassis or shield ground (to avoid large current flows or be shifted and prevent accurate reading of incoming voltage signals).

Normally, the handshaking lines are not used and "three-wire RS-232" connections are implemented like Fig. 17.62. To ensure the correct operation at all times,
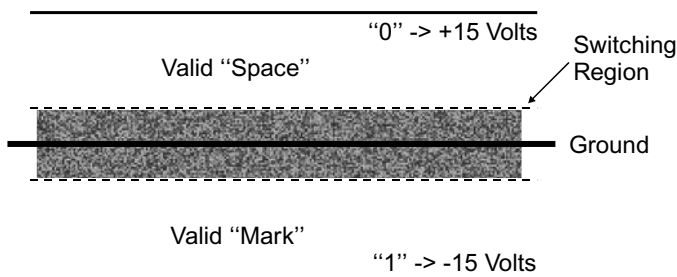
FIGURE 17.61   RS-232 voltage levels.

TABLE 17.35

RS-232 Handshake	Acronym	Function
Request to Send	RTS	From Sender, polling Receiver to see if the Receiver can accept data waiting to be sent.
Clear to Send	CTS	From Receiver, indicating to Sender that it can accept new data.
Data Transmitter Ready	DTR	From Sender, indicating to Receiver that data transfer is about to take place.
Data Set Ready	DSR	From Receiver, indicating to Sender that it is ready to communicate.
Data Carrier Detect	DCD	From Receiver, indicating that it has a connection with another DCE device.
Ring Indicator	RI	From Receiver/modem, indicating whether the connection with another DCE device is in the process of being made or if the line is busy.
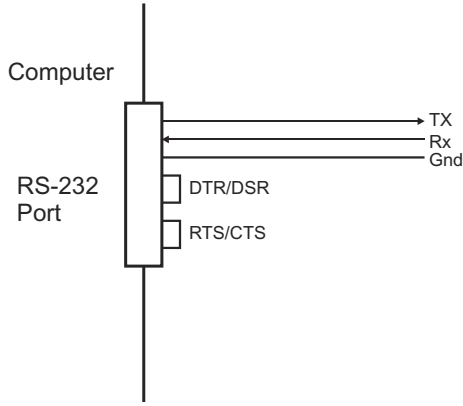
FIGURE 17.62   Typical RS-232 wiring.
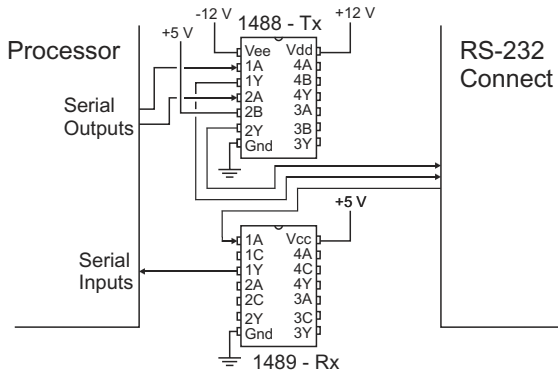
FIGURE 17.63   1488/1489 RS-232 connections.
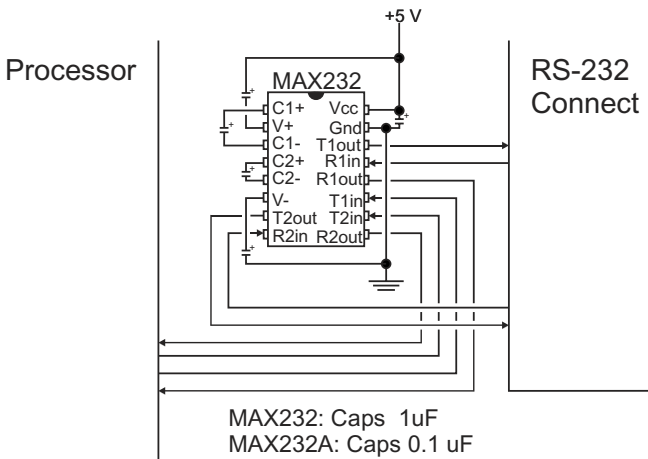
FIGURE 17.64   MAXIM MAX232 RS-232 connections.

the DTR/DSR and RTS/CTS lines should be shorted together at the DTE end of the cable. The DCD and RI lines are left unconnected.

With the three wire RS-232 protocol, there may be applications where software handshaking is required. A very popular standard that provides software handshaking is known as "XON/XOFF" protocol in which the receiver sends an "XOFF" ("DC3" or character 0x013) when it can't accept any more data. When it is able to receive data, it sends an "XON" ("DC1" or character 0x011) to notify the transmitter that it can receive more data.

There are a number of ways in which the RS-232 voltage level standard can be implemented. In the original IBM PC, the 1488/1489 RS-232 Level Converter Circuits were used for this purpose (see Fig. 17.63).

For applications where $+12/-12$ volts are not available, a charge-pump enabled device (like the Maxim MAX232) can be used for changing the levels (Fig. 17.64).

The Dallas Semiconductor DS275 will "steal" the negative voltage of the device being connected to for its negative mark ("1") voltage. This chip is used in an application as shown in Fig. 17.65.

The standard RS-232 data rates or "speeds" are:

Data Rate

110 bps
150 bps
300 bps
600 bps
1,200 bps
2,400 bps
4,800 bps
9,600 bps
19,200 bps
38,400 bps
57,600 bps
115,200 bps

RS-485   A "differential pair" serial communications electrical standard consists of a balanced driver with positive and negative outputs that are fed into a comparator, which outputs a "1" or a "0" depending on whether the "positive line" is
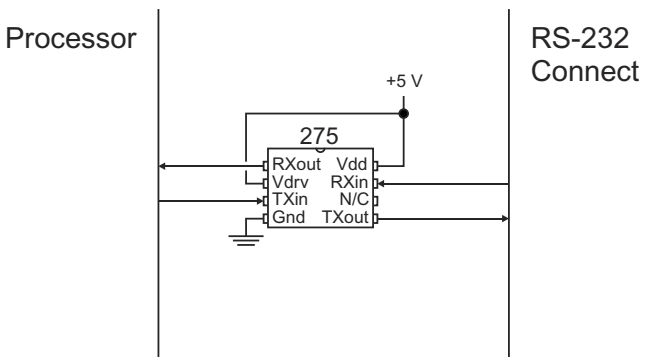
FIGURE 17.65   Dal Semi 275 RS-232 interface.

at a higher voltage than the negative. Fig. 17.66 shows the normal symbols used to describe a differential pair connection.

There are several advantages to this data connection method. The most obvious one is that the differential pair doubles the voltage swing sent to the receiver, and this increases its noise immunity. This is shown in Fig. 17.67. When the "positive" signal goes high, the negative voltage goes low. The change in the two receiver inputs is 10 volts, rather than the 5 volts of a single line. This is assuming the voltage swing is 5 volts for the positive and negative terminals of the receiver. This effective doubling of the signal voltage reduces the impact electrical interface has on the transmitted signal. Another benefit of differential pair wiring is that if on connection breaks, the circuit will operate (although at reduced noise reduction efficiency).

To minimize AC transmission line effects, the two wires should be twisted around each other. "Twisted pair" wiring can either be bought commercially or made by simply twisting two wires together. Twisted wires have a characteristic impedance of 75 ohms or greater.

RS-485 is very similar to RS-422, except it allows multiple drivers on the same network. The common chip is the "75176", which has the ability to drive and receive on the lines as shown in Fig. 17.68.

CAN The "Controller Area Network" protocol was originally developed by Bosch as a networking scheme that could be used to interconnect the computing systems used within automobiles. Before the advent of CAN (and J1850, which is similar North American standard), cars could have up to three miles of wiring weighing two hundred pounds interconnecting the various parts and systems within the car.

"CAN" was designed to be:

1. Fast (1 MBit/Second)
2. Insensitive to Electromagnetic Interference
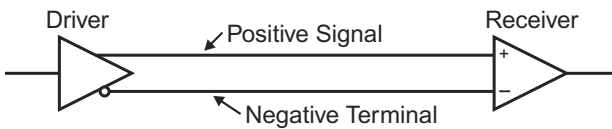
FIGURE 17.66   Differential pair serial data transmission.
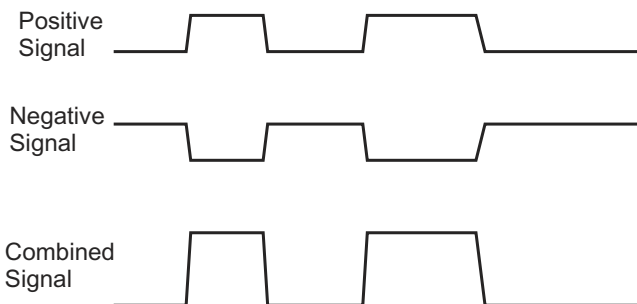
FIGURE 17.67   Differential data waveform.
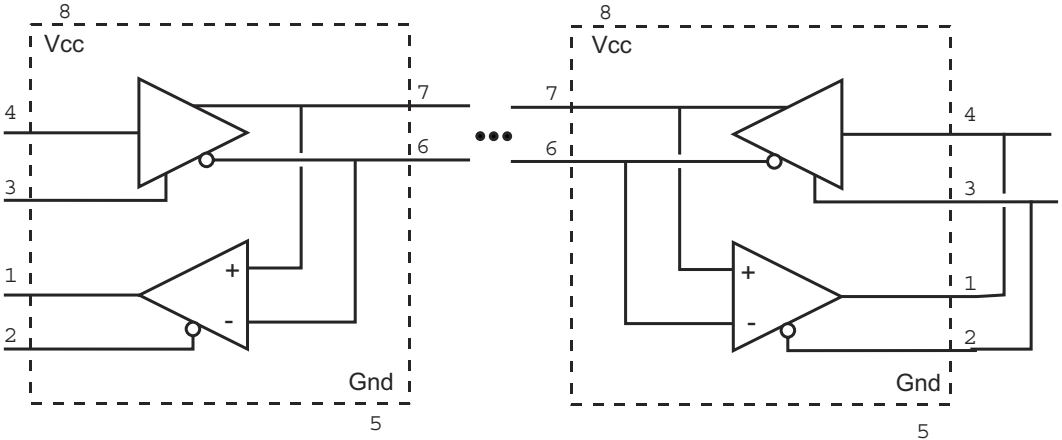
FIGURE 17.68 RS-485 connection using a 75176.

3. Simple with few pins in connectors for mechanical reliability

4. Devices could be added or deleted from the network easily (and during manufacturing)

While CAN is similar to J1850 and does rely on the same first two layers of the OSI seven layer communications model, the two standards are electrically incompatible.

The CAN Frame (Fig. 17.69) is transmitted as an asynchronous serial stream (which means there is no clocking data transmitted). This means that both the transmitter and receiver must be working at the same speed (typically data rates are in the range of 200 kbps to 1 Mbps).

In CAN, a "0" is known as a "Dominant" bit and a "1" is known as a "Recessive" bit.

The different fields of the frame are defined as:

       SOF = Start of Frame, A single Dominant Bit
Identifier = 11 or 19 Bit Message Identifier
      RTR = This Bit is set if the transmitter is also Tx'ing Data
    r1/r0 = Reserved Bits, Should always be Dominant
     DLC = Four Bits indicating the number of bytes that follow
     Data = Zero to 8 Bytes of Data, Sent MSB First
     CRC = 15 bits of CRC data followed by a recessive bit
     Ack = Two Bit field, Dominant/Recessive Bits
     EOF = End of Frame, at least 7 Recessive Bits
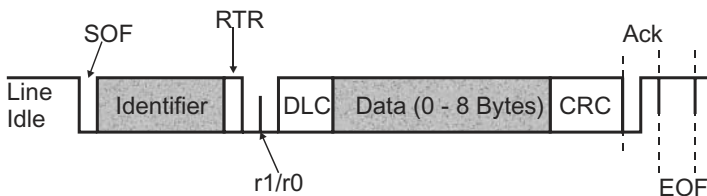
FIGURE 17.69 CAN 11 bit identifier frame.

The last important note about CAN is that devices are not given specific names or addresses. Instead, the message is identified (using the 11 or 19 bits Message Identifier). This method of addressing can provide you with very flexible messaging (which is what CAN is all about).

I2C   The most popular form of microcontroller network is "I2C", which stands for "Inter-Intercomputer Communications." This standard was originally developed by Philips in the late seventies as a method to provide an interface between microprocessors and peripheral devices without wiring full address, data, and control busses between devices. I2C also allows sharing of network resources between processors (which is known as "Multi-Mastering").

The I2C bus consists of two lines: a clock line ("SCL"), is used to strobe data (from the "SDA" line) from or to the master that currently has control over the bus. Both these bus lines are pulled up (to allow multiple devices to drive them). The two bus lines are used to indicate that a data transmission is about to begin as well as pass the data on the bus.

To begin a data transfer, a "Master" drives a "Start Condition" on the bus. Normally (when the bus is in the "Idle State"), both the clock and data lines are not being driven (and are pulled high). To initiate a data transfer, the Master requesting the bus pulls down the SDA bus line followed by the SCL bus line. During data transmission this is an invalid condition (because the data line is changing while the clock line is active/high).

Each bit is then transmitted to or from the "Slave" (the device the message is being communicated with by the "Master") with the negative clock edge being used to latch in the data as shown in Fig. 17.70. To end data transmission, the reverse is executed, the clock line is allowed to go high, which is followed by the data line.

Data is transmitted in a synchronous (clocked) fashion. The most significant bit is sent first, and after eight bits are sent, the master allows the data line to float (it doesn't drive it low) while strobing the clock to allow the receiving device to pull the data line low as an acknowledgment that the data was received. After the acknowledge bit, both the clock and data lines are pulled low in preparation for the next byte to be transmitted or a Stop/Start Condition is put on the bus. Fig. 17.71 shows the data waveform.

Sometimes, the acknowledge bit will be allowed to float high, even though the data transfer has completed successfully. This is done to indicate that the data
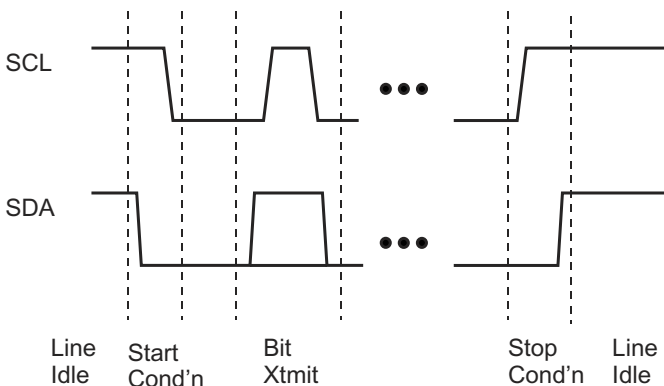
FIGURE 17.70   12C signals and waveforms.
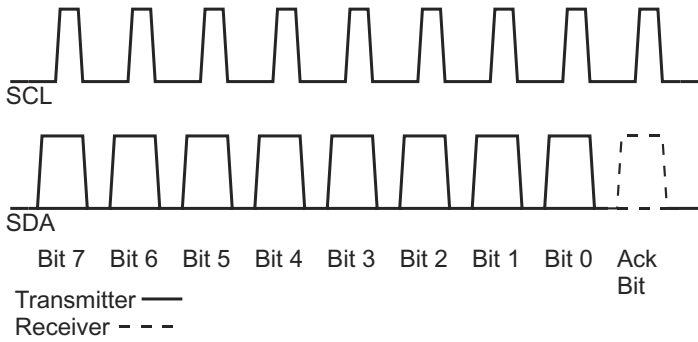
FIGURE 17.71 I2C data byte transmission.

transfer has completed and the receiver (which is usually a "slave device" that is unable to initiate data transfer) can prepare for the next data request.

There are two maximum speeds for I2C (because the clock is produced by a master, there really is no minimum speed), "Standard Mode" runs at up to 100 kbps and "Fast Mode" can transfer data at up to 400 kbps. Fig. 17.72 shows the timing specifications for both the "Standard" ("Std." or 100 kHz data rate) and "Fast" (400 kHz data rate).

A command is sent from the master to the receiver in the following format shown in Fig. 17.73.

The "Receiver Address" is seven bits long and is the bus address of the receiver. There is a loose standard to use the most significant four bits to identify the type of device, while the next three bits are used to specify one of eight devices of this type (or further specify the device type).

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.

Many of the existing techniques and equipment that are used for digital communication by telephone lines are not suited to local data communication needs.
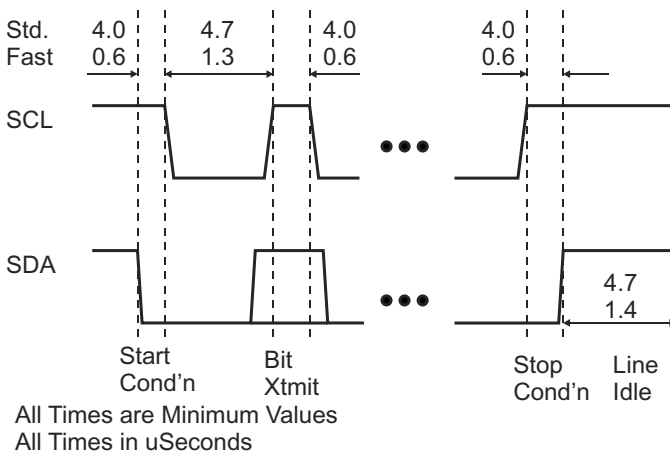
FIGURE 17.72 I2C signal timing.
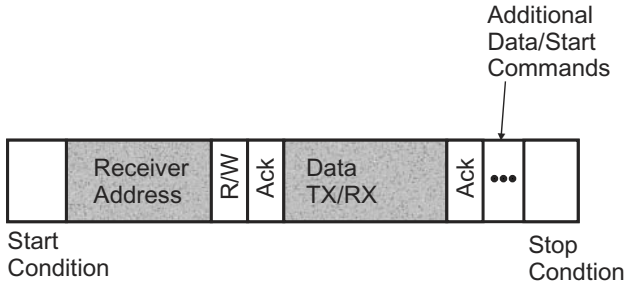
FIGURE 17.73 I2C data transmission.

Large-systems requirements place an overhead on the communication nodes that become burdensome. This reduces throughput and also introduces message setup delays that are unacceptable for many interactive computing situations. *Local area networks* (*LANs*) have been devised to eliminate some of these problems. LANs can operate over distances of up to 1000 ft or so with data rates of 106 baud. Table 17.36 lists several LANs types.

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 satisfied 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 600×10^6 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. 17.15. Single-phase bridge circuits are used for 5-hp drives and

TABLE 17.36 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./IBM
Interactive System	3M
Data Exchange	Amdax
System 20	Sytek/NRC
Wangnet	Wang Computer Corp.

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 electric

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.

WIRELESS COMMUNICATIONS

The Federal Communications Commission (FCC) regulates the use of radio-frequency transmission in the United States. This regulation is necessary to prevent interfering transmissions of radio signals. Some of the frequency allocations are given in Table 17.37. The frequency bands are also classified as shown in Table 17.38. 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. 17.74. 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. 17.75. 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

TABLE 17.37 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

TABLE 17.38 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.
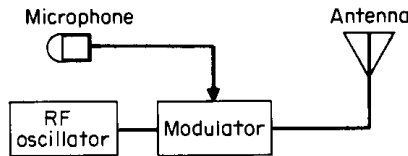
FIGURE 17.74 Radio transmitter.
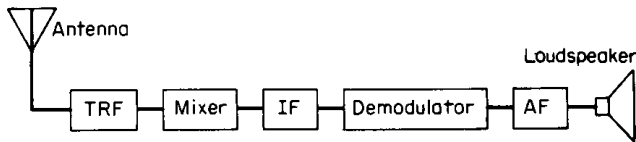
TRF Tuned radio frequency amplifier

IF Intermediate frequency amplifier

AF Audio frequency amplifier

FIGURE 17.75 Radio receiver.

which is the difference between the incoming radio-signal frequency and the local-oscillator frequency. Since this difference frequency is constant for all tuning positions, 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 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.

A block diagram of a radar system is shown in Fig. 17.76. 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 2000 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 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.

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.

1. 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. 17.77*a*. When the pulser operates, a sawtoothed wave produces a linear sweep voltage (Fig. 17.77*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. 17.77*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
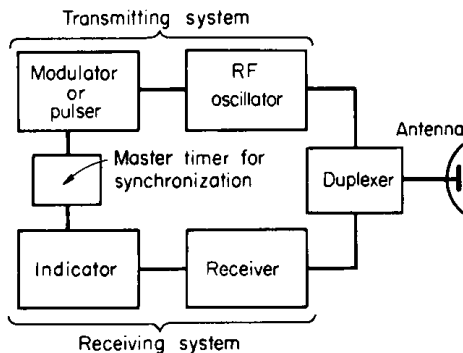
FIGURE 17.76 Block diagram of radar system.
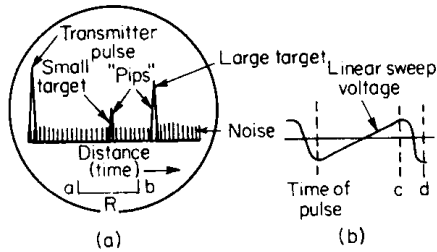
FIGURE 17.77   Type A presentation.

calibrated in distance such as miles. [The return of the spot to its initial starting position, produced by the sweep interval *cd* (Fig. 17.77*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.

2. Plan Position Indicator (PPI).   In the PPI (Fig. 17.78) 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
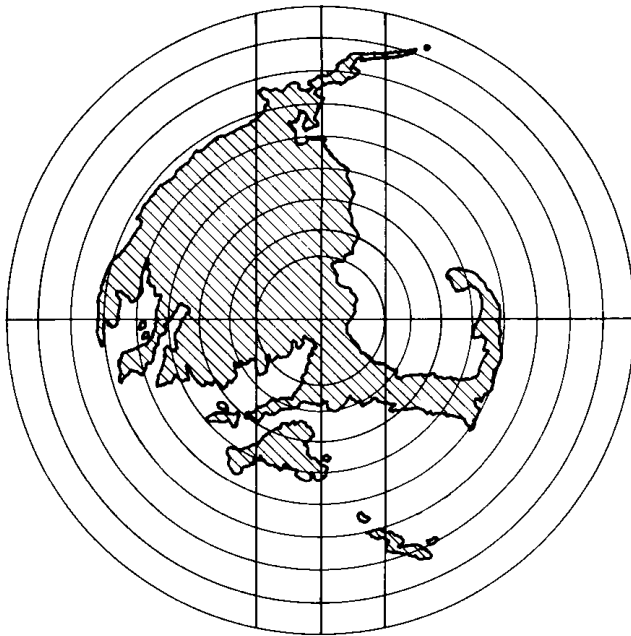
FIGURE 17.78   Plan position indicator (PPI) of southeastern Massachusetts.

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. 17.78.

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 screen be used in order that the operator may view the entire pattern. In Fig. 17.78 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* (*long-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.

3. 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. 17.79. 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 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. 17.79.

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. 17.79, for the sake of clarity, the distances between horizontal lines are greatly exaggerated and no attempt is made to maintain proportions.

4. 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 17.37 and 17.38).

FIGURE 17.79   Pattern of interlaced scanning.

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.

5. Kinescope.   The kinescope (Fig. 17.80) 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 8000 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 normally controlled by deflection coils, as shown in Fig. 17.80.

6. Television Receivers.   A block diagram for a television receiver is given in Fig. 17.81. 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 inter-
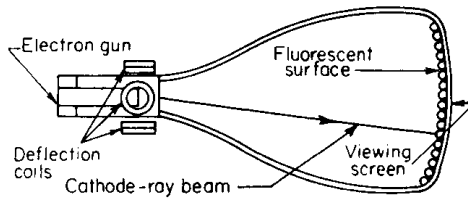
FIGURE 17.80   Kinescope for television receiver.
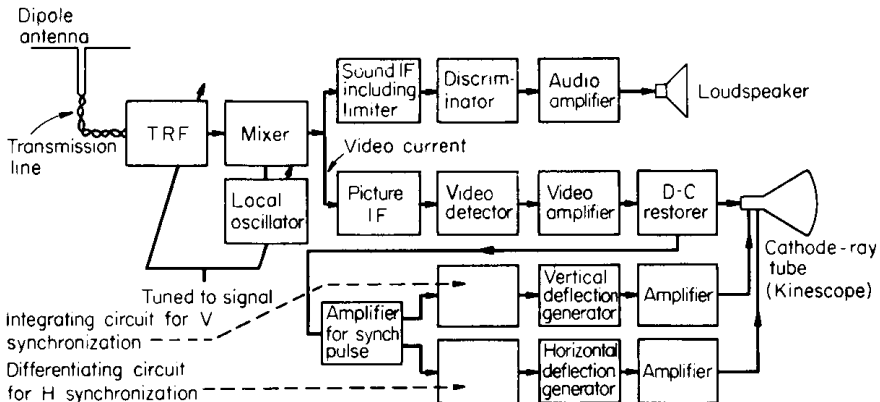
FIGURE 17.81   Block diagram for television receiver (TRF, tuned radio frequency; IF, intermediate frequency).

mediate frequency in the mixer. The sound channel is then conventional, a discriminator being used to demodulate the FM wave (Fig. 17.27). 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. 17.81. 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.

CHAPTER 18

RELIABILITY ENGINEERING, SYSTEMS ENGINEERING, AND SAFETY ENGINEERING

RELIABILITY ENGINEERING*

Reliability is the characteristic of a component, or system made up of many components, expressed by the probability that it will perform its particular function within a specific environment for a given period of time. Since the subject of reliability is obviously concerned with statistical analysis and the prediction of behavior based on tests, it might be considered, at first glance, to be a matter of guesswork or chance. However, reliability predictions have become a precise branch of industrial technology. Reliability engineering plays an invaluable part in the reduction of costly failures and the correct planning of overhaul and maintenance schedules.

TYPES OF FAILURES

Failure is defined as the inability of a component or system to carry out its specified function. Failures may be categorized in a number of ways according to the degree of failure, the reason for failure, the timing of the failure, and so on. The following are some relevant definitions:

Misuse failure is used to describe failure due to overloading or otherwise overstressing a component or system beyond its capability.

Inherent-weakness failure is used to describe failure due to inherent weakness of the component or system and occurring while the item is being correctly used.

Sudden failure is used to describe failures which could not have been anticipated.

Gradual failure is used to describe failures which could have been anticipated.

Partial failure is one in which the component or system may still function, but not to the limits of performance originally designed.

Complete failure results in total loss of the required function.

Catastrophic failure is one which is both sudden and complete.

Degradation failure is one which is both gradual and partial.

---

*This section is drawn from *An Introduction to Reliability Engineering,* by R. Lewis. Copyright © 1970. Used by permission of McGraw-Hill, Inc. All rights reserved. Updated 2003 by editors.

Chance failures is a term used generally to describe those failures which occur suddenly and at random during the anticipated useful life of a component or system. They may be due to a variety of causes, including inherent weakness, misuse, etc.; they are not due to the component having completed its normally anticipated useful life, i.e., to the component wearing out.

Wearout failure is another general term to describe failure due to the wearing out of a component which has more or less completed its anticipated useful life.

Both chance and wearout failures may be partial or complete.

FAILURE RATE

The number of failures occurring per unit time is known as the *failure rate.* As with all quantities describing change (speed, acceleration, etc.), an average value may be obtained by dividing the total number of failures which have occurred during a time interval by the length of the interval. The shorter the interval, the nearer the average value gets to the instantaneous failure rate. The *instantaneous failure rate* at any one time is the slope of the curve plotting failures against time at that particular time.

If, in determining the failure rate, the number of failures occurring during the time interval is expressed as a proportion of the number of survivors at the beginning of the time interval, then the failure rate obtained is called the *proportional failure rate.* It is denoted by the symbol λ. Unless otherwise stated the words *failure rate* within this handbook imply the proportional failure rate.

THE BATHTUB DIAGRAM

A typical graph plotting the percentage failure rate with respect to time is shown in Fig. 18.1. It is often referred to as a *bathtub diagram* because of its shape.

During the burn-in period, a high failure rate exists, owing to the presence of substandard components in the sample tested. After the weak components have died out, the failure rate stabilizes at an approximately constant value; this period is called the *useful-life period.*
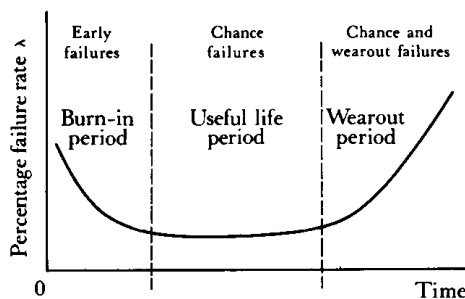
FIGURE 18.1 Bathtub diagram.

Eventually wearout failures begin to occur, and the failure rate rises again. During this wearout period, chance failures may, of course, still be occurring.

In order to make reasonably accurate predictions of reliability, failures due to chance and wearout must be studied and an analysis made of the times in relation to the type of failure.

CONSTANT-FAILURE-RATE CASE

When the failure rate is constant, reliability prediction is made much easier mathematically, since it is possible to use exponential curves to assist analysis.

As was shown above, the failure rate may be assumed to be constant when failures are due to chance alone; this can be achieved by correct overhaul schedules, which eliminate wearout failure. It is also possible, however, to achieve a constant failure rate, and thus simplify the mathematics involved, by a process of immediate replacement on wearout. This latter case is not as obvious but has been conclusively demonstrated. It should be noted that replacement on failure is a procedure which cannot always be adopted, since certain systems or components cannot be allowed to fail even temporarily.

RELIABILITY EQUATIONS AND CURVES WHEN FAILURE RATE IS CONSTANT

The probability of no failures occurring in a given time can be expressed by the following equation, provided that the failure rate is constant:

$$R = e^{-\lambda t} \tag{18.1}$$

where R is the probability of no failures in time t (i.e., the reliability), e is the exponent 2.7183, and λ is the constant failure rate. (This is in fact the no-event term of the Poisson probability function.) The *unreliability* Q is defined as the probability of total failure. It follows logically that

$$R + Q = 1 \tag{18.2}$$

and that
$$Q = 1 - e^{-\lambda t} \tag{18.3}$$

where Q is the probability of total failure in time t. A graph of R and Q against time yields the familiar exponential curves shown in Fig. 18.2.

Notice that at time $t = 0$, $R = 1$, and $Q = 0$, at time $t = 1/\lambda$, $R = 0.37$, and $Q = 0.63$ (from tables of values of e raised to various powers). A graph of survivors, i.e., the number of components still alive at time t against time, will yield the same shape as the reliability curve in Fig. 18.2. A graph of failures against time yields the same shape as the unreliability curve (see Fig. 18.3). The equation of the graph of survivors versus time (survival curve) is

$$N_s = N_0 e^{-\lambda t} \tag{18.4}$$

where N_s is the number of survivors at time t, and N_0 is the original number in the
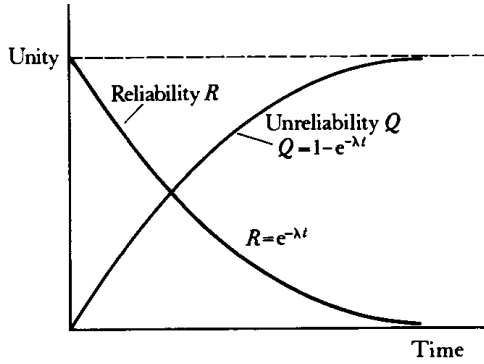
FIGURE 18.2 Reliability and unreliability curves.
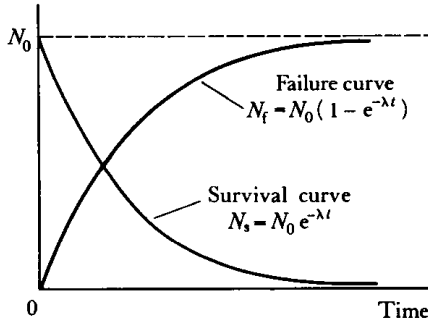
FIGURE 18.3 Failure and survival curves.

test sample. Similarly, the equation of the failures versus time graph (failure curve) is

$$N_f = N_0(1 - e^{-\lambda t}) \tag{18.5}$$

where N_f is the number of failures at time t, and N_0 is as before.

Using these curves and the preceding equations, the reliability or unreliability, number of survivors or failures, may be calculated at any time, provided that the failure rate λ is known (*and* constant, since the exponential analysis only applies when this is so).

FAILURE INTERVALS

1. Mean Time Between Failures (MTBF). The *mean time between failures* (MTBF) of a component or system is an extremely important characteristic in reliability predictions. It is defined as the mean or average time which elapses between failures, and it usually refers to a situation in which the failure rate λ is constant,

i.e. due to chance failures or the adoption of the replacement-on-failure technique described earlier. The symbol for MTBF in these cases is m.

To appreciate the theoretical derivation of m, we must look more closely at the survival curve and the reliability curve and examine the significance of the area contained between the curves and the axes. First, consider a survival curve which is not exponential, such as the one shown in Fig. 18.4.

This is in fact a very simple function indeed and shows a constant number of survivors, i.e., a zero failure rate. At time T_1 shown, the survivors number N_{s_1}. Multiplying N_{s_1} by T_1 gives us the *survivor-hours,* i.e., the total number of hours of survival by all components. This value $N_{s_1} T_1$ is clearly the area bounded by the curve, the axes, and the line $t = T_1$. The area beneath a survival curve, whatever its shape, is in fact always equal to the total survival hours of all the components. To return to the exponential survival curve, which we have seen is obtained for a constant (nonzero) failure rate, the same principle applies. In this case, after a very long time (depending on the failure rate) all components N_0 have failed and the survival curve reaches the time axis. The area under the curve represents the total survival hours of all N_0 components. If we now divide this area by the total failures, that is, N_0, this gives the average time between failures, the MTBF m. The area under a nonlinear curve such as this is, of course, obtained by integration and in this case

$$\text{MTBF} = \left[-\frac{e^{-\lambda t}}{\lambda} \right]_0^{\mathbf{x}}$$

$$\text{MTBF} = 0 - (-1/\lambda) = 1/\lambda \tag{18.6}$$

and we find that the MTBF is, in fact, the *reciprocal* of the failure rate.

Since the survival curve is the same curve as the reliability curve, except that the vertical axis of the reliability curve has been multiplied by N_0 to give the survivors axis of the survival curve, then it follows that the MTBF is in fact equal to the area under the reliability curve divided by unity, i.e., equal to the area itself.

Since

$$m = \frac{1}{\lambda}$$

the equations for reliability R and unreliability Q at any time can be rewritten as
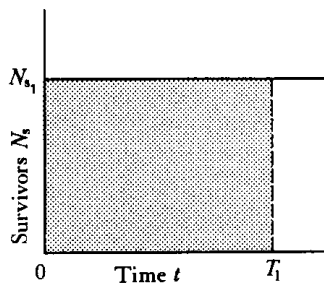
FIGURE 18.4 A nonexponential survival curve.

$$R = e^{-t/m} \tag{18.7}$$

and
$$Q = 1 - R = 1 - e^{-t/m} \tag{18.8}$$

and the equations for the number of survivors N_s or failures N_f at any time can be written

$$N_s = N_0 e^{-t/m} \tag{18.9}$$

$$N_f = N_0(1 - e^{-t/m}) \tag{18.10}$$

where N_0 is the number of components alive at the start of the test.

Notice that since $m = 1/\lambda$, and since, as was shown earlier, when $t = 1/\lambda$ the reliability has fallen to 0.37 and the survivors to 0.37 N_0, the MTBF for the exponential case is the time at which

$$R = 0.37 \qquad Q = 0.63$$

$$N_s = 0.37N_0 \qquad N_f = 0.63N_0$$

2. Measurement of MTBF (Chance Failures).   The two main types of tests for measuring m for chance failures are the nonreplacement and the replacement. The latter is seldom used because it involves constant observation to ascertain exact moments of failure. Nonreplacement methods require observation only at the beginning and end of the test time. To exclude wearout failures, the test is truncated (cut off) before the wearout probability is too high. It has been demonstrated by Epstein that the best estimate of m for a truncated test is given by

$$m = \frac{\text{test hours for failures} + \text{test hours for survivors}}{\text{number of failures}}$$

that is,

$$m = \frac{\text{total component test hours (survival hours)}}{\text{total number of failures}} \tag{18.11}$$

It will be appreciated that the figure obtained is only an estimate; the confidence which one can have in such an estimate may be determined using standard statistical methods.

For a constant failure rate, the MTBF also may be determined by finding the reciprocal of the failure rate, as shown above.

Worked examples on the determination of failure rate, MTBF, reliability, and survivors, etc., using the formulas given so far, are provided in standard reference works on the subject.

3. Wearout Failures: Mean Wearout Life.   As has been shown, chance failures are distributed exponentially, approximately 63 percent occurring before a time equal to the MTBF and approximately 37 percent occurring afterwards. Failures due to wearout, i.e., to the component having completed its anticipated useful life, are not distributed in this manner. A graph of wearout failures against time has the shape shown in Fig. 18.5.

This type of distribution is well known in statistical analysis; it is called the *Gaussian* or *normal* distribution and has certain defined characteristics which clearly distinguish it from other distributions.
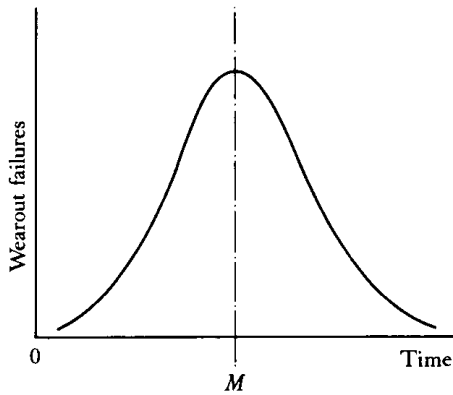
FIGURE 18.5   Wearout failure curve.

The failures due to wearout are clustered about an average or *mean* value of time. Since each point on the curve corresponding to, say, x failures occurring at time t means that x of the components failed after a useful life t, the mean value M corresponds to the *mean wearout life* of the components. The individual lives of the components are scattered *normally* about the mean. Mean wearout life must not be confused with mean time between chance failures, discussed above. The MTBF (chance) tells us the average time anticipated between chance failures *during* the useful life of a component or system; the mean wearout life, on the other hand, tells the average value of the anticipated useful life assuming failure due to chance does not occur. Suppose, for example, a component has a value of M equal to 10,000 h; its mean time between chance failures (computed from a test which is truncated long before 10,000 h) may be as high as 100,000 h. These two figures indicate that provided the component is used within the useful life period, i.e., up to 10,000 h, the probability of chance failure using the formulas described above and a failure rate λ of $1/m$, that is, $1/100,000$ or 0.00001, is quite low. After the anticipated useful life is over, the probability of failure rises very rapidly due to a very much increased failure rate, which is, itself, due to wearout beginning to take place. Wearout and chance failures can be distinguished from one another by careful examination of the physical characteristics of the dead component.

An important parameter concerned with any normal distribution is the *standard deviation σ*. This is computed by finding the square root of the mean of the square of the deviations of the measured characteristic from the average value. That is,

$$\sigma = \sqrt{\frac{\text{sum of squares of deviations from average}}{\text{total number of observations}}} \qquad (18.12)$$

or, for wearout tests, if the life of components 1, 2, 3, . . . , n is indicated by t_1, t_2, t_3, . . . , t_n and the total number used is n, then

$$\sigma = \frac{\sqrt{(t_1 - M)^2 + (t_2 - M)^2 + \cdots + (t_n - M)^2}}{n} \qquad (18.13)$$

where M is the mean (wearout) life.

For a normal distribution of failures it can be shown that approximately 68 percent of the failures occur within a period $M \pm \sigma$, that is, between time $M - \sigma$

and time $M + \sigma$; approximately 95 percent occur within a period $M \pm 2\sigma$; and approximately 99.7 percent occur within a period $M \pm 3\sigma$. This is useful when establishing the confidence that one can have in estimates of wearout life (or indeed of any variable which has a normal distribution). See Fig. 18.6.

4. Measurement of Mean Wearout Life. For a wearout life test, a sample of components is put on test under the conditions they will experience in service, and the test is run until the components fail. Careful examination of both physical characteristics of the dead components and of their life length eliminates chance failures. (It is clear, for example, that if a group of components has lives centering around, say, 10,000 h, then a component surviving only 1000 h is probably not a wearout failure.)

The mean life is then computed as follows:

$$M = \frac{\text{sum of lives of components}}{\text{number of components}} \qquad (18.14)$$

The standard deviation may be determined using the equation already given.

SYSTEM RELIABILITY

1. Inclusion of Subunits in Systems: Intermittent Operation. A system very often contains subunits which are not required to function the whole time during the system operating period. In these cases, care must be taken when assessing the MTBF, failure rate, and reliability to take into account the unit operating time and the fact that it is not equal to the system operating time.

We shall consider the exponential case only since this is the most commonly found. Consider a component or subunit having a reliability R_c for t_c component operating hours. R_c is given by
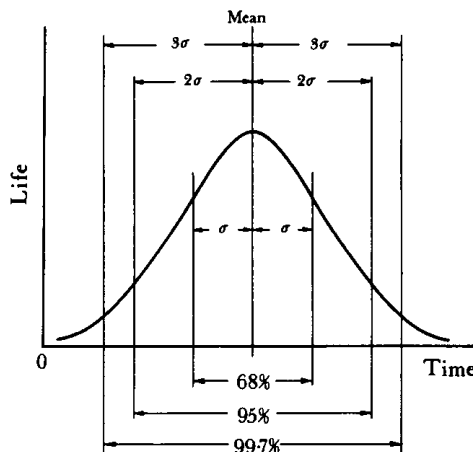
FIGURE 18.6 Percentage points on the normal distribution curve.

$$R_c = e^{-t_c/m_c}$$

where m_c is the mean time between failure of the component expressed in component operating hours. If now this component is inserted in an exponentially failing system which is operating for t_s hours, of which the component is to operate only t_c hours, then the component reliability for t_s system operating hours will be the same as for t_c component operating hours. The reliability of the component, expressed as a function of system hours, is given by

$$e^{-t_s/m_s}$$

where m_s is the mean time between failure of the component expressed in system operating hours.

The reliability of the component is the same, however it is expressed, since it is operated for t_c hours whether alone or as a part of a system. Thus

$$R_c = e^{-t_c/m_c} = e^{-t_s/m_s}$$

and therefore

$$\frac{t_c}{m_c} = \frac{t_s}{m_s}$$

or

$$\frac{m_c}{m_s} = \frac{t_c}{t_s} \tag{18.15}$$

The ratio t_c/t_s expresses the fraction of system time during which the component is required to operate and is called the *duty cycle* of the component, indicated by d. Thus we have

$$m_s = \frac{m_c}{d} \qquad \text{where } d = \frac{t_c}{t_s} \tag{18.16}$$

2. System Reliability.   In calculating the reliability of a system made up of a number of components or individual complete units, each having their own reliability, the *type* of system must first be determined. There are two types—series systems and parallel systems.

A *series system* is one in which failure of one of the subunits or components means failure of the system as a whole. A *parallel system* is one which does not fail until all subunits or components have failed.

It will be remembered that reliability is the probability of survival; in computing system reliability, various well-established laws of the mathematics of probability are used. The relevant laws are discussed below.

3. Laws of Probability Relevant to Reliability Calculations

1. If X and Y are two independent events and P_x is the probability that X will occur and P_y is the probability that Y will occur, then the probability that *both* events X and Y will occur P_{xy} is given by

$$P_{xy} = P_x P_y \tag{18.17}$$

2. If the two events can occur simultaneously, the probability that either X or Y or both X and Y will occur P_{x+y} is given by

$$P_{x+y} = P_x + P_y - P_x P_y \tag{18.18}$$

Both these laws may be extended to cover more than two events.

4. Reliability of a Series System. The reliability of a series system or probability of survival of the system is the probability of *all* the components surviving, since a failure of only one component means overall system failure.

If R_s is the system reliability and R_1, R_2, etc. are the reliabilities of the system components for the same time period, then using law 1,

$$R_s = R_1 R_2 \tag{18.19}$$

In this case, the "event" described in law 1 is survival. This expression is called the *product law of reliabilities* for series systems. For a system having more than two components, the additional reliabilities, R_3, R_4, etc., are included in the product term.

5. Unreliability of a Series System. The unreliability of a series system or probability of failure of the system is the probability of at least *one* of the system components failing.

If Q_s is the system unreliability and Q_1, Q_2, etc. are the unreliabilities of the system components for the same time period, then using law 2,

$$Q_s = Q_1 + Q_2 - Q_1 Q_2 \tag{18.20}$$

for a system having two components. It will be recalled that unreliability $= 1 -$ reliability, that is, in general, $Q = 1 - R$, so that

$$Q_1 = 1 - R_1$$

$$Q_2 = 1 - R_2$$

and substituting in the equation for Q_s, we see that

$$Q_s = (1 - R_1) + (1 - R_2) - (1 - R_1)(1 - R_2)$$
$$= 1 - R_1 + 1 - R_2 - 1 + R_1 + R_2 - R_1 R_2$$
$$= 1 - R_1 R_2$$
$$= 1 - R_s \tag{18.21}$$

which is to be expected. In this case, the "event" described in law 2 is failure.

The unreliability expression for a system having more than two components is more complex than that above, but using $Q_s = 1 - R_s$, the reliability may first be determined using Eq. (18.19) and then the unreliability determined using Eq. (18.21).

6. Reliability of a Parallel System. The reliability of a parallel system or probability of survival of the system is the probability of at least *one* component surviving, since, provided that at least one component survives in a parallel system, the system will not fail.

If R_p is the system reliability and R_1, R_2 are the component reliabilities, then using law 2,

$$R_p = R_1 + R_2 - R_1R_2 \tag{18.22}$$

In this case, the "event" described in law 2 is survival. For a system having more than two components, the reliability expression is more complex, but as we shall see, an easy method of calculation is to determine the unreliability first and use Eq. (18.24) given below.

7. Unreliability of a Parallel System. The unreliability of a parallel system or probability of system failure is the probability of *all* components failing, since for a parallel system, even if only one component survives, the system has not failed.

If Q_p is the system unreliability and Q_1 and Q_2 are the component unreliabilities, then using law 1,

$$Q_p = Q_1Q_2 \tag{18.23}$$

In this case, the "event" described in law 1 is failure. This expression is called the *product law of unreliabilities* for parallel systems. For a system having more than two components, the additional unreliabilities are included in the product term.

Since $Q_1 = 1 - R_1$ and $Q_2 = 1 - R_2$,

$$
\begin{aligned}
Q_p &= (1 - R_1)(1 - R_2) \\
&= 1 - R_1 - R_2 + R_1R_2 \\
&= 1 - (R_1 + R_2 - R_1R_2) \\
&= 1 - R_p
\end{aligned}
\tag{18.24}
$$

which is to be expected. This equation is most useful for determining the system reliability R_p after having found the system unreliability Q_p. It is easier to compute Q_p than R_p in the first instance, because Q_p is contained in the product law of unreliabilities while the expression for R_p becomes increasingly complex as the number of subunits is increased.

8. Systems Containing Exponentially Failing Units. Whether or not a system made up of units which are failing exponentially behaves overall in an exponential fashion depends on the system type. It is found that series systems do behave exponentially and their reliability may be expressed in the familiar $e^{-\lambda t}$ form, whereas parallel systems do not and the form of their reliability equation depends on the number of subunits.

9. Series Systems. The reliability equation as shown above is

$$R_s = R_1R_2R_3 \cdots$$

where R_s is the system reliability and R_1, R_2, R_3, etc. are the subunit reliabilities.

If $R_1 = e^{-\lambda_1 t}$, $R_2 = e^{-\lambda_2 t}$, $R_3 = e^{-\lambda_3 t}$ etc., where λ_1, λ_2, λ_3, etc. are the respective failure rates of the subunits, then

$$R_s = e^{-\lambda_1 t} \, e^{-\lambda_2 t} \, e^{-\lambda_3 t}$$

$$= e^{-(\lambda_1 + \lambda_2 + \lambda_3)t} \tag{18.25}$$

which is, of course, of general exponential form, the system failure rate being the *sum* of the individual failure rates. The system failure rate clearly increases as the number of subunits is increased.

Since the overall behavior is exponential, that is, we can express the reliability R_s in the form

$$R_s = e^{-\lambda_s t}$$

where λ_s, the system failure rate, is given by

$$\lambda_s = \lambda_1 + \lambda_2 + \lambda_3 \tag{18.26}$$

then the MTBF for a series system m_s is equal to the reciprocal of the system failure rate:

$$m_s = \frac{1}{\lambda_s} = \frac{1}{\lambda_1 + \lambda_2 + \lambda_3} \tag{18.27}$$

This reciprocal equation applies only to components or systems whose reliability is expressible in exponential form, i.e., having constant failure rate.

For a series system containing n similar units of equal reliability,

System reliability

$$R_s = e^{-n\lambda t} \tag{18.28}$$

System failure rate

$$\lambda_s = n\lambda \tag{18.29}$$

System MTBF

$$m_s = \frac{1}{n\lambda} \tag{18.30}$$

10. Parallel Systems.   As was stated earlier, the form of the reliability expression for a parallel system depends on the number of subunits. For a simple two-unit system,

$$R_p = R_1 + R_2 - R_1 R_2 \tag{18.31}$$

For a three-unit system,

$$R_p = R_1 + R_2 + R_3 - R_1 R_2 - R_2 R_3 - R_1 R_3 + R_1 R_2 R_3 \tag{18.32}$$

And the expression becomes more complex as the number of units is increased.

If we write $R_1 = e^{-\lambda_1 t}$, $R_2 = e^{-\lambda_2 t}$, etc., as we did for series systems, then the system reliability for a two-unit system is given by

$$R_p = e^{-\lambda_1 t} + e^{-\lambda_2 t} - e^{-(\lambda_1 + \lambda_2)t} \tag{18.33}$$

For a three-unit system, by

$$R_p = e^{-\lambda_1 t} + e^{-\lambda_2 t} + e^{-\lambda_3 t}$$

$$+ e^{-(\lambda_1 + \lambda_2)t} + e^{-(\lambda_2 + \lambda_3)t}$$

$$+ e^{-(\lambda_1 + \lambda_3)t} + e^{-(\lambda_1 + \lambda_2 + \lambda_3)t} \tag{18.34}$$

Clearly, these equations are not of simple exponential form, and we cannot express the overall system reliability in the form $e^{-\lambda_p t}$, where λ_p is a constant system failure rate. The MTBF may still be obtained by integration of the reliability expression, as was done in the series case, since, as was shown earlier, this method of finding the MTBF does not rely on the expression being of exponential form. However, for a parallel system, the MTBF is not the reciprocal of the system failure rate, but depends on the number of subunits.

It can be shown that the MTBF m_p for a two-unit system is given by

$$m_p = \frac{1}{\lambda_1} + \frac{1}{\lambda_2} - \frac{1}{\lambda_1 + \lambda_2} \tag{18.35}$$

For a three-unit system, it is given by

$$m_p = \frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} - \frac{1}{\lambda_1 + \lambda_2} - \frac{1}{\lambda_2 + \lambda_3} - \frac{1}{\lambda_1 + \lambda_3} + \frac{1}{\lambda_1 + \lambda_2 + \lambda_3} \tag{18.36}$$

where λ_1, λ_2, and λ_3 are the unit failure rates, respectively, and for an n-unit system, each unit having the same

$$m_p = \frac{1}{\lambda} + \frac{1}{2\lambda} + \frac{1}{3\lambda} + \cdots + \frac{1}{n\lambda} \tag{18.37}$$

The system failure rate for a parallel system is not constant but is time-dependent.

11. Series and Parallel Systems: Reliability Compared.

In both system types, the overall reliability is dependent on the number of subunits. To illustrate the effect of the series and parallel connection of the system subunits, consider again the reliability expressions. For a two-unit system having unit reliabilities R_1, R_2,

$$\text{Series system reliability} = R_1 R_2$$

$$\text{Parallel system reliability} = R_1 + R_2 - R_1 R_2$$

Bearing in mind that reliability can never exceed unity, examination of the two expressions shows that

$$R_1 + R_2 - R_1 R_2 \geq R_1 R_2$$

the two being equal only when $R_1 = R_2 = 1$, at all other values of R_1 and R_2 the left-hand side being greater than the right-hand side. This means, then, that the reliability for a two-unit parallel system is greater than that for a two-unit series system except when the subunits have equal reliabilities equal to unity. In this case,

series and parallel systems will have equal reliabilities. A similar observation may be made concerning reliabilities of three-unit systems:

Series system reliability = $R_1 R_2 R_3$

Parallel system reliability = $R_1 + R_2 + R_3 - R_1 R_2 - R_2 R_3 - R_1 R_3 + R_1 R_2 R_3$

Clearly the first six terms of the parallel reliability expression must have an overall value of zero or greater than zero, since R_1, R_2, or R_3 must all be unity or less than unity. The seventh term is, in fact, the series reliability, so we have

Parallel reliability = (terms ≥ 0) + (series reliability)

which is obviously greater than or equal to the series reliability. It can be shown that for any number of units the parallel reliability is greater than or equal to the series reliability. Physically, this is logical, since for a parallel system, all units except one can fail before system failure, whereas for a series system, only one unit needs to fail for system failure. Typical sets of reliability curves for one-, two-, and three-unit systems in both connection modes are shown in Fig. 18.7.

Notice that for a series system, the reliability at any time t_s is reduced as the number of subunits is increased. The reliability curve is always exponential regardless of the number of subunits. For a parallel system, the reliability at any time t_p is increased as the number of subunits is increased. For systems containing more than one unit, the reliability curve ceases to be exponential, as was explained earlier.

This improved reliability for parallel-connected systems is used in schemes to reduce risk of failure, such as standby systems, in which several equal units are allowed to stand idle ready to take over should the operational unit fail. Such schemes are said to employ *parallel redundancy.*

SUMMARY OF RELEVANT FORMULAS

For a constant proportional failure rate,

$$R = e^{-\lambda t} \tag{18.38}$$

$$R + Q = 1 \tag{18.39}$$

$$Q = 1 - e^{-\lambda t} \tag{18.40}$$

$$N_s = N_0 e^{-\lambda t} \tag{18.41}$$

$$N_f = N_0(1 - e^{-\lambda t}) \tag{18.42}$$

$$m = \frac{1}{\lambda} \tag{18.43}$$

$$R = e^{-t/m} \tag{18.44}$$

$$Q = 1 - e^{-t/m} \tag{18.45}$$

$$N_s = N_0 e^{-t/m} \tag{18.46}$$

$$N_f = N_0(1 - e^{-t/m}) \tag{18.47}$$
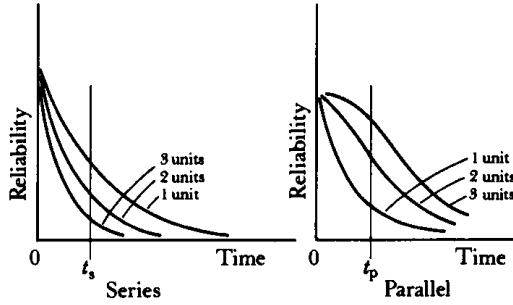
FIGURE 18.7 Reliability curves for series and parallel systems.

where R = reliability
   Q = unreliability
   λ = proportional failure rate (i.e., failure rate expressed as a proportion of N_0)
   N_s = number of live components (survivors)
   N_f = number of dead components (failures)
   N_0 = initial number of live components
   m = mean time between (chance) failures
   t = time

For a test to determine MTBF m,

$$m = \frac{\text{total survival hours}}{\text{number of failures}} \tag{18.48}$$

For a normally distributed variable x, the standard deviation σ is given by

$$\sigma = \sqrt{\frac{\Sigma(x - x_m)^2}{n}} \tag{18.49}$$

where x_m is the mean value of n observations of x.
   In a test to determine mean wearout life M,

$$M = \frac{\text{sum of lives}}{\text{number of components}} \tag{18.50}$$

For an exponentially distributed variable,

$$\text{Upper confidence limit} = \frac{2nm}{\chi^2_{1-\alpha/2,n}} \tag{18.51}$$

$$\text{Lower confidence limit} = \frac{2nm}{\chi^2_{\alpha/2,n}} \tag{18.52}$$

at a level of confidence given by $100(1 - \alpha)$ percent, where n denotes number of failures, m denotes an estimate of the mean value of the variable, and χ^2 denotes the value of chi-squares (given in tables) for values of n and α or $(1 - \alpha/2)$.

For a component or unit which forms part of a system

$$\frac{m_c}{m_s} = \frac{t_c}{t_s} \qquad (18.53)$$

$$m_s = \frac{m_c}{d} \qquad (18.54)$$

where m_c = component MTBF in component operating hours
m_s = component MTBF in system operating hours
t_c = component operating hours
t_s = system operating hours
d = duty cycle ($d = t_c/t_s$)

The probability of both events x and y occurring P_{xy} is given by

$$P_{xy} = P_x P_y \qquad (18.55)$$

And the probability of either event x or event y occurring P_{x+y} is given by

$$P_{x+y} + P_x + P_y - P_x P_y \qquad (18.56)$$

where P_x = the probability of x occurring
P_y = the probability of y occurring

The following equations refer to series and parallel systems. The symbols used are as above for reliability, unreliability, failure rate, etc., with the addition of appropriate subscripts as follows:

Subscript s denotes series
Subscript p denotes parallel
Subscripts 1, 2, 3, etc. denote components or subunits 1, 2, 3, etc.

$$R_s = R_1 R_2 \cdots \qquad (18.57)$$

$$Q_s = Q_1 + Q_2 - Q_1 Q_2 \qquad (18.58)$$

$$Q_s = 1 - R_s \qquad (18.59)$$

$$R_p = R_1 + R_2 - R_1 R_2 \qquad (18.60)$$

$$Q_p = Q_1 Q_2 \qquad (18.61)$$

$$Q_p = 1 - R_p \qquad (18.62)$$

$$R_s = e^{-(\lambda_1 + \lambda_2 + \lambda_3 + \cdots)t} \qquad (18.63)$$

$$\lambda_s = \lambda_1 + \lambda_2 + \lambda_3 + \cdots \qquad (18.64)$$

$$m_s = \frac{1}{\lambda_s} \qquad (18.65)$$

$$R_s = e^{-n\lambda t} \qquad (18.66)$$

$$\lambda_s = n\lambda \tag{18.67}$$

$$m_s = \frac{1}{n\lambda} \tag{18.68}$$

$$R_p = R_1 + R_2 - R_1R_2 \tag{18.69}$$

$$R_p = R_1 + R_2 + R_3 - R_1R_2 - R_2R_3 - R_1R_3 + R_1R_2R_3 \tag{18.70}$$

$$R_p = e^{-\lambda_1 t} + e^{-\lambda_2 t} + e^{-(\lambda_1+\lambda_2)t} \tag{18.71}$$

$$R_p = e^{-\lambda_1 t} + e^{-\lambda_2 t} + e^{-\lambda_3 t} + e^{-(\lambda_1+\lambda_2)t} + e^{-(\lambda_2+\lambda_3)t}$$
$$+ e^{-(\lambda_1+\lambda_3)t} + e^{-(\lambda_1+\lambda_2+\lambda_3)t} \tag{18.72}$$

$$m_p = \frac{1}{\lambda_1} + \frac{1}{\lambda_2} - \frac{1}{\lambda_1 + \lambda_2} \tag{18.73}$$

$$m_p = \frac{1}{\lambda_1} + \frac{1}{\lambda_3} + \frac{1}{\lambda_3} - \frac{1}{\lambda_1 + \lambda_2} - \frac{1}{\lambda_2 + \lambda_3} - \frac{1}{\lambda_1 + \lambda_3}$$
$$+ \frac{1}{\lambda_1 + \lambda_2 + \lambda_3} \tag{18.74}$$

$$m_p = \frac{1}{\lambda} + \frac{1}{2\lambda} + \frac{1}{3\lambda} + \cdots + \frac{1}{n\lambda} \tag{18.75}$$

where n in Eqs. (18.67), (18.68), and (18.75) denotes the number of components or subunits having equal failure rates.

For a system, the utilization factor U is given by

$$U = \frac{\text{operating time}}{\text{maintenance time} + \text{idle time} + \text{operating time}} \tag{18.76}$$

the availability (maximum utilization factor) A is given by

$$A = U_{\text{max}}$$
$$= \frac{\text{operating time}}{\text{minimum maintenance time} + \text{operating time}} \tag{18.77}$$

For any two sets of operating conditions denoted by x and m, respectively, the voltages V_x and V_m, temperatures t_x and t_m, and failure rates λ_x and λ_m are related by the equation

$$\lambda_x = \lambda_m \left(\frac{V_x}{V_m}\right)^n K^{t_x - t_m} \tag{18.78}$$

where n and K are constants over a limited range of conditions and may be determined by the equations

$$K = \left(\frac{\lambda_x}{\lambda_m}\right) \frac{1}{t_x - t_m} \tag{18.79}$$

for a constant voltage test, and

$$n = \frac{\ln (\lambda_x/\lambda_m)}{\ln (V_m/V_x)} \tag{18.80}$$

for a constant temperature test.

SYSTEMS ENGINEERING*

SYSTEMS ENGINEERING

The design of a complex interconnection of many elements (a system) to maximize an agreed-upon measure of system performance is the basis of *systems engineering.* Systems engineering, also referred to as *system engineering,* includes two parts: modeling, in which each element of the system and the criterion for measuring performance are described, and optimization, in which adjustable elements are set at values that give the best possible performance.

The systems approach can be applied to problems ranging from the very simple to those so complex that the human mind is unable to comprehend the reasons for system behavior. A simple problem is the scheduling of the preparation of a family meal. A much more complex system problem is the control of the timing of hundreds or thousands of traffic lights in a city.

The techniques of systems engineering have been applied to a tremendous range of current problems, from industrial automation to control of weapons and space vehicles.

Modeling refers to the determination of a quantitative picture of the important system characteristics. This model may be in the form of collected data or analytical studies, or it may be a representation of the important system characteristics in a laboratory or computer simulation of the actual system.

A specific and rather narrow systems problem is the control of traffic flow through a tunnel. The designer has to maximize the number of cars permitted per hour through the tunnel. To do this, the designer can vary the speed limit through the tunnel and control the number of cars within the tunnel (or the rate at which cars enter). Before a system design can be decided on, measurements are made of the way in which cars move under various speeds and degrees of congestion. These data are combined with the known behavior characteristics of a driver. On the basis of this quantitative information, a computer simulation of the system can be constructed from which the engineer can study the effects of different operating rules, the results of a breakdown in one lane of the tunnel, or the effect of a driver who stays below the minimum speed limit. See the section "Simulation" below.

The engineer attempts to design a system that is optimum according to a quantitative criterion which measures the quality of performance. For example, in the tunnel traffic problem, the entry and speed of cars are controlled in such a way that the tunnel handles the maximum possible number of cars per hour. In many situations, the criterion is probabilistic, and the engineer must optimize on the basis of probable behavior of the system. The routing of emergency vehicles through a network of city streets is a typical example. See the section "Optimization" below.

Three technological developments have enormously broadened the scope of problems amenable to the systems approach. The first of these is the understanding of the principles of automatic control (or the principles underlying automation). Systems can be designed in which the desired automatic control is realized by the appropriate choice of input signals and system configuration (including the use of feedback, with the actual system response compared automatically with the desired response and the error used to modify the response toward the desired value). In other words, technology has developed automatic goal-seeking systems to replace the human role. The automatic control of elevators in a busy office building is a familiar example. The second is the revolution in communications. Enormous quantities of data can be transmitted over great distances with nearly perfect fidelity. The third fundamental technological development has been the high-speed electronic digital computer. The computer lies at the core of modern systems engineering, since it permits both the modeling and the optimization of systems vastly more complex than the human being alone can handle.

SIMULATION

The development and use of computer models for the study of actual or postulated dynamic systems is known as *simulation.* The essential characteristic of simulation is the use of models for study and experimentation rather than the actual system modeled. In practice, it has come to mean the use of computer models because modern electronic computers are so much superior for most kinds of simulation that computer modeling dominates the field. *Systems,* as used in the definition, refers to an interrelated set of elements, components, or subsystems. *Dynamic systems* are specified because the study of static systems seldom justifies the sophistication inherent in computer simulation.

Postulated systems as well as *actual* ones are included in the definition because of the importance of simulation for testing hypotheses, as well as designs of systems not yet in existence. The *development* as well as *use* of models is included because, in the empirical approach to system simulation, a simplified simulation of a hypothesized model is used to check educated guesses and thus to develop a more sophisticated and more realistic simulation of the simuland. The *simuland* is that which is simulated, whether real or postulated.

Among the systems which have been simulated are aerospace systems; chemical and other industrial processes; structural dynamics; physiological and biological systems; automobile, ship, and submarine dynamics; social, ecological, political, and economic systems; corporations and small businesses; traffic and transportation systems; electrical, electronic, optical, and acoustic systems; electrical energy systems and other energy related systems; and learning, thinking, and problem-solving systems.

Mathematical modeling is a recognized and valuable adjunct, and usually a precursor, of computer simulation. Mathematical modeling does not necessarily pre-

cede simulation, however; sometimes the simuland is not well enough understood to permit rigorous mathematical description. In such cases, it is often possible to postulate a functional relationship of the elements of the simuland without specifying mathematically what that relationship is. This is the building-block approach, for which analog computers are particularly well suited. Parameters related to the function of the blocks can be adjusted until some functional criteria are met. Thus the mathematical model can be developed as the result of, rather than as a requirement for, simulation.

Analog computers, in which signals are continuous and are processed in parallel, were originally the most popular for simulation. Their modular design made it natural to retain the simulation-simuland correspondence, and their parallel operation gave them the speed required for real-time operation. The result was unsurpassed human-machine rapport. However, block-oriented digital simulation languages were developed which allow a pseudo-simulation-simuland correspondence, and digital computer speeds have increased to a degree that allows real-time simulation of all but very fast or very complex systems.

Hybrid simulation, in which both continuous and discrete signals are processed, both in parallel and serially, is the result of a desire to combine the speed and human-machine rapport of the analog computer with the precision, logic capability, and memory capacity of the digital computer. Hybrid simulation now makes possible the simulation of a new array of systems which require combinations of computer characteristics unavailable in either all-analog or all-digital computers.

SYSTEMS ANALYSIS

The application of mathematics to the study of systems is known as *systems analysis.* The term *operations research* is reserved for the study of part of a system. See the section "Operations Research" below.

The basic idea is that a mathematical model of the system under study is constructed, a mathematical analysis is done of the mathematical model, and the results of this analysis are applied to the original system.

A great deal of experience is needed to construct the mathematical model and to interpret the results of the analysis. The mathematical analysis, usually involving a computer, is seldom routine. The procedures are so complex that often an individual systems analysis is required for the systems analysis of a particular system. Often particular parts of the process are carried out by different people. Ideally, there should be an interaction between the mathematical analysis, the construction of the mathematical model, and the interpretation of the results.

Systems analysis is thus a part of applied mathematics. What makes the difference between this and conventional applied mathematics is that the systems studied often involve human beings. The presence of humans and the application of the results to human systems introduce a great deal of complication.

There are three principal difficulties in the application of conventional mathematical techniques: dimensionality, the presence of both "hard" and "soft" variables, and conflicting objectives. *Dimensionality* refers to the possibility that the description of a system may involve many state variables. This feature presents serious difficulties, but they may be circumvented by means of various techniques. Much more serious is the presence of hard and soft variables and various combinations. Thus, for example, in a study of a business, the number of employees is

a *hard variable,* while the quality of management is a *soft variable.* Hard and soft variables may be handled by the theory of fuzzy systems. In most situations, there are many objectives. In many cases, these objectives are in partial or complete conflict. Thus it is impossible to write down a single criterion for performance. Fortunately, these questions can be studied by means of simulation. Simulation allows various possibilities to be studied in electronic time without disturbing the system or the people affected by the system. See discussion above.

OPTIMIZATION

In its most general meaning, *optimization* covers the efforts and processes of making a decision, a design, or a system as perfect, effective, or functional as possible. Formal optimization theory encompasses the specific methodology, techniques, and procedures used to decide on the one specific solution in a defined set of possible alternatives that will best satisfy a selected criterion. Because of this decision-making function, the term *optimization* is often used in conjunction with procedures which more appropriately belong in the more general domain of decision theory. Strictly speaking, formal optimization techniques can be applied only to a certain class of decision problem known as *decision making under certainty.*

1. Concepts and Terminology. Conceptually, the formulation and solution of an optimization problem involves the establishment of an evaluation criterion based on the objectives of the optimization problem, followed by determination of the optimum values of the controllable or independent parameters that will best satisfy the evaluation criterion. This determination is accomplished either objectively or by analytical manipulation of the so-called criterion function, which relates the effects of the independent parameters on the dependent evaluation criterion parameter. In most optimization problems, there are a number of conflicting criteria and a compromise must be reached by a tradeoff process which makes relative value judgments among the conflicting criteria. Additional practical considerations encountered in most optimization problems include so-called functional and regional constraints on the parameters. The former represents physical or functional interrelationships which exist among the independent parameters (that is, if one is changed, it causes some changes in the others); the latter limits the range over which the independent parameters can be varied.

Optimization techniques work only for a specific system or configuration which has been described to a point where all criteria and parameters are defined within the system and are isolated and independent from other parameters outside the boundaries of the defined system. Formal optimization is not a substitute for creativity in that it depends on a clear definition of the system to be optimized, and even though an optimum solution for a given system configuration is obtained, this does not guarantee that a better solution is not available.

2. Problem Formulation. Definition is a critical part of problem formulation and consists of

1. A description of the system configuration to be optimized, including definition of system boundaries to an extent that system parameters become isolated and independent of external parameters.

2. Definition of a single, preferably quantitative, parameter which will serve as the overall evaluation criterion for the specific optimization problem. This is the dependent parameter (it depends on the choice of the optimum solution) which measures how well the solution satisfies the desired objectives of the problem and this is the parameter which will be maximized or minimized to satisfy the objectives.

3. Definition of controlled or independent parameters that will have an effect on the criterion. These are the parameters whose values determine the value of the criterion parameter, and these should include all controllable parameters which influence the criterion and are within the boundaries of the defined problem.

The most critical aspect of formulating a formal optimization problem is the establishment of a satisfactory criterion function that describes the behavior of the evaluation criterion as a function of the independent parameters. The *criterion function,* often also referred to as the *pay-off* or *objective function,* usually takes the form of a penalty or cost function (which attempts are made to minimize) or a merit, benefit, or profit function (which attempts are made to maximize by choosing the optimum values of the independent variables). In order to apply formal solution methods, the criterion function should be expressed graphically or analytically.

Physical principles of operation which govern the relationship among the various independent parameters of the problem represent functional constraints. Most optimization problems involve practical limits on the range over which each parameter or function of the parameters can be varied. They represent the regional constraints.

3. Solution of Optimization Problems.   The reliability and the sophistication of the solutions of optimization problems increase as the problem becomes better defined. In the early stages of optimization of a complex problem, the considerations may be primarily objective judgments based on the optimizer's judgments of the relevant parameters. As alternative subsystem and system configurations evolve through a process of conceptualization, analysis and evaluation, and elimination of undesirable system configurations, a clearer configuration of parts of the problem can be defined, together with the relevant criteria, parameters, and constraints. Procedures for defining competing systems or alternative strategies and formulating the problem in order to apply formal optimization techniques are also common to operations research, systems analysis, and systems design.

OPERATIONS RESEARCH

The application of scientific methods and techniques to decision-making problems is known as *operations research.* A decision-making problem occurs where there are two or more alternative courses of action, each of which leads to a different and sometimes unknown end result. Operations research is also used to maximize the utility of limited resources. The objective is to select the best alternative, that is, the one leading to the best result.

To put these definitions into perspective, the following analogy might be used. In mathematics, when solving a set of simultaneous linear equations, one states that if there are seven unknowns, there must be seven equations. If they are independent and consistent and if it exists, a unique solution to the problem is found. In oper-

ations research there are figuratively "seven unknowns and four equations." There may exist a solution space with many feasible solutions which satisfy the equations. Operations research is concerned with establishing the best solution. To do so, some measure of merit, some objective function, must be prescribed.

In the current lexicon there are several terms associated with the subject matter of this program: *operations research, management science, systems analysis, operations analysis,* and so forth. While there are subtle differences and distinctions, the terms can be considered nearly synonymous.

1. Methodology. The success of operations research, where there has been success, has been the result of the following six simply stated rules: (1) formulate the problem, (2) construct a model of the system, (3) select a solution technique, (4) obtain a solution to the problem, (5) establish controls over the system, and (6) implement the solution.

The first statement of the problem is usually vague and inaccurate. It may be a cataloging of observable effects. It is necessary to identify the decision maker, the alternatives, goals, and constraints, and the parameters of the system. A statement of the problem properly contains four basic elements that, if correctly identified and articulated, greatly ease the model formulation. These elements can be combined in the following general form: "Given (the system description), the problem is to optimize (the objective function), by choice of the (decision variable), subject to a set of (constraints and restrictions)."

In modeling the system, one usually relies on mathematics, although graphical and analog models are also useful. It is important, however, that the model suggest the solution technique, and not the other way around.

With the first solution obtained, it is often evident that the model and the problem statement must be modified, and the sequence of problem-model-technique-solution-problem may have to be repeated several times. The controls are established by performing sensitivity analysis on the parameters. This also indicates the areas in which the data-collecting effort should be made.

Implementation is perhaps of least interest to the theorists, but in reality it is the most important step. If direct action is not taken to implement the solution, the whole effort may end as a dust-collecting report on a shelf.

2. Mathematical Programming. Probably the one technique most associated with operations research is linear programming. The basic problem that can be modeled by linear programming is the use of limited resources to meet demands for the output of these resources. This type of problem is found mainly in production systems, but it is not limited to this area.

3. Stochastic Processes. A large class of operations research methods and applications deals with *stochastic processes.* These can be defined as processes in which one or more of the variables take on values according to some, perhaps unknown, probability distribution. These are referred to as *random variables,* and it takes only one to make the process stochastic.

In contrast to the mathematical programming methods and applications, there are not many optimization techniques. The techniques used tend to be more diagnostic than prognostic; that is, they can be used to describe the "health" of a system, but not necessarily how to "cure" it.

4. Scope of Application.   There are numerous areas where operations research has been applied. The following list is not intended to be all-inclusive, but is mainly to illustrate the scope of applications: optimal depreciation strategies; communication network design; computer network design; simulation of computer time-sharing systems; water resource project selection; demand forecasting; bidding models for offshore oil leases; production planning; classroom size mix to meet student demand; optimizing waste treatment plants; risk analysis in capital budgeting; electric utility fuel management; optimal staffing of medical facilities; feedlot optimization; minimizing waste in the steel industry; optimal design of natural-gas pipelines; economic inventory levels; optimal marketing-price strategies; project management with CPM/PERT/GERT; air-traffic-control simulations; optimal strategies in sports; optimal testing plans for reliability; optimal space trajectories.

SAFETY ENGINEERING*

Safety in any environment in which people work is the responsibility of the person in charge of the facility. This person may have various titles in different industries. The generic title, however, is *plant engineer*. This title is used throughout this section, since it applies to essentially every industry known today.

SAFETY

A well-organized plant safety program should encompass all phases of the plant environment and operations. From the standpoint of this handbook, we are primarily concerned with safety controls and devices in the environment and only make reference to supervisory controls of the safety program.

Safeguarding against industrial hazards has always been one of the principal tasks of the plant engineer. Production depends on the ability to maintain a continuous flow of materials without the interruptions caused by accidents. Engineering controls and safeguards serve to protect inexperienced workers or those workers who are distracted or who suffer from fatigue.

The basic elements of a good safety program include

1. Management leadership
2. Assignment of responsibility
3. Maintenance of safe working environment
4. Training program
5. Record system
6. Medical follow-through

The plant engineer is basically concerned with items 1, 2, and 3, but items 4, 5, and 6 are also important.

LEGAL ASPECTS OF SAFETY

The early legal aspects of industrial safety were limited to laws which were developed in conjunction with workers' compensation acts. They were primarily aimed at providing for just compensation to the injured party, but they also established accident investigation procedures and some regulation of hazards.

The past two decades have seen a greater public demand for safety in the workplace. With it came a proliferation of federal laws and regulations which include the following.

1. Laws Applicable to Industrial Plants

1. Environmental Protection Agency

- Toxic Substances Control Act of 1976
- Marine, Protection, Research and Sanctuaries Act of 1972
- Safe Drinking Water Act of 1974
- Water Pollution Control Act of 1972
- Clean Air Act of 1970
- National Environment Policy Act of 1969
- Atomic Energy Act of 1954
- Clean Water Act
- Endangered Species Act of 1973
- Energy Supply and Environmental Coordination Act
- Fish and Wildlife Coordination Act

2. Solid Waste Disposal Act of 1965

3. Occupational Safety and Health Act of 1971

4. Department of Transportation

- Transportation Safety Act
- Hazardous Materials Transportation Act of 1974
- Ports and Waterways Safety Act of 1972

5. Resource Conservation and Recovery Act of 1976

6. Federal Insecticide, Fungicide, and Rodenticide Act of 1972

7. Consumer Product Safety Commission

- Federal Hazardous Substances Act
- Consumer Product Safety Act
- Poison Prevention Packaging Act

8. Mine Safety and Health Act

The Occupational Safety and Health Act will probably have a far greater effect on business and industry than the other legislation. The law requires all employers in the private sector of business to furnish their employees with a workplace free from those recognized hazards likely to cause physical harm or death. Standards are published or referenced in the act. This 800-page document is entitled *OSHA*

Safety and Health Standards (29 CFR 1910). It is available from the Superintendent of Documents, Washington, DC 20402, and is recommended as a reference for plant engineers.

The adoption of specific federal legislation has also led to an increase in the number of civil suits. The violation of a federal standard may become prima facie evidence of negligence. Such cases have been very costly.

Recent developments at OSHA have placed greater emphasis on the health aspects of worker's compensation. What started in specialized industries, such as coal mining and asbestos, is spreading to the chemical industry and others. Much more of the social responsibility for the well-being of the employee is being directed to the employer.

All employers must keep accurate records of work-related injuries, illnesses, and deaths. Any injury that involves medical treatment, loss of consciousness, restrictions of motion or work, or job transfer must be recorded. Required information must be logged and specific information regarding each case detailed. At the end of the calendar year, a summary of logged cases must be posted in each plant.

INSTRUMENTATION AND CONTROLS

Some of the most important advances in safety over the past two decades have been achieved in the area of instrumentation and controls. Standards of the Instrument Society of America (ISA) should be followed. The plant engineer is directly concerned with these devices, since they are involved in controlling the operations, materials flow, and environment of the plant. Instrumentation and controls fall into the following categories:

1. Temperature systems
2. Pressure systems
3. Flow systems
4. Analytical and testing systems (i.e., gas analysis, solids and dust analysis, reaction product tests, oxygen analysis, pollution instrumentation, electronic components, and system checkout equipment)
5. Weighing, feeding, and batching systems
6. Force and power systems
7. Motion and geometric systems (i.e., speed, velocity, vibrations)
8. Humidity and moisture determinations
9. Radiation systems
10. Electrical determinations
11. Communication systems
12. Automatic process controllers (i.e., computer process control, safety in instrumentation and control systems, electronic, pneumatic and hydraulic control devices)
13. Controlling elements (i.e., valves, actuators, electric motor drives and controls, mercury switches, etc.)

The importance of instrumentation and controls cannot be overestimated. From the standpoint of safety and environmental control, the following applications are dominant:

1. Detection of leaks in equipment
2. Survey of operating areas for escape of toxic materials
3. Detection of flammable or explosive mixtures in the atmosphere or process lines
4. Monitoring plant stacks and other areas for the accidental discharge of toxic gases, vapors, or smokes.
5. Analysis of waste streams for toxic or other objectionable material.
6. Control of waste-treatment or product-recovery facilities

Automatic process control, unlike manual control, provides continuous monitoring and corrective action. However, automatic controllers cannot react to new conditions, nor can they predict beyond the data programmed into them. Therefore, the ability to foresee an upset condition in a process and the capability of developing fail-safe actions may well determine whether a serious accident is averted.

Analytical and testing instrumentation provides the plant engineer with much needed information. For instance, decisions can be made in these areas:

1. Purity of raw materials entering a process. Contaminants may be hazardous.
2. Process control by automatization of sampling.
3. Process troubleshooting by continuous analysis to prevent process upsets.
4. Determination of product quality to provide an impetus to a product safety program.

PLANT ENGINEER'S FUNCTION WITHIN THE SAFETY COMMITTEE

A safety program is a management method to develop specific safety objectives, assign responsibility, and obtain desired results. The plant engineer should participate with the safety committee from the standpoint of developing adequate inspection and control procedures in order to maintain a safe work environment and control of safety devices. Safety devices may be anything from a pressure-relief valve to a machine guard.

1. Inspection and Maintenance.   In many plants an inspection committee is appointed. A safety inspection may be done in conjunction with the regular plant inspection and maintenance program. This is an opportunity for the plant engineer to check all safety devices. The safety inspection should focus attention on those items directly concerned with accident prevention. The plant engineering staff should have a working knowledge of safety standards if their inspections are to cover safety as well as normal wear and tear of equipment. A general plant inspection should include

1. All buildings and physical equipment
2. Inspection of new machinery before it is placed in operation
3. Inspection of walking and working surfaces and means of egress
4. Special equipment such as powered platforms, personnel lifts, and vehicle-mounted work platforms
5. Materials handling and storage facilities, elevators, cranes
6. Machinery, machinery guards, and electrical equipment
7. Compressed gas and air equipment
8. Special equipment such as pressure vessels, drums and furnaces, and welding, cutting, and brazing equipment
9. Hand and portable powered tools and other hand-held equipment
10. Environmental control, equipment, ventilation, and pollution controls for toxic and hazardous substances

A systematic method of inspection and maintenance is preferred. Some companies require *preventive* maintenance programs which call for the replacement of *critical* equipment periodically regardless of inspection results. Others require general inspections of the entire premises, annually.

Areas of the plant that have the potential of developing into catastrophic hazards require special inspection procedures. Items that may develop into a catastrophe include (1) structural failure, (2) fire or explosion, and (3) release of hazardous gases or vapors.

Some elements of a plant may require inspection and maintenance on the basis of federal, state, or local laws. Items such as elevators, boilers, and unfired pressure vessels are in this category. This equipment may require the use of specially trained and licensed inspectors either from a governmental agency or an insurance carrier.

A careful record should be kept of all inspections and recommendations. This is particularly important in the event that an accident occurs at this location which results in litigation. Some companies assign inspection tasks to maintenance personnel, electricians, and others who are charged with the job of repairing the equipment. Supervisors should continuously examine their own work areas to make sure that tools, machinery, and other types of equipment are safe to handle.

Inspection methods should be established for all new equipment and processes. Nothing should be placed into operation until all safeguards have been checked and operations evaluated by the plant engineer. Safe operating instructions should be given to all workers concerned.

2. Inspection Procedures.   The following inspection criteria should be applied:

1. Inspectors must be familiar with the company's safety and health policies as well as the particular laws and regulations that pertain. Frequently, these regulations are only minimum requirements and it may be necessary to exceed them to secure adequate safety.
2. Inspectors should have available an analysis of all accidents that have occurred in the plant within the past year.
3. The inspectors should utilize all aids available, including inspection checklists, report forms, and other pertinent information.

An inspection report should be divided into three areas of interest:

1. A report on imminent hazards which require immediate corrective action.
2. A routine report on unsatisfactory (nonemergency) conditions which need corrective action.
3. A general report on the overall safety conditions of the facility.

ACCIDENT PREVENTION

The four basic steps for preventing accidents are as follows:

1. Elimination of the hazard
2. Control of the hazard
3. Training of personnel to be aware of and avoid the hazard
4. Utilization of personal protective equipment

The plant engineer is primarily concerned with steps 1 and 2 by ensuring a safe design for plant, physical facilities, and machinery.

BUILDING STRUCTURE

The plant engineer is directly concerned with the inspection and maintenance of the building structures. Slippery floors, stairs, runways, ramps, and other means of access are involved in many serious plant accidents. About one-fifth of the industrial injuries result from falls. Many of these can be avoided by careful design, construction, and maintenance. Applicable standards and codes should be applied.

1. Basic Rules. Housekeeping is a prime consideration in minimizing the hazard of falls. Some of the basic rules in this area are as follows:

1. All places of employment should be kept clean and orderly. Floors shall be maintained in a clean and dry condition with adequate drainage.
2. All passageways including aisles, ramps, and stairways should be kept clear and in good repair.
3. Floor loading should be kept well within prescribed limits.
4. All floor and wall openings should be guarded by standard railings or properly constructed closures.
5. Stairways should conform to acceptable standards.
6. Exits should be sufficient in number and properly located so that the building can be evacuated quickly in an emergency.
7. All ladders should conform to the applicable standards and codes and they should be properly maintained.
8. Ladders should not take the place of fixed stairways.

9. Workers should follow good practices in the use of ladders with regard to placement, support, angle between horizontal base and vertical plane of support, and proximity to other hazards (i.e., electrical).

10. Scaffolds, which are in effect elevated working platforms, should be designed with an adequate factor of safety and protection for the workers.

11. Scaffolds must be maintained, inspected, and guarded on all exposed sides. Metal scaffolds should not be constructed near electrical equipment.

2. Standards.   Refer to ANSI standards that apply to building structures.

MEANS OF EGRESS FOR INDUSTRIAL OCCUPANCIES

A *means of egress* is a continuous and unobstructed way to exit from any point in a building or structure to a public way. It includes vertical and horizontal ways of travel and intervening room spaces and other areas. The number of exit facilities required is specified in standards and codes, and it depends on the structure, occupants, and hazard exposures.

Inspections of means of egress have shown the following items at fault in such facilities:

1. Improper number of exits and locations

2. Poor illumination—lack of emergency lighting

3. Lack of directional signs

4. Poor housekeeping—obstructions

5. Improper and hazardous floor surfaces

6. Faulty operation of exit doors

1. Standard.   Refer to NFPA 101 1970, *Life Safety Code,* which applies to egress for industrial occupancies.

POWERED PLATFORMS, PERSONNEL LIFTS, AND VEHICLE-MOUNTED WORK PLATFORMS

Powered platforms, personnel lifts, and vehicle-mounted work platforms are means of elevating workers on suspended operated work platforms or personnel lifts. The platforms are used for exterior building maintenance work, the personnel lifts are used to lift workers to an elevated jobsite, and the vehicle-mounted work platforms are used to position personnel to an elevated work site. All these installations should be in conformance with the appropriate ANSI standards.

Every work platform should be tested and inspected frequently. New platforms should be tested before they are placed in service. Each installation should be inspected and tested at least every 12 months and should undergo a maintenance inspection and test every 30 days. Results of all inspections and tests should be logged, including the date, time, and inspector.

Items which require special inspections are as follows:

1. Governors
2. Initiating devices
3. Both independent braking systems
4. Interlocks and emergency electric devices
5. Electric systems
6. Emergency communication system

Maintenance is required specifically on all parts of the equipment related to safe operations. Broken or worn parts and electric devices should be replaced promptly. Gears, shafts, bearings, brakes, and hoisting drums should be maintained in proper alignment. Gears should be replaced when there is evidence of appreciable wear. All parts should be kept free from dirt. Wires or ropes should be replaced when they are damaged or in a deteriorated condition. Guardrails and toeboards are required for working platforms. Load-rating plates must be conspicuously posted and adhered to.

The personnel lifts require frequent inspections. Their use should be limited to *personnel.* They should not be used to lift construction materials. The inspections should include but are not limited to the following items:

Steps	Drive pulley
Rails and supports	Electrical systems
Belt and belt tension	Vibration—alignment
Handholds	Brake systems
Floor landing	Warning lights
Limit switches	Pulleys—clearance

1. Standards.   Refer to ANSI standards that apply to platforms, personnel lifts, and vehicle-mounted work platforms.

VENTILATION

Ventilation systems protect the health and environment of plant workers by removing objectionable dusts, fumes, vapors, or gases. They are essential safety devices and must be properly installed and maintained. Ventilation systems can be divided into two primary groups, local systems and general systems.

1. Local Systems.   Local exhaust systems prevent the accumulation of toxic or flammable materials near the process unit. Due to the variety of work and the types of equipment, it is necessary to design hoods specifically developed to be as close to the operation as possible. Air-sampling devices are available to determine whether the atmosphere has been adequately cleared or constitutes a hazard. Safety standards may specify local exhaust ventilation for particular processes.

After contaminated air passes from the hood to exhaust ducts, it is sent through a cleaning system or to the outdoors. Such exhaust air must conform to the regulations of the EPA. Dusts may be collected by means of cyclone dust collectors or

electrostatic precipitators. Many types of air-cleaning devices are used. The selection is determined by the degree of the hazard associated with the dust. Gases or vapors may be removed from the waste air by absorption or adsorption processes which dissolve or react with the waste product chemically. Some gases and vapors are rendered harmless by passing them through a combustion chamber where they are changed to acceptable gases. In other cases condensers are utilized to liquefy toxic vapors for removal.

When ventilation is used to control potential exposure to workers it is adequate to reduce the concentration of contaminant so that the hazard is removed.

Local exhaust ventilation is used for a great many industrial operations such as anodizing, pickling, metal cleaning, and open-surface tank operations. There is a standard developed for this purpose. Other uses include spray booths; dip tanks; fume controls in electric welding and grinding operations; cast-iron machining; dust control in foundries; ventilation of internal combustion engines; and kitchen range hoods.

One of the basic maintenance items in the operation of local exhaust ventilation systems is to ensure that the flow of air is unobstructed. A minimum maintained velocity must exist in order to meet the health and safety requirements. Where flammable gases or vapors are removed, the electric equipment must conform to code requirements.

At intervals of not more than 3 months the hood and duct system should be inspected for evidence of corrosion, damage, or obstruction. In any event if the airflow is found to be less than required, it should be increased to that required.

2. General Systems.   General ventilation in the workplace contributes to the comfort and efficiency of the employees. It also serves to clear the air of hazardous contaminants and excessive heat or humidity. When local exhaust ventilation cannot be applied due to the many sources of vapor release, general ventilation is prescribed. The publication of the American Society of Heating, Refrigerating and Air Conditioning Engineers, *Handbook of Fundamentals*, defines the prescribed methods and requirements for the development of acceptable ventilation systems. The plant engineer should inspect this equipment frequently to make sure that it performs as required.

3. Standards.   Refer to ANSI standards pertaining to ventilation.

COMPRESSED GASES

Plant engineers should be aware of the safety requirements for compressed-gas handling and storage. Pressure-relief devices for gas cylinders, portable tanks, and cargo tanks should be installed and maintained in accordance with Compressed Gas Association (CGA) pamphlets S1.1-1963 and S1.2-1963.

MATERIALS HANDLING AND STORAGE

1. Powered Industrial Trucks.   Safety requirements for powered industrial trucks involve the following:

1. The plant engineer is directly concerned with the safety requirements relating to fire protection: design, maintenance, and use of fork trucks, tractors, platform lift trucks, motorized hand trucks, and other specialized industrial trucks.

2. All new industrial trucks should meet the requirements of ANSI B56.1-1969, *Powered Industrial Trucks.*

3. In locations used for the storage of hazardous liquids or liquefied or compressed gases, only approved trucks designated as DS, ES, GS, or LPS may be used. In areas containing combustible dusts approved EX or ES trucks may be used.

4. Any power-operated industrial truck not in safe operating condition should be removed from service.

2. Overhead and Gantry Cranes.   Safety requirements for overhead and gantry cranes include

1. All new and existing overhead and gantry cranes should meet the design specifications of ANSI B30.2.0-1976, *Safety Standard for Overhead and Gantry Cranes* (*Top Running Bridge, Multiple Girder*).

2. Exposed moving parts such as gears, set screws, projecting keys, chains, chain sprockets, and reciprocating components which might constitute a hazard under normal operating conditions should be guarded.

3. Both holding brakes and control brakes should be tested to determine whether they are within required tolerances.

4. Brakes on trolleys and bridges should be examined and adjusted when necessary.

5. Inspections fall into two classes: *frequent inspections,* which fall into daily or monthly intervals, and *periodic inspections,* which fall into 1- to 12-month intervals. *Frequent inspections* should include (*a*) all functional operating mechanisms for maladjustments interfering with daily operations, (*b*) deterioration or leakage in air or hydraulic systems, (*c*) hooks and hoist chains and ropes including end connections, (*d*) all functional operating mechanisms. *Periodic inspections* should include (*a*) all items listed under frequent inspections, (*b*) deformed, cracked, or corroded members, (*c*) worn, cracked, or distorted parts, (*d*) excessive wear on brake system, (*e*) improper performance of power plants, (*f*) excessive wear of chain drive sprockets and excessive chain stretch, (*g*) defective electrical apparatus.

6. Preventive maintenance based on crane manufacturers' recommendations should be established.

7. Any unsafe conditions—including those in ropes—disclosed in the inspection requirements should be corrected before operation of the crane is allowed.

3. Crawler, Locomotive, and Truck Cranes.   Safety requirements for crawler, locomotive, and truck cranes include

1. All new and existing locomotive and truck cranes should meet the design specifications of ANSI B30.5-1968, *Safety Code for Crawler, Locomotive, and Truck Cranes.*

2. Load ratings should not exceed the stipulated percentages for cranes with indicated types of mountings.

3. Inspections are frequent and periodic as cited under overhead gantry cranes.

4. Derricks.   All new derricks constructed after August 31, 1971 should meet the design specifications of ANSI B30.6-1977, *Safety Code for Derricks.* Those derricks constructed prior to that date should be modified to conform. The plant engineer should check the standard to assure compliance.

Inspection procedures are similar to those required for cranes under frequent or periodic inspections. Derricks not in regular use for a period of 6 months should be given a complete inspection before use.

Maintenance and repair is an important part of the plant engineer's assignment and should include the following steps:

1. Any unsafe conditions shown in the inspection should be corrected.
2. Adjustments should be made to assure correct functioning of all components.
3. Ropes and wires should be thoroughly inspected and replaced or repaired.
4. No derrick should be loaded beyond the rated load.

5. Helicopter Cranes. Helicopter cranes should comply with the applicable regulations of the Federal Aviation Administration. The weight of an external load should not exceed the helicopter manufacturer's rating. All equipment should be checked frequently. The cargo hooks should be tested prior to each day's operation to determine that the release functions properly, both electrically and mechanically.

6. Slings. All slings used in conjunction with materials handling equipment made from alloy-steel chains, wire rope, metal mesh, natural or synthetic fiber rope, and synthetic web (nylon and polypropylene) should conform to the applicable standards. They should be inspected daily for damage or defects.

7. Standards. Refer to ANSI standards pertaining to materials handling and storage.

MACHINERY AND MACHINE GUARDING

One or more methods of machine guarding should be provided to protect the operator and other employees in the machine area from hazards such as those created by point of operation, ingoing nip points, rotating parts, flying chips, and sparks. Examples are barrier guards, two-hand tripping devices, electronic safety devices, etc.

1. General. The following machines usually require point-of-operation guarding:

Guillotine cutters
Shears
Alligator shears
Power presses
Milling machines
Pointer saws
Jointers
Forming rolls and calenders
Revolving drums, barrels, and containers
Exposed blades

2. Woodworking Machinery.   Woodworking machinery includes automatic cut-off saws, circular saws, hinged saws, revolving double arbor saws, hand-fed ripsaws, hand-fed crosscut saws, circular resaws, swing-cut saws, radial saws, bandsaws, jointers, tenoning machines, boring and mortising machines, wood shapers, planers, lathes, sanding machines, guillotine veneer cutters, and miscellaneous woodworking machines. These machines should be properly guarded in accordance with appropriate ANSI standards. All belts, pulleys, gears, shafts, and moving parts should be guarded in accordance with requirements.

3. Abrasive Wheel Machinery.   Abrasive wheel machinery includes cylindrical grinders, surface grinders, swing frame grinders, and automatic snagging machines. The guard design and specifications should be in accordance with appropriate ANSI standards.

MILLS AND CALENDERS IN THE RUBBER AND PLASTICS INDUSTRIES

All new and existing installations of mills and calenders should comply with the OSHA requirements and pertinent standards.

Mill safety controls consisting of safety trip controls, pressure-sensitive body bars, safety trip rods, safety trip wire cables, or wire center cords should be installed in all mills either singly or in combination. Fixed guards should be used where applicable.

Calender safety controls should be provided on all calenders within reach of the operator and the bite. Calenders and mills should be installed so that persons cannot normally reach over or under, or come in contact with, the roll bite. All trip and emergency switches should require manual resetting. A suitable alarm should be provided in conjunction with the safety devices.

MECHANICAL POWER PRESSES

All mechanical power presses (excluding pneumatic power presses, bulldozers, hot-banding and hot-metal presses, forging presses, hammers, riveting machines, and similar types of fastener applicators) should conform to the OSHA regulations and applicable standards. Mechanical power presses require particular emphasis as they are involved in many accidents. The machine components should be designed, secured, and covered to minimize hazards. Safeguards should include

1. Brakes or blocks capable of stopping the motion of the slide
2. Foot pedals, protected to prevent unintended operation
3. Hand-operated levers to prevent premature or accidental tripping
4. Two-hand trips to have individual operator hand controls arranged to require the use of both hands to trip the press
5. Air-controlling equipment to be protected against foreign material
6. Brake monitoring systems installed on the press to indicate when the performance of the braking system has deteriorated

7. Point-of-operation guards installed and used in accordance with the standard's requirements

8. A regular program of periodic and regular inspections on a weekly basis of power presses

9. Reports of injuries to employees operating mechanical power presses required by OSHA, with an analysis of each accident on a 30-day basis

FORGING MACHINES

The safety requirements apply to use of lead casts and other uses of lead in the forge or die shop. All equipment should comply with OSHA requirements and the applicable standards. This should include

1. Thermostatic control of heating elements required for melting lead

2. Industrial hygiene and personal protective equipment required due to the toxicity and heat condition of the lead

3. All presses and hammers controlled by operators who are protected by required guards

4. Devices included to lock out the power when dies are being changed

MECHANICAL POWER TRANSMISSION

This section pertains to mechanical power transmission equipment (see "Standards"), with the exception of small belts operating at a reduced speed and reduced requirements for the textile industry to prevent accumulation of combustible lint. Standard guards are usually secured with the following materials: expanded metal, perforated or solid sheet metal on a frame of angle iron, or an iron pipe securely fastened to the floor or frame of the machine. The standard requirements include guards for the following:

1. Flywheels located 7 ft or less from the floor

2. Cranks and connecting rods when exposed to contact

3. Tail rods or extension piston rods

4. Shafting—horizontal, vertical, and inclined

5. Power transmission apparatus and pulleys

6. Belts, ropes, chain drums, gears, sprockets, and chains

7. Shafting, collars, couplings, belt shifters, clutches, perches, or fasteners

Periodic inspection is required of all power transmission equipment. Intervals not exceeding 60 days are recommended.

1. Standards.   Refer to ANSI standards that apply to mechanical power transmission.

HAND AND PORTABLE POWERED TOOLS AND OTHER HAND-HELD EQUIPMENT

Portable tools must be equipped with adequate guards and should comply with OSHA and the appropriate standards. This includes

1. Portable circular saws, saber, scroll, and jigsaws, portable belt sanding machines, portable abrasive wheels, explosive actuated fastening tools, power lawnmowers, jacks.
2. Warning instructions should be supplied with the equipment. All equipment should be inspected periodically.

1. **Standards.**   Refer to ANSI standards that pertain to hand-held equipment.

WELDING, CUTTING, AND BRAZING

Most welding and cutting operations are mobile and are generally used for construction, demolition, maintenance, and repair. Production-line welding and cutting equipment is permanently installed; the hazards can be controlled through proper design and operation. The oxygen and acetylene used for welding and cutting must be handled and stored in accordance with the pertinent standards.

Compressed-gas cylinders should be handled with care. The fusible safety plugs in acetylene cylinders melt at about 212°F (100°C). This is an obvious fire hazard if the cylinder is subjected to heat. Oxygen, on the other hand, will react strongly with oils or other hydrocarbons upon contact.

Regulators and pressure gauges should be used on the appropriate cylinders. Leaking cylinders must be removed from the building and taken away from sources of ignition. All equipment should be kept free of oily or greasy substances.

Piping systems should be tested and proved gas-tight at 1½ times the maximum operating pressure. Service piping systems should be protected by pressure-relief devices discharging upward to a safe location. Backflow protection should be provided by an approved device that will prevent oxygen from flowing into the fuel system.

Acetylene generators should be of approved construction and clearly marked with the maximum rate of acetylene production and the pressure limitations. Relief valves should be regularly operated to ensure proper functioning. Storage of all chemicals should conform to the requirements of the standards.

ARC-WELDING AND CUTTING EQUIPMENT

Arc-welding apparatus should comply with the requirements of the National Electrical Manufacturers, *Standard for Electric Arc-Welding Apparatus,* NEMAEW-1 1962, and for ANSI/UL551-1976. *Safety Standards for Transformer-Type-Arc-Welding Machines.* The design requirements include the following items of concern to the plant engineer:

1. Input power terminals, tip change devices, and live metal parts should be completely enclosed.
2. Terminals for welding leads should be protected from accidental electrical contact by personnel or by metal objects. The frame or case of the welding machine should be grounded as specified.
3. Printed rules and instructions concerning operation of the equipment supplied by the manufacturer should be strictly followed.
4. All parts of the equipment should be frequently inspected. Cables with damaged insulation or exposed bare conductors should be replaced.

RESISTANCE-WELDING EQUIPMENT

All resistance-welding equipment should be installed in accordance with article 630D of the *NEC*.* The following items of particular interest to the plant engineer are cited:

1. Controls on all automatic or air and hydraulic clamps should be arranged or guarded to prevent accidental actuation.
2. All doors and access and control panels should be kept locked and interlocked to prevent access by unauthorized persons to live portions of the equipment.
3. All press-welding machine operations, where there is a possibility of the operator's fingers being under the point of operation, should be effectively guarded in a manner similar to that prescribed for punch-press operations.
4. Flash-welding equipment should be equipped with hoods to control flying flash and ventilation of fumes.
5. Combustible material must be removed from the welding area and the basic fire-prevention requirements followed in accordance with ANSI/NFPA 51B- 1977. Fire watchers are required wherever a fire problem exists. A welding permit system should be instituted to control hazardous exposures.
6. Welding operators should use all of the protective equipment specified in the standard, and ventilation should be used where required.

1. **Standards.**   Refer to ANSI standards for resistance-welding equipment.

SPECIAL INDUSTRIES

Plant engineers involved in the following special industries should consult the OSHA regulations for specific requirements pertinent to their industry.

1. Pulp, paper, and paperboard mills
2. Textiles

---

*NEC is a registered trademark of the National Fire Protection Association, Quincy, MA 02269.

3. Bakery equipment

4. Laundry machinery and operations

5. Sawmills

6. Pulpwood logging

7. Agricultural operations

8. Telecommunications

1. Standards.   Refer to ANSI special-industry standards.

ELECTRIC EQUIPMENT

All electric equipment in the plant should comply with OSHA regulations and the *National Electrical Code* (NFPA 70-1981).* It is particularly important for the plant engineer to be knowledgeable in the following sections:

1. Article 250, Grounding

2. Article 500, Hazardous (Classified) Locations

TOXIC AND HAZARDOUS SUBSTANCES

The exposure of employees to any of approximately 600 chemicals (see ''Standards'' at end of this section) should at no time exceed the ceiling value given for that material. To obtain compliance with this section of the OSHA regulations, it is necessary to develop administrative and engineering controls. When such controls are not feasible, protective equipment or other protective measures should be used.

Controls can be developed by environmental monitoring, personal monitoring, and employee observation. The plant engineer has an important part to play in the inspection and maintenance of the environmental monitoring equipment.

One method for controlling exposure is by local exhaust-ventilation and dust-collection systems. They should be constructed, installed, and maintained in accordance with ANSI Z9.2-1979, *Fundamentals Governing the Design and Operation of Local Exhaust Systems,* which is incorporated by reference in this section of the handbook.

Twenty-two materials require special care in handling since they are in the category of regulated substances which present a cancer hazard. These are listed below:

Asbestos	beta-Propiolactone
Coal-tar pitch volatiles	2-Acetylaminofluorene
4-Nitrobiphenyl	4-Dimethylaminoayobenzene
alpha-Naphthylamine	*N*-Nitrosodimethylamine

***National Electrical Code* is a registered trademark of the National Fire Protection Association, Quincy, MA 02269.

Methylchloromethyl ether	Vinyl chloride
3,3-Dichlorobenzidene	Inorganic arsenic
bis(Chloromethyl) ether	Benzene
beta-Naphthylamine	Coke-oven emissions
Benzidine	Cotton dust
4-Amino diphenyl	1,2-Dibromo-3-chloropropane
Ethylenimine	Acrylonitrile

Controls on these materials are very detailed since the materials must be contained in a closed-system operation. The operating area must be restricted to authorized personnel only. Warning signs and instructions should be posted.

CHAPTER 19

MEASUREMENTS IN ENGINEERING

This chapter is a selection of methods and techniques for measurements of basic physical variables and selected engineering parameters.

LENGTH MEASUREMENT

1. Engineer's Rule. These are made form-hardened and tempered steel marked off with high accuracy in length from about 10–30 cm with folding rules up to 60 cm. They are used for marking off, setting calipers and dividers, etc.

2. Vernier Caliper Gauge This is used for internal and external measurements. It has a long flat scale with a fixed jaw and a sliding jaw, with a scale, or cursor, sliding along the fixed scale and read in conjunction with it. Two scales are provided to allow measurement inside and outside of the jaws (Fig. 19.1).

3. Micrometers. Micrometers are used for the measurement of internal and external dimensions, particularly of cylindrical shape. Measurements are based on the advance of a precision screw (Fig. 19.2).

4. Gage Blocks. Gage blocks represent industrial dimension standards. They are small steel blocks with highly polished parallel surfaces. Gage blocks are available in a range of thicknesses that make it possible to stack them so than any required dimension can be obtained in increments. Table 19.1 gives sets of gage blocks.

The literature of manufacturers of gage blocks furnishes an excellent source of information on the measurement techniques that are employed in practice.[1]*

5. Measurement of Large Bores. The size of very large bores may be measured by means of a gauge rod of known length slightly less than the bore. The rod is placed in the bore and the "rock" noted. The bore can be determined from the amount of rock and the road length.

---

*Number is related to reference list.
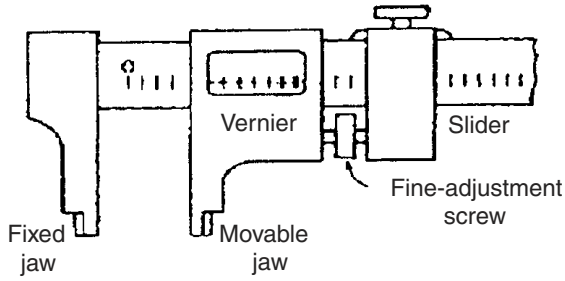
FIGURE 19.1   Vernier calliper.
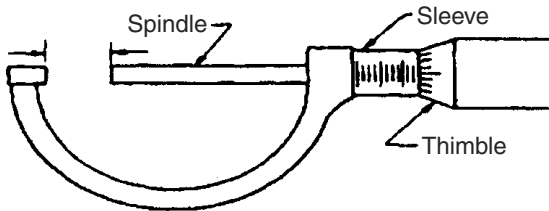
FIGURE 19.2   Micrometer calliper.

TABLE 19.1   Gage Block Sets

	No. blocks
Set M78	
1.01–1.49 mm in 0.01-mm steps	49
0.05–9.50 mm in 0.50-mm steps	19
10, 20, 30, 40, 50, 75, 100 mm	7
1.0025 mm	1
1.005 mm	1
1.0075 mm	1
Set M50	
1.01–1.09 mm in 0.01-mm steps	9
1.10–1.90 mm in 0.01-mm steps	9
1–25 mm in 1-mm steps	25
50, 75, 100 mm	3
1.0025, 1.0050, 1.0075 mm	3
0.005 mm	1

Source: BS 888.

$$\text{Bore diameter } D = L + \frac{a^2}{8L}$$

where (Fig. 19.3); L = gauge length, a = "rock"

Table 19.2 gives the accuracy of different methods of linear measurement.

6. Optical Methods. An optical method for measuring dimensions very accurately is based on the principle of light interference. The instument based on this principle is called an *interferometer* and is used for the calibration of gage blocks and other dimensional standards. Other optical instruments in wide use are various types of microscopes and telescopes, including the conventional surveyor's transit, which is employed for measurements of large distances.

ANGLE MEASUREMENT

In chapter 3 (section on trigonometry), circular and angular measure of a plane angle is defined. Two engineering methods are given here.

1. Combination angle slip gauges. Precision angle blocks are available with faces inclined to one another at a particular angle accurate to one second of arc. The gauges may be wrung together as with slip gauges, and angles may be added or subtracted to give the required angle (see Fig. 19.4).
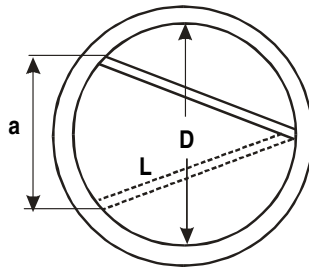
FIGURE 19.3 Large bore sketch.

TABLE 19.2 Accuracy of Linear Measurement

Instrument	Use	Accuracy (mm)
Steel rule	Directly	± 0.25
Vernier calipers	External	± 0.03
	Internal	± 0.05
25-mm micrometer	Directly	± 0.007
	Preset to gage blocks	± 0.005
Dial gauge	Over complete range	± 0.003–0.03
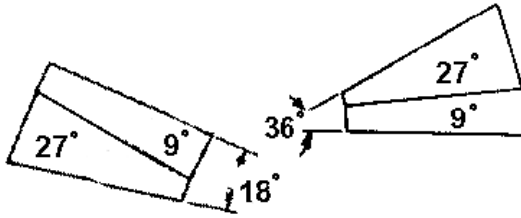
FIGURE 19.4 Examples of angle slip gauges.

Typical 13-block set includes:
Degrees: 1, 3, 9, 9, 27, 41
Minutes: 1, 3, 9, 27
Seconds: 3, 9, 27
Plus 1 square block.

2. Measurement of Angle of Tapered Bores. The method of measuring the angle of internal and external bore tapers is shown in Fig. 19.5 using precision balls, rollers and slip gauges.

STRAIN MEASUREMENT

An essential requirement of engineering design is the accurate determination of stresses and strains in components under working conditions. The change in the value of a linear dimension of the test material element, say x, divided by the original value of the element dimension, say l, is called strain ε, i.e. $\varepsilon = x/l$.

(a)   (b)

$$\theta = \tan^{-1} \frac{(b_2 - b_1)}{2h}$$

$$\theta = \sin^{-1} \frac{(D_1 - D_2)}{(2h_2 + D_2) - (2h_1 + D_1)}$$
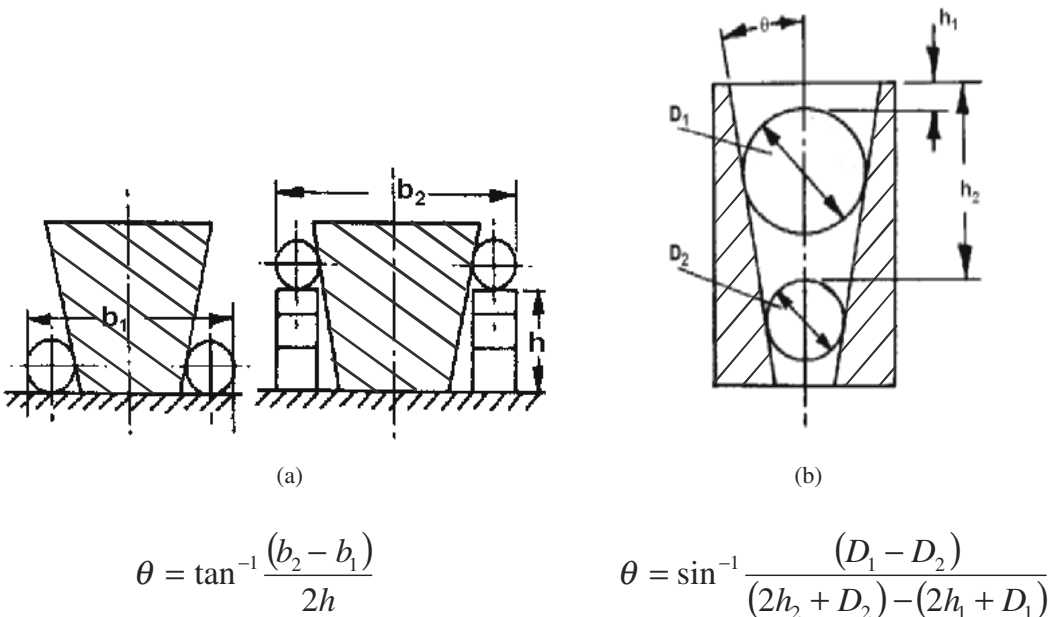
FIGURE 19.5 (a) External taper (using rollers and slip gauges) and (b) Internal taper (using two balls).

If, on removal of external forces, a test element recovers its original shape and size, the material is said to be *elastic.* If it does not return to its original shape, it is said to be *plastic.*

1. Extensiometer.   An extensiometer is an instrument used in engineering and metallurgical design to measure the elastic extensions of materials, in order to forecast their behavior during use. Several different designs of extensiometer with instruction of measurements are given in the engineering literature.[1,6–8]

2. Strain Gauges.   Strain gauges are devices (Fig. 19.6) that experience a change in electrical resistance due to the change in length accompanying the application of stress. They are able to detect very small displacements, usually in the range 0–50 μm, and are typically used as part of other transducers. Measurement inaccurances as low as $\pm 0.15\%$ of full-scale reading are achievable and the quoted life expectancy is usually three million reversals. The strain occurring in the gauge, which is connected into the arm of a Wheatstone bridge, is detected on a galvanometer G (Fig. 19.7).

(a)   (b)
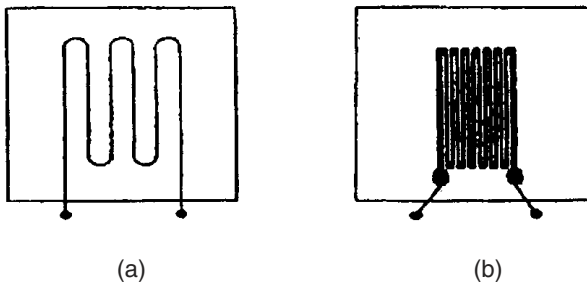
FIGURE 19.6   Strain gauges (*a*) wire type, (*b*) foil type.

(a)   (b)
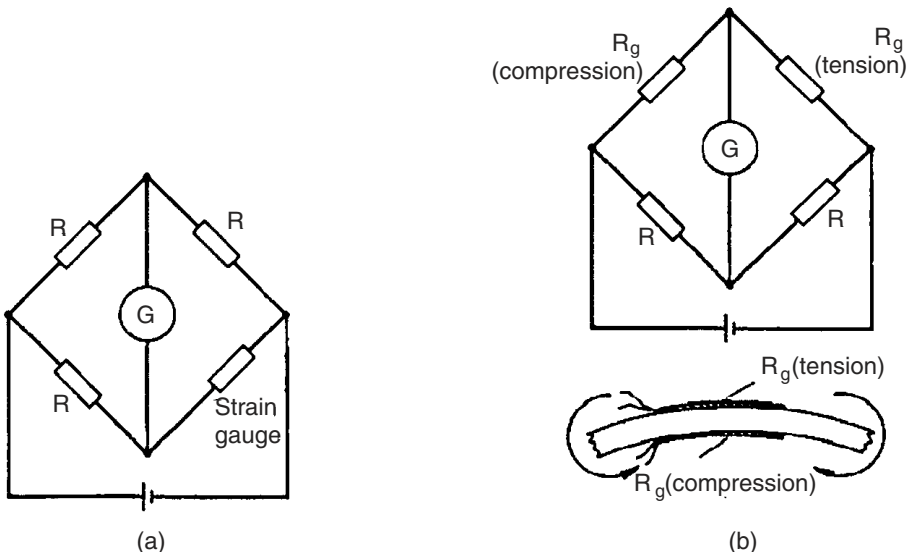
FIGURE 19.7   Wheatstone bridge circuit for strain measurements (*a*); bending (two active gauges, one in tension, one in compression) (*b*).
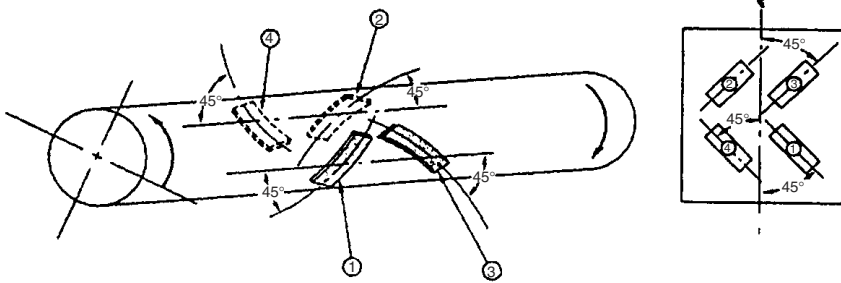
FIGURE 19.8 Torque measurement—gauge positions.

If the resistance changes occur in a strain gauge as a result of applied strain, than the bridge galvanometer deflection is a measure of the amount of strain.

3. Torque Measurement. Two gauges (Fig. 19.8) are mounted on a shaft at 45° to its axis and perpendicular to one another. Under torsion, one gauge is under tension and the other under compression, the stresses being numerically equal to the shear stress. The gauges are connected in a bridge circuit, as for bending. To eliminate bending effects, four gauges may be used, two being on the opposite side of the shaft.

TEMPERATURE MEASUREMENT

A change in temperature of a substance can often result in a change in one or more of its physical properties. Some properties of substances used to determine changes in temperature include changes in dimensions, electrical resistance, state, type and volume of radiation, and color.

1. Liquid-In Glass Thermometer. A liquid-in glass thermometer uses the expansion of a liquid with increase in temperature as its principle of operation (Fig. 19.9).

The commonest type of thermometer uses mercury, which has a freezing point of −39°C and a boiling point of 357°C, although it can be used up to 500°C since the thermometer may contain an inert gas under pressure. The advantages of this thermometer are: good visibility, linear scale, non-wetting, good conductor of heat and pure mercury is easily available. The disadvantages are: it is fragile, slow cooling of glass, long response time, and errors arise due to non-uniform bore and incorrect positioning.
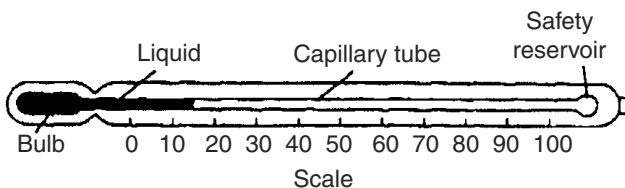
FIGURE 19.9 Liquid-in glass thermometer.

Alcohol can be used as liquid for glass thermometers down to $-113°C$ (freezing point), but its boiling point is only 78°C. The alcohol needs coloring.

2. Thermocouples.   When two dissimilar metals are brought into intimate contact, a small voltage, known as a thermal electomotive force (e.m.f.) exists across it, which increases, usually linearly, with temperature. The basic circuit includes a "cold junction" and a sensitive measuring device—e.g. a galvanometer, G, which indicates the e.m.f. The cold junction must be maintained at a known temperature as a reference—e.g. by an ice bath or a thermostatically controlled oven (Fig. 19.10). If two cold junctions are used, then the galvanometer may be connected by ordinary copper leads (Fig. 19.11).

For practical use, thermocouples are protected by a metal sheath with ceramic beads as insulation.

A copper-constantant thermocouple can measure temperature from $-250°C$ up to about 400°C and is used typically with boiler flue gases, food processing and with sub-zero temperature measurement. An iron-constantant thermocouple can measure temperature from $-200°C$ to about 850°C, and is used typically in paper and pulp mills, re-heat and annealing furnaces and in chemical reactors. A chromel-alumel thermocouple can measure temperatures from $-200°C$ to about 1,100°C and is used typically with blast furnace gases, brick kilns and in glass manufacture. Thermocouples made of platinum-platinum/rhodium are capable of measuring temperature up to 1,400°C, or tungsten–molybdenum, which can measure up to 2,600°C. (See Fig. 19.12.)
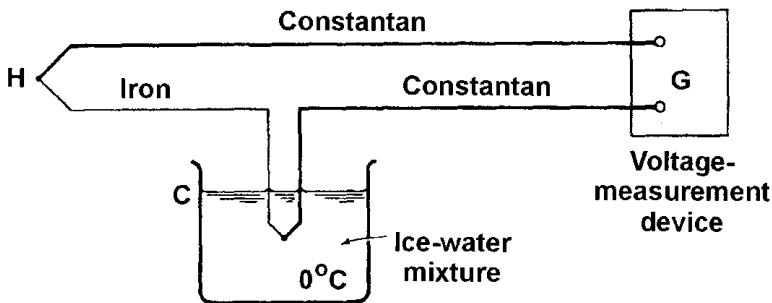
FIGURE 19.10   Thermocouple circuit with ice bath.
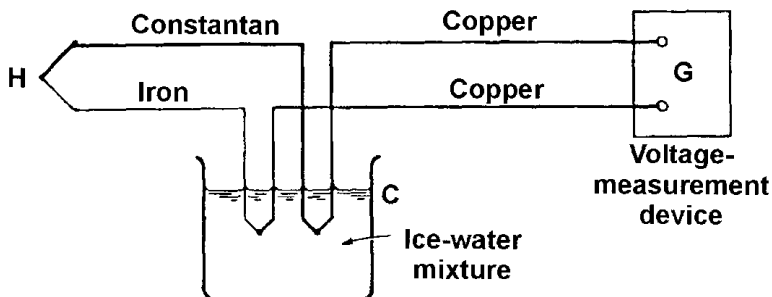
FIGURE 19.11   Thermocouple circuit with extension leads (hot junction H, cold junction C).

FIGURE 19.12   Emf temperature relations for thermocouple materials (Ref. 7).

TABLE 19.3   Thermoelectric Sensitivity of Thermocouple Material Relative to Platinum (Reference Junction at 0°C)

Metal	Sensitivity (μV °C^{-1})	Metal	Sensitivity (μV °C^{-1})
Bismuth	−72	Silver	6.5
Constantan	−35	Copper	6.5
Nickel	−15	Gold	6.5
Potassium	−9	Tungsten	7.5
Sodium	−2	Cadmium	7.5
Platinum	0	Iron	18.5
Mercury	0.6	Nichrome	25
Carbon	3	Antimony	47
Aluminum	3.5	Germanium	300
Lead	4	Silicon	440
Tantalum	4.5	Tellurium	500
Rhodium	6	Selenium	900

Source: Ref. 25.

The advantage of thermocouples are: they are simple in construction (Fig. 19.13) compact, robust and relatively cheap, they are suitable for remote control, automatic systems and recorders since they have a short response time.

The disadvantages are that they suffer from errors due to voltage drop in the leads, variation in cold-junction e.m.f. and stray thermoelectric effects in leads.

3. Resistance Thermometers.   Resistance thermometers use the change in electrical resistance caused by temperature change. Therefore, resistance thermometers are based on the fact that the electrical resistance of a metal wire varies with temperature. The metals most used are platinum and nickel, for which the resistance increases with temperature in a linear manner. If R_0 is the resistance at 0°C, than the resistance R_t at t °C is:

$$R_t = R_0 \cdot (1 + \alpha \cdot t)$$

or

$$t = (R_t - R_0)/\alpha \cdot R_0$$

where α = temperature coefficient of resistance. The value of α is given for a number of metals as well as electrolytes and semi-conductors in the table below.

Fig. 19.14 shows the construction of a typical resistance thermometer. It consists of a small resistance coil enclosed in a metal sheath with ceramic insulation beads. The small resistance change is measured by means of a Wheatstone bridge (see section on electrical measurements), and dummy heads eliminate temperature effects on the element leads.

The resistance thermometer is used for heat treatment and annealing furnaces and for calibration of other thermometers. The main disadvantages are fragility and slow response.
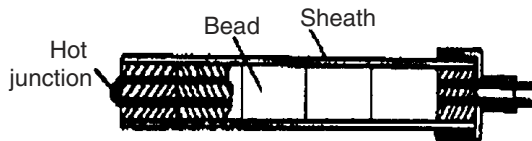
FIGURE 19.13   Thermocouple (practical arangement).

TABLE 19.4   Resistance Temperature Coefficients (at Room Temperature), 1/°C

Material	α (°C^{-1})	Material	α (°C^{-1})
Nickel	0.0067	Gold	0.004
Iron	0.002–0.006	Platinum	0.00392
Tungsten	0.0048	Mercury	0.00099
Aluminum	0.0045	Manganin	± 0.00002
Copper	0.0043	Carbon	-0.0007
Lead	0.0042	Electrolytes	-0.02 to -0.09
Silver	0.0041	Semi-conductor (thermistor)	-0.068 to $+0.14$

Source: Ref. 25.
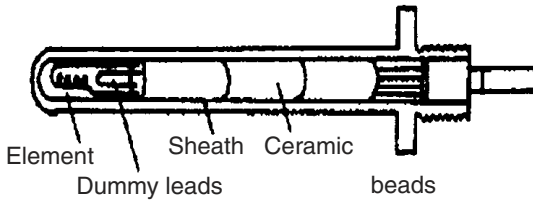
FIGURE 19.14   Typical resistance thermometer.

4. Thermistors.   A thermistor is a semi-conducting material, such as mixtures of oxides of copper, manganese, cobalt, etc., in the form of a fused bead connected to two leads. As its temperature is increased, its resistance rapidly decreases. Typical resistance/temperature curves for three thermistor materials compared with platinum is shown in Fig. 19.15.

A "thermistor" is a bead of such material (e.g. oxides of copper, manganese and cobalt), with leads connected to a measuring circuit. The resistance of a typical thermistor can vary from 400 Ω at 0°C to 100 Ω at 140°C.

The main advantages of a thermistor are its high sensitivity and small size. It provides an inexpensive method of measuring and detecting small changes in temperature.

5. Pyrometers.   At very high temperatures where thermometers and thermocouples are unsuitable, temperature can be deducted from the measurement of radiant energy from a hot source. The radiation is passed down a tube and focused, using a mirror, onto a thermocouple that is shielded from direct radiation.

Total radiation pyrometer is shown in Fig. 19.16.

Radiant energy from a hot source, such as a furnace, is focused on to the hot junction of a thermocouple after reflection from a concave mirror. The temperature
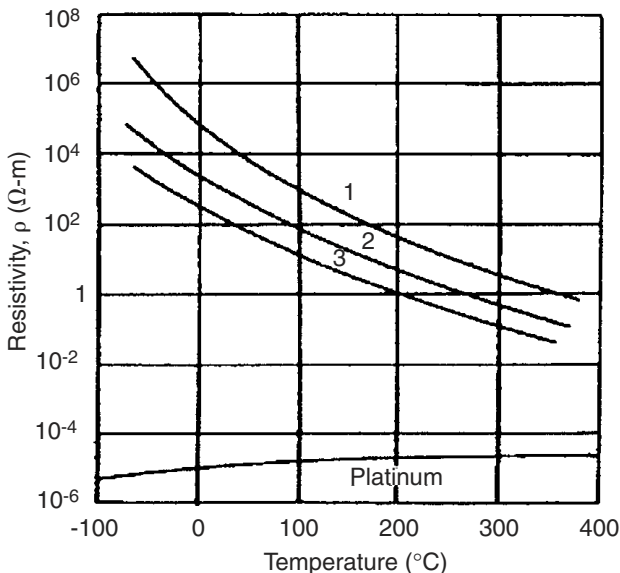
FIGURE 19.15   Resistance vs. temperature (Ref. 7).
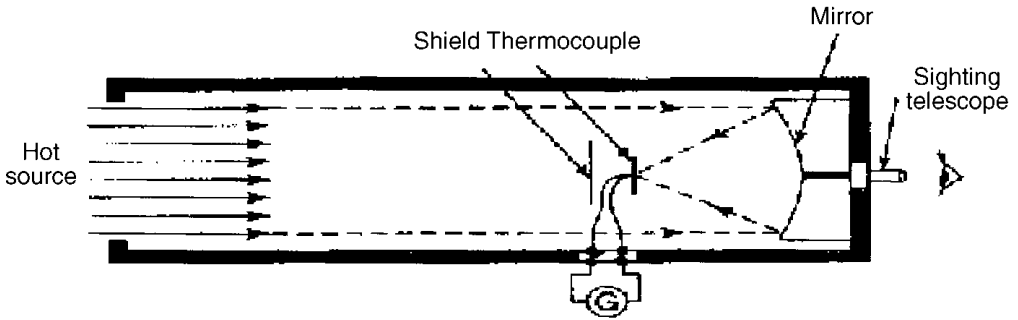
FIGURE 19.16   Total radiation pyrometer.

rise recorded by the thermocouple depends on the amount of radiant energy re-
ceived, which in turn depends on the temperature of the hot source. The galvanom-
eter G connected to the thermocouple records the current that results from e.m.f.
developed and may be calibrated to give a direct reading of the temperature of the
hot source. Total radiation pyrometers are used to measure temperature in the range
700°C to 2,000°C.

Disappearing-filament pyrometer is shown in Fig. 19.17.

The brightness and color of a hot body varies with temperature, and, in the case
of the disappearing-filament pyrometer, it is compared with the appearance of a
heated lamp filament. The radiation is focused onto the filament, the brightness of
which is varied by means of a calibrated variable resistor until the filament appears
to vanish. A red filter protects the eye.

Manual adjustments are necessary for all pyrometers. A reasonable amount of
skill and care is required in calibrating and using a pyrometer. For each new mea-
suring situation, the pyrometer must be re-calibrated.

6. Temperature Indicating Paints and Crayons.   Temperature indicating paints
contain substances that change their color when heated to certain temperatures. This
change is usually due to chemical decomposition, such as loss of water, in which
the change in color of the paint after having reached the particular temperature willl
be a permanent one.

Kits are available of paints and crayons made of chemicals that change color at
definite temperatures. The range is from about 30°C to 700°C, with an accuracy of
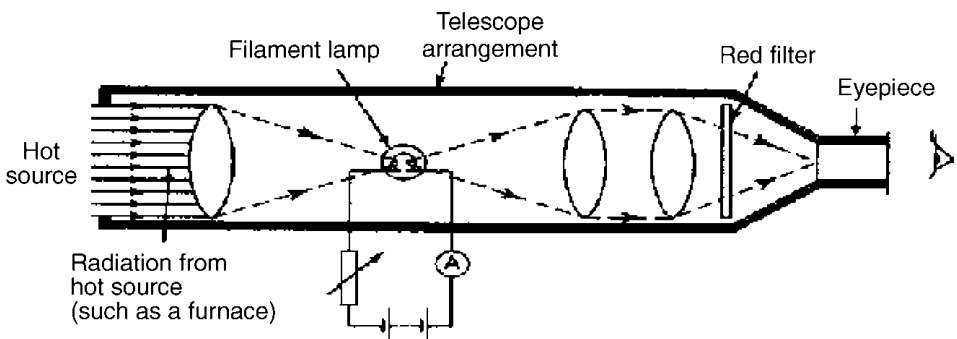
FIGURE 19.17   Disappearing-filament pyrometer.
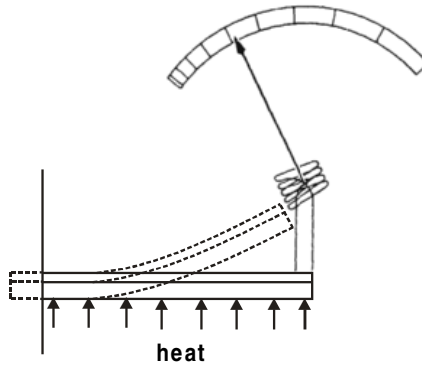
FIGURE 19.18 Bimetallic thermometer sketch.

TABLE 19.5 Fixed-Point Temperatures

	Temperature (°C)
Boiling point of liquid oxygen	-182.97
Melting point of ice	0.00
Triple point of water	0.01
Boiling point of water	100.00
Boiling point of liquid sulphur	444.60
Melting point of silver	960.80
Melting point of gold	1063.00

about 5%. Several paints are required to cover the range. Crayons are the easiest to use. The method is suitable for inaccessible places.

7. Bimetallic Thermometers.   Bimetallic thermometers depend on the expansion of metal strips that operate an indicating pointer (Fig. 19.18).

Two thin metal strips of differing thermal expansion are welded or riveted together and the curvature of the bimetallic strip changes with temperature change. Bimetallic thermometers are useful for alarm and over-temperature applications where extreme accuracy is not essential. The normal upper limit of temperature measurement by this thermometer is about 200°C, although with special methods the range can be extended to about 400°C.

8. Fixed-Point Temperatures.   Table 19.5 gives fixed-point temperatures known to a high degree of accuracy from which instruments can be calibrated.

PRESSURE MEASUREMENT

Definitions of absolute and gauge pressure are given in Chapter 8 (section on fluid statics). Pressure-indicating instruments are made in a wide variety of forms because of their many different applications.

1. Barometers.   A barometer is an instrument for measuring atmospheric pressure. A simple barometer consists of a glass tube, just under 1 m in length, sealed at one end, filled with mercury and then inverted into a trough containing more mercury. Care must be taken to ensure that no air enters the tube during this latter process. Such a barometer is shown in Fig. 19.19, and it is seen that the level of the mercury column falls, leaving an empty space, called a vacuum. Atmospheric pressure acts on the surface of the mercury in the trough as shown, and this pressure is equal to the pressure at the base of the column of mercury in the inverted tube, i.e. the pressure of the atmosphere is supporting the column of mercury. If the atmospheric pressure falls, the barometer height h decreases.

Another arrangement of a typical barometer is shown in Fig. 19.20 where a U-tube is used instead of an inverted tube and trough.

If, instead of mercury, water was used as the liquid in a barometer, then the barometer light h at standard atmospheric pressure would be 13.6 times more than for mercury (due to density ratio), i.e. about 10.4 m high, which is not practicable.

2. Manometers.   A manometer is a device for measuring or comparing fluid pressures and is the simplest method of indicating such pressures. Manometers are passive instruments that give a visual indication of pressure values.

The U-tube manometer may be used to measure a pressure relative to atmospheric pressure, or the difference between two pressures. If one leg is much larger in diameter than the other, a single-leg manometer is obtained and only a single
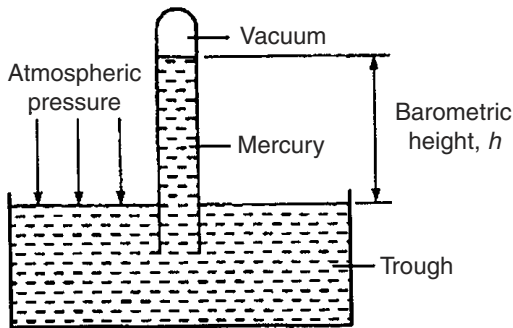
FIGURE 19.19   Mercury barometer.
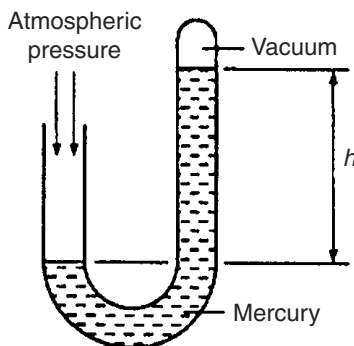
FIGURE 19.20   U-tube barometer.

reading is required (as for barometer). The inclined single-leg manometer gives greater accuracy. See section on fluid statics (Chapter 8, sec. 5) for specific ΔP relations for different U-tube configurations.

When the manometer fluid is less dense than the fluid, the pressure of which is to be measured, an inverted manometer is used (Fig. 19.21).

3. Bourdon-Tube Pressure Gauge.   The Bourdon gauge is the most commonly used pressure device (Fig. 19.22). It consists of a flattened tube of spring bronze or steel bent into a circle. Pressure inside the tube tends to straighten it. Since one end of the tube is fixed to the pressure inlet, the other end moves proportionally to
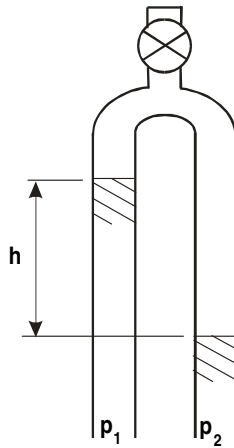
FIGURE 19.21   Inverted U-tube manometer.
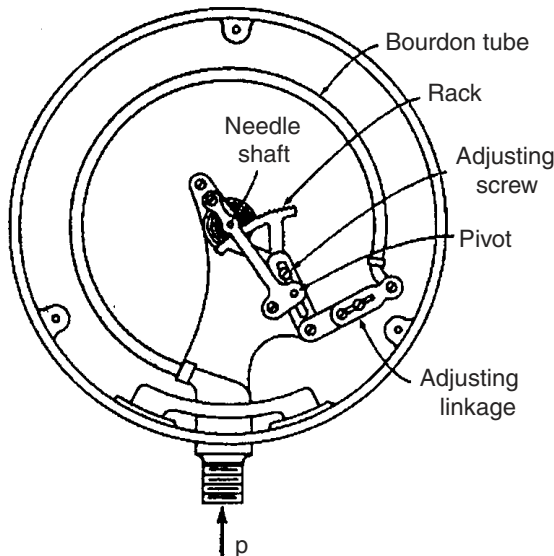
FIGURE 19.22   Bourdon-tube gauge.

the pressure difference existing between the inside and outside of the tube. The motion rotates the pointer through a pinion-and-sector mechanism. For amplification of the motion, the tube may be bent through several turns to form spiral or helical elements as are used in pressure recorders.

This gauge can be used for liquids or gases from a fraction of a bar pressure up to 10^4 bar.

4. Diaphragm Gauge. In the diaphragm gauge (Fig. 19.23), the pressure acts on a diaphragm in opposition to a spring or other elastic member. The deflection of the diaphragm is therefore proportional to the pressure. Since the force increases with the area to the diaphragm, very small pressures can be measured by the use of large diaphragms. The diaphragm may be metallic (brass, stainless steel) for strength and corrosion resistance, or nonmetallic (leather, neoprene, rubber) for high sensitivity and large deflection. With a stiff diaphragm, the total motion must be very small to maintain linearity.

5. Pressure Transducers. A wide range of transducers is available that convert the deflection of a diaphragm into an electric signal that gives a reading on an indicator or is used to control a process, etc. Transducers cover a wide range of pressure and have a fast response. Types include piezo-crystal, strain gauge, variable capacity, and variable inductance.[3,9]

FLOW VELOCITY MEASUREMENT

1. Pitot-Static Tube The pitot-static tube consists of two concentric tubes, the central one with an open end pointing upstream of the fluid flow (Fig. 19.24) and the other closed at the end but with small holes drilled at right angles to the direction of flow. The opening in front of the probe senses the stagnation pressure, while the small holes around the outer periphery of the tube sense the static pressure.

The central tube pressure is equal to the static pressure plus the "velocity pressure", whereas the outer tube pressure is the static pressure only.

A manometer or other differential pressure measuring device measures the pressure difference between the tubes, which is equal to the "velocity pressure". For
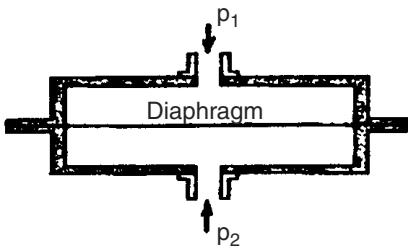
FIGURE 19.23 Schematic of a diaphragm gauge.
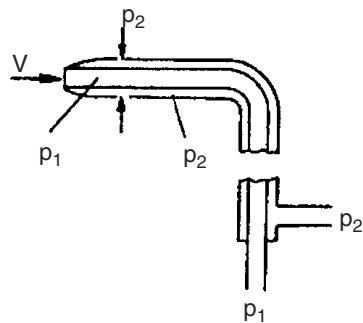
FIGURE 19.24 Pitot-static tube.

large pipes or ducts, traversing gear is used and an average value of velocity cal-
culated.

Fluid velocity:

$$V = \sqrt{\frac{2(p_2 - p_1)}{\rho_f}}$$

where ρ_f = fluid density.

2. Hot-Wire Anemometer. A simple hot-wire anemometer consists of a small
piece (Fig. 19.25) of wire that is heated by an electric current and positioned in
the air or gas stream whose velocity is to be measured.

The stream passing the wire cools it, the rate of cooling being dependent on the
flow velocity. The resulting change in resistance of the element is measured by a
bridge circuit and is related to velocity by calibration.

A hot-wire anemometer is small in size and has great sensitivity.

3. Anemometer. An anemometer is used to measure velocity, usually of air. Its
hemispherical cups are attached to a rotating shaft (Fig. 19.26a). The shape of the
cups gives a greater drag on one side than the other and the result in speed of
rotation is approximately proportional to the air speed. Velocity is found by mea-
suring revolutions over a fixed time.
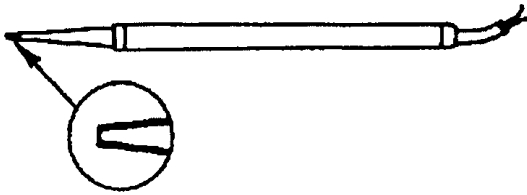
FIGURE 19.25   Hot-wire anemometer.

(a)   (b)
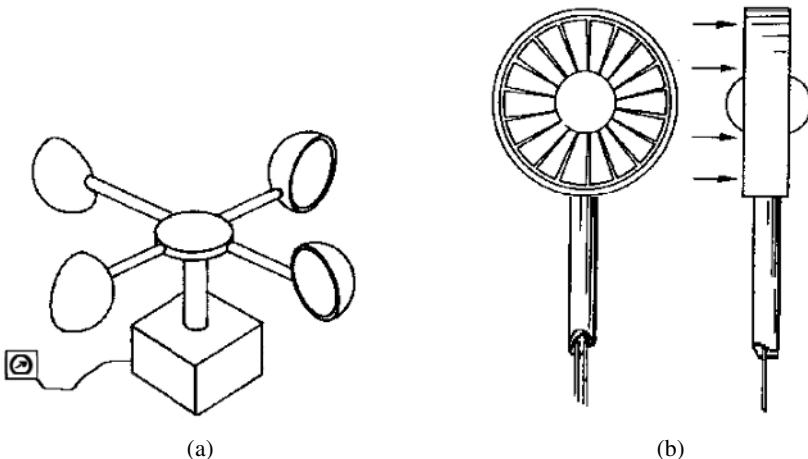
FIGURE 19.26   Anemometer, (*a*) cap-tube anemometer; (*b*) vane anemometer.

The "vane anemometer" has an axial impeller attached to a handle with extensions and an electrical pick-up, which measures the revolutions (Fig. 19.26*b*).

4. Rotational-Speed Measurements.   *Stroboscope* has an electronic flash tube that flashes at a variable rate and that is adjusted to coincide with the rotational speed so that the rotational object, or a suitable mark on it, appears to stand still. The flash-rate control is calibrated in rotational speed.

5. Laser Doppler Anemometer (LDA).   The laser anemometer is a device that offers the nondisturbance advantages of optical methods while affording a very precise quantitative measurement of velocities. The device is capable of rapid response and is suitable for measurement of high-frequency turbulence fluctuations.

One possible schematic of a LDA is shown in Fig. 19.27. The laser beam is focused on a small-volume element in the flow through lens L_1. In order for the device to function, the flow must contain some type of small particles to scatter the light, but the particle concentration required is very small. Ordinary tap water, for example, contains enough impurities to scatter the incident beam. Two additional lenses L_2 and L_3 are positioned to receive the laser beam that is transmitted through the fluid (lens L_3) and some portion of the beam that is scattered through the angle θ (lens L_2). The scattered light experiences a Doppler shift in frequency that is directly proportional to the flow velocity.

It is clear that the LDA measures the velocity of the scattering particles. If they are sufficiently small, the slip between particles and fluid velocities will be small, and thus an adequate indication of fluid velocity will be obtained. Laser anemometers that measure more than one velocity component simultaneously have been developed.
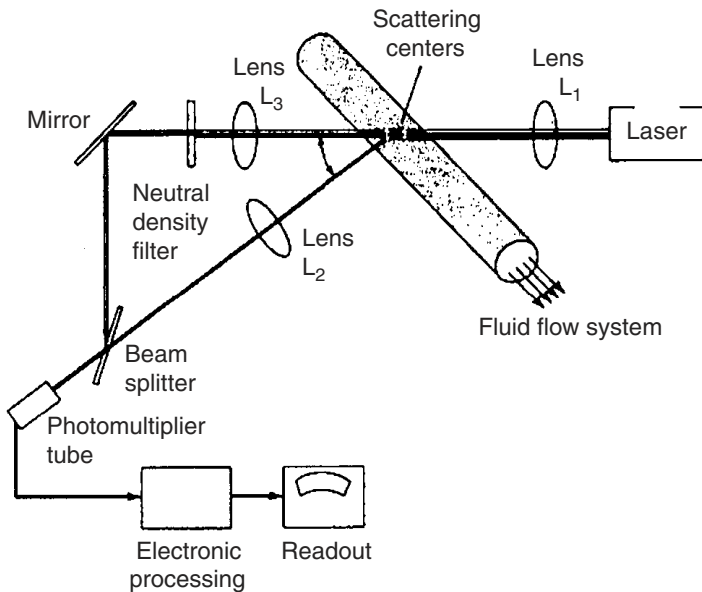
FIGURE 19.27   Schematic of laser-anemometer for fluid flow-measurement.

MEASUREMENT OF FLUID FLOW

Fluid flow is one of the most difficult of industrial measurements to carry out, since flow behavior depends on a great many variables concerning the physical properties of a fluid.

Fluid flow measuring instruments are generally called flowmeters. Two main categories of flowmeters are a) differential pressure flowmeters and b) mechanical flowmeters.

Differential Pressure Flowmeters.   Flowmeters installed in pipelines often cause an obstruction to the fluid flowing in the pipe by reducing the cross-section area of the pipeline. This causes a change in the velocity of the fluid with a related change in pressure. Figure 19.28 shows a schematic of three typical "obstruction" flowmeters *a*) venturi, *b*) flow nozzle and *c*) orifice. The flow rate of the fluid may be determined from measurement of the difference between the pressure on the wall at the pipe, a specified upstream location 1 and downstream location 2.

Therefore:

$$Q = \frac{A_2}{\sqrt{1 - (A_2/A_1)^2}} \cdot \sqrt{\frac{2}{\rho_f} \cdot (p_1 - p_2)}$$

where Q = volumetric flow rate (m³/s)
  ρ_f = fluid density (kg/m³)
 A_1, A_2 = cross-section areas (m²)

The pressure difference $(p_1 - p_2)$ is measured using a manometer connected to appropriate pressure tapping points at cross-section areas 1 and 2.
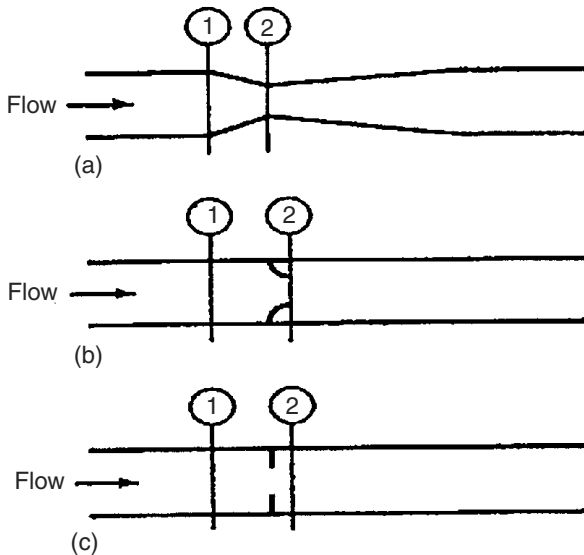
FIGURE 19.28   Schematic of three typical obstruction meters (*a*) venturi; (*b*) flow nozzle, (*c*) orifice.

Calibration of the manometer depends (Fig. 19.29) on the shape of the obstruction (orifice plate, flow nozzle or venturi), the position at the pressure tapping points and the physical properties of the fluid. In industrial applications, the pressure difference is detected by a differential pressure cell, the output from which is either an amplified pressure signal or an electrical signal.

1. Venturi Tube Meter. A typical arrangement (ASME Standard[5]) of a section through such a device is shown in Fig. 19.30 and consists of a short converging conical tube called the inlet or upstream cone, leading to a cylindrical portion called the throat. The diverting section called the outlet or recovery cone follows this.

Flow meter causing obstruction in fluid flow
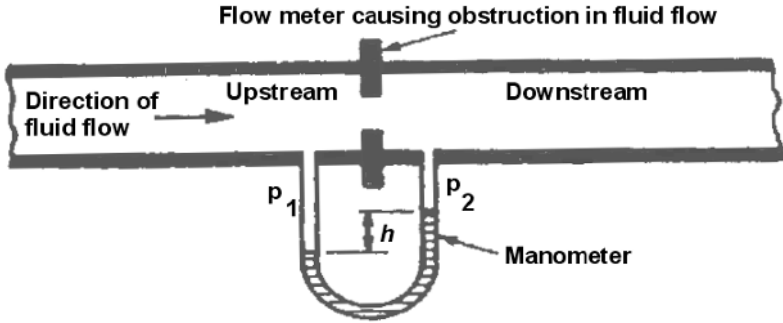
FIGURE 19.29   Pipeline section into which a flowmeter is inserted.
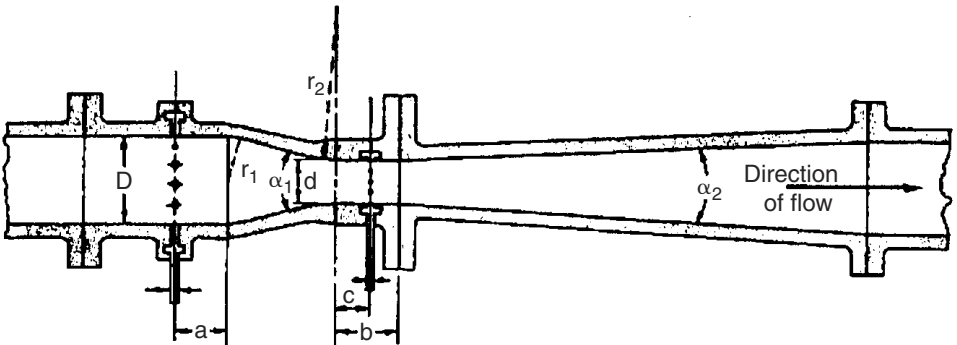
D = Pipe diameter inlet and outlet
d = Throat diameter as required
a = 0.25D to 0.75D for 4″ $\leqq D \leqq$ 6″, 0.25D to 0.50 D for 6″ $< D \leqq$ 32″
b = d
c = $d/2$
δ = $^3/_{16}$ in. to $^1/_2$ in. according to D. Annular pressure chamber
     with at least four piezometer vents
r_2 = 3.5d to 3.75d
r_1 = 0 to 1.375D
α_1 = 21° ± 2°
α_2 = 5° to 15°
1″ = 25.4 m

FIGURE 19.30   Venturi flowmeter.

Pressure tapping, are made at the entry and at the throat and the pressure difference *h* which is measured using a manometer, a differential pressure cell or similar gauge, is dependent on the flow rate through the meter.

Usually, pressure chambers are fitted around the entrance pipe and the throat circumference with a series of tapping holes made in the chamber to which the manometer is connected. This ensures that an average pressure is recorded. The loss of energy due to turbulence that occurs just downstrean with an orifice plate is largely avoided in the venturi meter due to the gradual divergence beyond the throat. Venturimeters are usually made a permanent installation in a pipeline and are manufactured usually from stainless steel, cast iron, monel metal or polyester glass fibre.

Advantages of venturimeters are high accuracy and low pressure loss in the tube (2% to 3% in a well-proportioned tube). Disadvantages are high manufacturing costs and installation tends to be rather long.

2. Orifice Plate Meter.   An orifice plate meter consists of a circular, thin, flat plate with a hole (or orifice) machined through its center to fine limits of accuracy. Recommended location of pressure taps for use with concentric orifice is given in Fig. 19.31.

The variation of pressure near an orifice plate is shown in Fig. 19.32 together with the orifice plate and diaphragm-type meter. The position of minimum pressure is located downstream from the orifice plate where the flow stream is narrowest. This point of minimum cross-sectional area of the jet is called the "vena contracta". Beyond this point the pressure rises but does not return to the original upstream value and there is a permanent pressure loss. Advantages of orifice plate are: relative, inexpensive and fit between a standard pair of pipe flanges. Orifice plates are usually used in medium and large pipes and are best suited to the indication and control of essentially constant flow rate.

3. Flow Nozzle Meter.   The flow nozzle lies between an orifice plate and the venturimeter both in performance and cost. A typical section through a flow nozzle is shown in Fig. 19.33 with ASME standard. The fluid flow does not contract any further as it leaves the nozzle and the pressure loss created is considerably less than that occurring with orifice plate. Flow nozzles are suitable for use with high velocity flows for they do not suffer the wear that occurs in orifice plate edges during such flows.
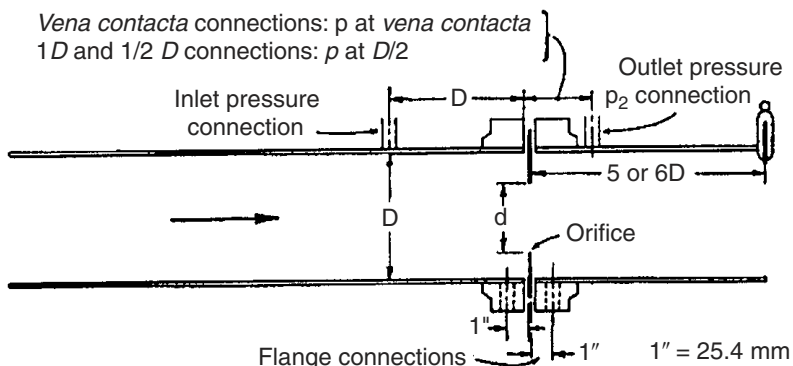
FIGURE 19.31   Orifice location.
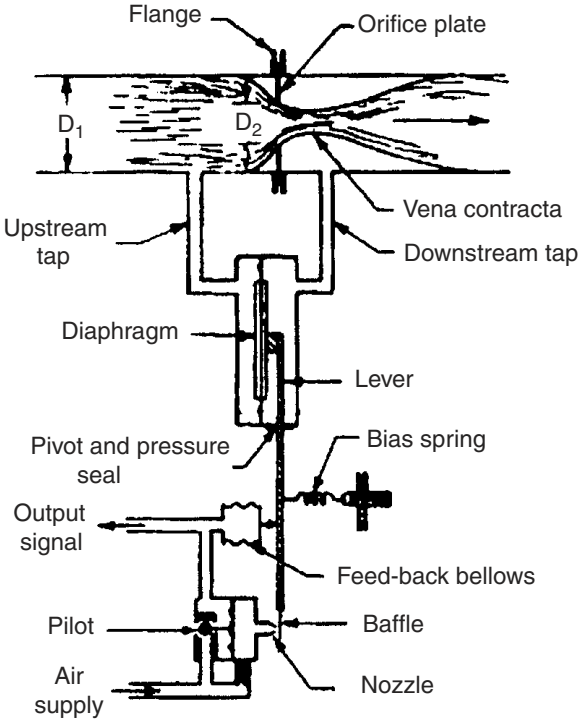
FIGURE 19.32 Orifice plate; diaphragm-type meter; pressure drop variation.

Low β series: $\beta < 0.5$
$r_1 = d$
$r_2 = {}^2/3d$
$L_t = 0.6d$
$1/8'' \gtrless t \gtrless 1/2''$
$1/8'' \gtrless t_2 \gtrless 0.15D$

High β series: $\beta > 0.25$
$r_1 = {}^1/2D$
$r_2 = {}^1/2(D - d)$
$L_t \gtrless 0.6d$ or $L_t \gtrless {}^1/2D$
$2t \gtrless D - (d + {}^1/8'')$
$1/8'' \gtrless t_2 \gtrless 0.15D$

Optional designs
of nozzle outlet

$\beta = d/D$

$1'' = 25.4$ mm
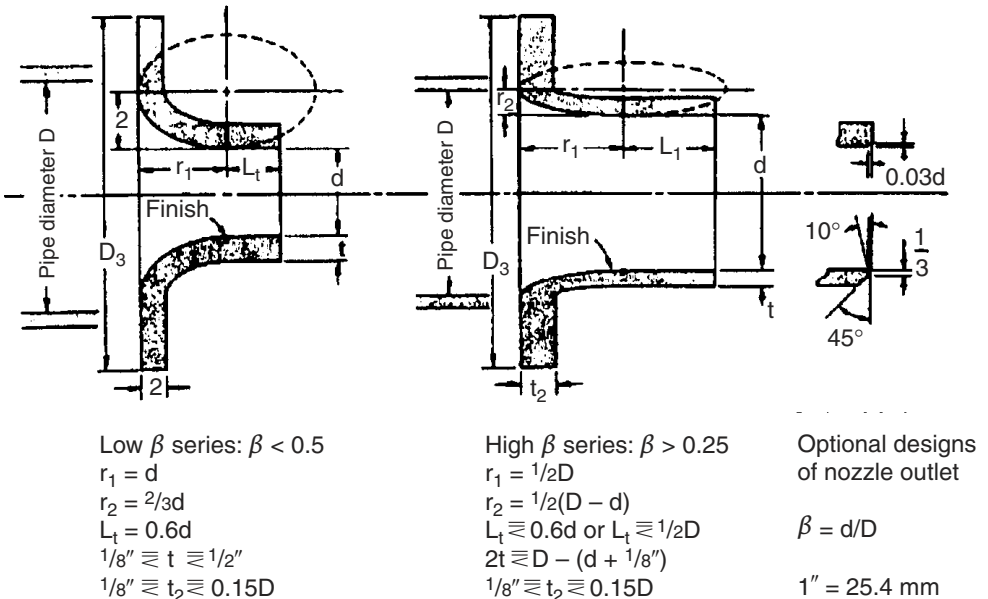
FIGURE 19.33 Flow nozzle—ASME recommended proportion.

For all three differential pressure flowmeters (venturi, orifice, flow nozzle) volumetric flow

$$Q = Co\sqrt{p_1 - p_2}$$

where values of Co "discharge contrant" depends on the detailed meter/channel geometry and the flow Reynolds number.[3,5,7]

Mechanical Flowmeters. With mechanical flowmeters, the fluid flowing past it displaces a sensing element situated in pipeline. Typical examples of mechanical flowmeters are: deflecting vane flowmeter and turbine-type meters.

4. Deflecting Vane Flowmeter. The deflecting vane flowmeter consists basically of a pivoted vane, suspended in the fluid flow stream (Fig. 19.34). When a stream of fluid impinges on the vane it deflects from its normal position by an amount proportional to the flow rate. The movement of the vane, is indicated on a scale that may be calibrated in flow units. This type of meter is used for measuring liquid flow rates in open channels or for measuring the velocity of air in ventilation ducts.

5. Turbine Type Meters. An axial or tangential impeller mounted in a pipe rotates (Figs. 19.35 and 36) at a speed roughly proportional to the velocity, and hence the flow, of the fluid in the pipe. The rotational speed is measured either mechanically or electronically to give flow or flow rate.

Float and Taped-Tube Meter. With the differential flowmeters (orifice . . .), the area of the opening in the obstruction is fixed and any change in the flow rate produces a corresponding change in pressure. With the float and tapered-tube meter,
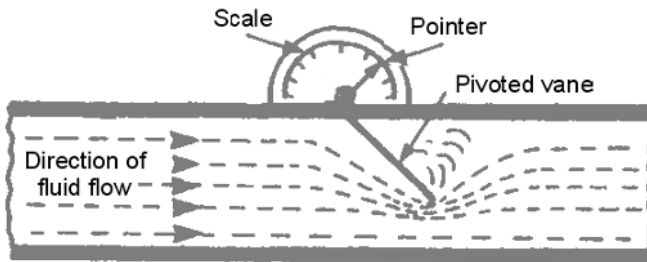
FIGURE 19.34 Deflecting vane flowmeter.
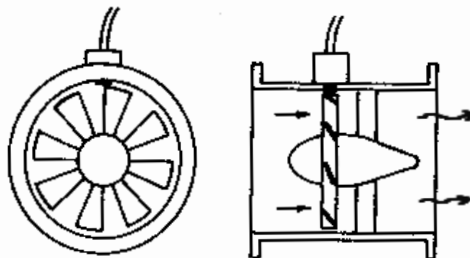
FIGURE 19.35 Axial-impeller flowmeter.
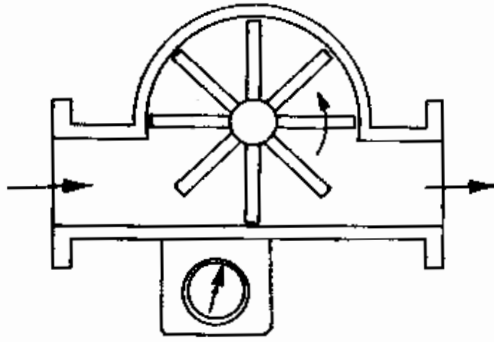
FIGURE   19.36 Tangential-impeller   flow-
meter.

the area at the restriction may be varied so as to maintain a steady pressure differ-
ential. A typical meter of this type is shown in Fig. 19.37, where a vertical tapered
tube contains a "float" that has a density greater than the fluid.

The fluid, which may be a liquid or gas, flows through the annular space between
the float and the tube.

As the flow is increased, the float moves to a greater height. The movement is
roughly proportional to flow, and calibration is usually carried out by the supplier.
Angled grooves in the rim at the float cause rotation and give the float stability.
Rotometers have a very simple design, can measure very low flow rates, and can
be made direct reading. They can only be installed vertically in a pipeline.

Electromagnetic Flowmeter.   The flow rate of fluids that conduct electricity, such
as water or molten metal, can be measured using an electromagnetic flowmeter
whose principle of operation is based on the laws of electromagnetic induction.
When a conductor of length l moves at right angles to a magnetic field of density
B at a velocity v, an e.m.f. E is generated as:
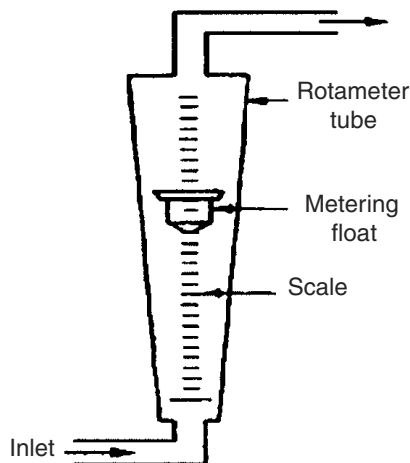
FIGURE 19.37   Rotometer.
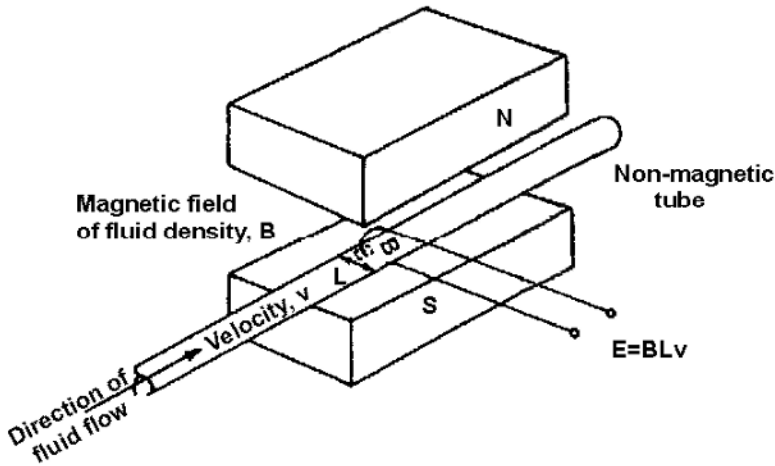
FIGURE 19.38   Electromagnetic flowmeter.

$$E = BLv$$

or

$$v = E/(B \cdot L)$$

Knowing the flow area of the tube, the volumetric flow rate is directly calculated (measured). The specific electromagnetic flowmeter arrangement is shown in Fig. 19.38, the fluid is the conductor and the e.m.f. is detected by two electrodes placed across the diameter of the non-magnetic tube. An advantage of this meter is a linear relationship between the fluid flow and the induced e.m.f.

ELECTRICAL MEASUREMENTS

(*Current, Voltage, Resistance, Frequency, Power . . .*)

1. Introduction.   The three main effects of an electric current are:

- *magnetic effect* (applications: motors, relays, bells, generators, transformers, telephones, lifting magnets, car-ignition . . .);
- *heating effect* (applications: water heaters, electric fires, cookers, irons . . .);
- *chemical effect* (application: primary and secondary cells and electroplating . . .).

These effects are applied for development of basic electrical measurements and sensing devices. In order to detect electrical quantities such as current, voltage, resistance or power, it is necessary first to transform an electrical quantity into a visible indication. (Also, to measure other physical quantities it is practical to convert these quantities into electrical variables.)

This is accomplished with instruments (or meters) that indicate the magnitude of quantities either by the position of a pointer moving over a graduated scale (an analog instrument) or in the form of a decimal number (a digital instrument).

In general, there is a sharp distinction between a.c. (alternating current) and d.c. (direct current) devices used in measurements. Consequently, it is often desirable to transform an a.c. signal to an equivalent d.c. value, and vice versa. An a.c. signal is converted to d.c. (rectified) by use of selenium rectifiers, silicon or germanium diodes, or electron-tube diodes.

2. Basic Instruments

Ammeter.   An ammeter is used to measure current and must be connected in series with the circuit (Fig. 19.39). Since all the current in the circuit passes through the ammeter, it must have a very low resistance. When an ammeter is used to measure currents of larger magnitude, a proportion of the current is diverted through a low-value resistance connected in parallel with the meter (Fig. 19.40). Such a diverting resistor is called a shunt. From Fig. 19.40, $V_{1-4} = V_{2-3}$ hence $i_a \cdot r_a = i_s \cdot R_s$.

Therefore:

$$R_s = \frac{i_a \cdot r_a}{i_s} \quad (\Omega)$$

3. Galvanometer.   A sensitive milli-ammeter with center zero position setting is called a *galvanometer* (G). It is used mostly for the null measurements. The null method of measurements is a simple and accurate method where the instrument is
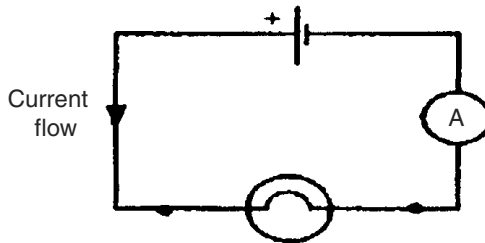
FIGURE 19.39   Ammeter.
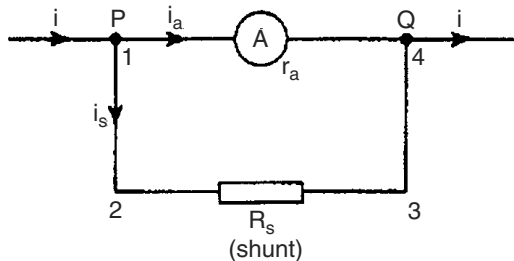
FIGURE 19.40   Diverting resistor.

adjusted to read zero current. If there is any deflection then some current is flowing through the circuit. If there is no deflection, then no current flows (i.e. we have a null condition). Examples where this method is used are in the d.c. potentiometer, a.c. bridges, the Wheatstone bridge, etc.

The mechanical structure of the galvanometer is shown in Fig. 19.41. The input signal is applied across a coil mounted in a jeweled bearing or on a tautband suspension so that it is free to rotate between the poles of a permanent magnet.

Current in the coil produces a magnetic moment that tends to rotate the coil. The rotating deflection of the coil θ is proportional to the current i

$$\theta = \frac{n \cdot B \cdot W \cdot L}{K} \cdot i$$

where n = number of turns of coil; W, L = coil width and length, respecively; B = magnetic field intensity; K = spring constant of the hairspring. Galvanometer deflection is indicated by a balanced pointer attached to the coil. In very sensitive applications, the pointer is replaced by a mirror reflecting a spot of light onto a groundglass scale. The bearings and hairspring are replaced by a torsion-wire suspension. More on similar moving-iron and moving-coil rectifier configuration is available in various references.[6,9,10]

3. Voltmeter. A voltmeter is an instrument used to measure *potential difference* and must be connected in parallel with the part of the circuit whose potential difference is measured. In Fig. 19.42, a voltmeter is connected in parallel with the lamp to measure the potential difference across it. To prevent a significant current flowing through it, a voltmeter must have a high resistance.

4. Potentiometer. The d.c. potentiometer is a null-balance instrument used for determining values of the electromotive forces (e.m.f.) and a potential difference (p.d.) by comparison with a known e.m.f. or p.d. Fig. 19.43*a* shows a standard cell of known e.m.f., the slider S is moved until balance is obtained (i.e. the galvanometer deflection is zero), shown as length l_1. The standard cell is now replaced by a
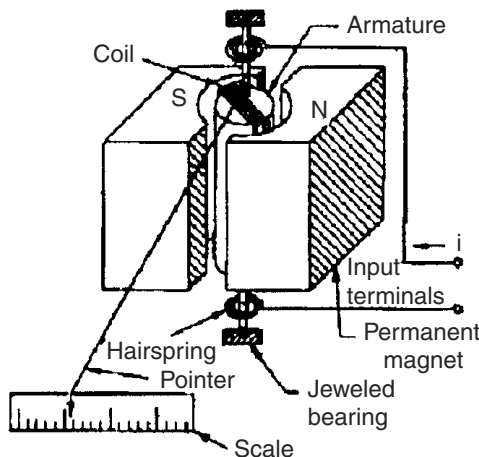
FIGURE 19.41 Galvanometer.
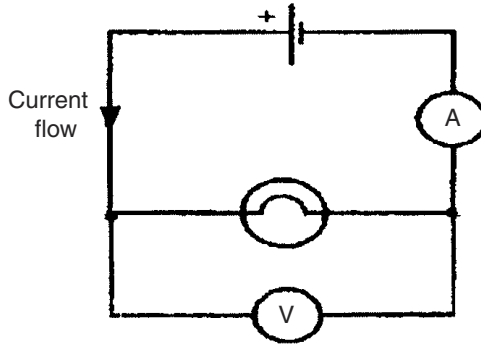
FIGURE 19.42 Voltmeter.
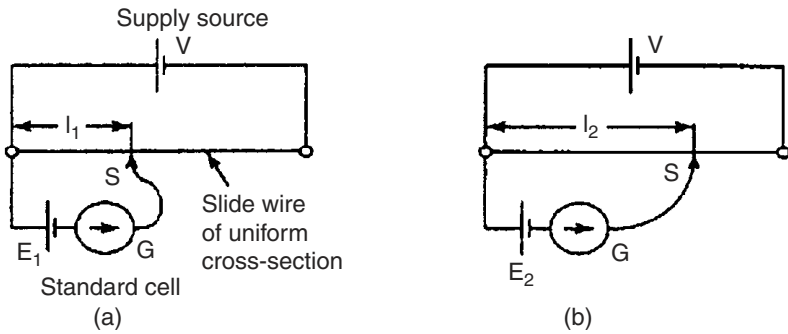
FIGURE 19.43 Potentiometer.

cell of unknown e.m.f. E_2 (Fig. 19.43*b*) and again balance is obtained as length l_2. Therefore:

$$\frac{E_1}{E_2} = \frac{l_1}{l_2}$$

or

$$E_2 = E_1 \cdot \frac{l_1}{l_2} \quad \text{(V)}$$

5. Ohmmeter. An ohmmeter is an instrument for measuring electrical resistance (Fig. 19.44). Unlike the ammeter or voltmeter, the ohmmeter circuit does not receive the energy necessary for its operation from the circuit under test. In the ohmmeter, this energy is supplied by a self-contained source of voltage, such as a battery. Initially, terminals 1-1 are short-circuited and R adjusted to give full-scale deflection of the milli-ammeter. If current i is at a maximum value and voltage E is constant, then the resistance $R = E/i$ is at a minimum value. The full-scale deflection on the milli-ammeter is made zero on the resistance scale. If terminals 1-1 are open-circuited, no current flows and $R = (E/0 \rightarrow \infty)$ is infinity. The calibrated scale gives the correct reading of any unknown resister.
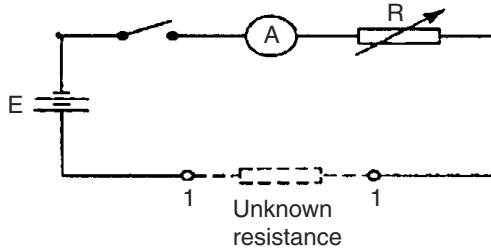
FIGURE 19.44 Ohmmeter.

5. Wheatstone Bridge. A Wheatstone bridge circuit (shown in Fig. 19.45) is normally used for the *comparison and measurement of resistances* in the range of 1 Ω to 1 MΩ. The cornerstone of the bridge consists of the four resistances (R_1, R_2, R_3, R_x), which are arranged in a diamond shape. An unknown resistance R_x is compared with others of known values, i.e., R_1 and R_2, which have fixed values, and R_3 which is variable. R_3 is varied until zero deflection is obtained on the galvanometer G. No current then flows through the meter, $V_A = V_B$, and the bridge is balanced, i.e.;

$$R_1 \cdot R_x = R_2 \cdot R_3$$

or

$$R_x = \frac{R_2 \cdot R_3}{R_1} \quad (\Omega)$$

The last equation gives R_x in terms of the other three known resistors in the bridge.

6. Cathodic-Ray Oscilloscope. The cathodic-ray oscilloscope (c.r.o.) is a useful and versatile device characterized by high input impedence and wide frequency range. It is used in the observation of waveforms and for the measurements of *voltage, current, frequency, phase* and *periodic time.*
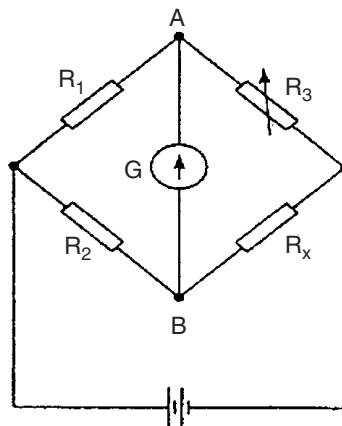
FIGURE 19.45 Wheatstone bridge.

The heart of any oscilloscope is the cathode-ray tube which is shown schematically in Fig. 19.46. Electrons are released from the hot cathode and accelerated toward the screen by the use of a positively charged anode. An appropriate grid arrangement then governs the focus of the electron beam on the screen. The exact position of the spot on the screen is controlled by the use of the horizontal and vertical deflection plates.

A voltage applied on one set of plates produces the x deflection, while a voltage on the other set produces the y deflection. Thus, with appropriate voltages on the two sets of plates, the electron beam may be made to fall on any particular spot on the screen of the tube. The screen is coated with a phosphorescent material, which emits light when struck by the electron beam.

For example, a sinusoidal waveform is displaced on a c.r.o. screen for alternating voltage measurements (see Fig. 19.47).

If the time/cm switch is on, say 5 ms/cm, then the periodic time $\bar{t}$ of the sine wave is 5 ms/cm × 4 cm, i.e. 20 ms or 0.02 s.

Then frequency
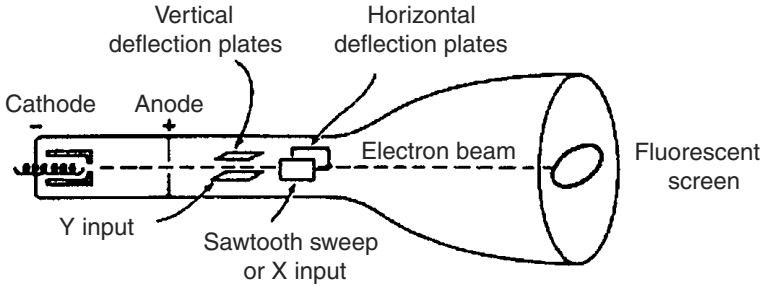
FIGURE 19.46   Schematic diagram of a cathode-ray tube.
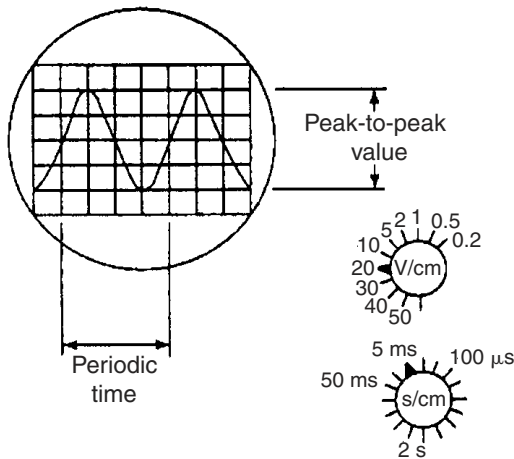
FIGURE 19.47   The c.r.o. screen measurements (example).

$$f = \frac{1}{t} = \frac{1}{0.02} = 50 \text{ Hz}$$

If the volts/cm switch is on, say 20 volts/cm, then the amplitude or peak volume of the sine wave shown is 20 volts/cm $\times$ 2 cm = 40 V.

$$\text{Therefore the r.m.s. voltage} = \frac{\text{peak voltage}}{\sqrt{2}} = \frac{40}{\sqrt{2}} = 28.28 \text{ (V)}.$$

Double beam oscilloscopes are useful whenever two signals are to be compared simultaneously.

In adjustment and use of the c.r.o., reasonable skill is needed. However, the c.r.o.'s greatest advantage is in observing the shape of a waveform—a feature not possessed by other measuring instruments.

7. Q-Meter.   The simplified circuit of a Q-meter, used for measuring Q-factor, is shown in Fig. 19.48. The Q-factor for a series L-C-R circuit is the voltage magnification at resonance, i.e.:

$$\text{Q-factor} = \frac{\text{voltage across capacitor}}{\text{supply voltage}} = \frac{V_c}{V}$$

Current from a variable frequency oscillator flowing through a very low resistance R_0 develops a variable frequency voltage, V_{R0}, which is applied to a series L-R-C circuit. The frequency is then varied until resonance causes voltage V_c to reach a maximum value. At resonance V_{R0} and V_c are noted, then:

$$\text{Q-factor} = \frac{V_c}{V_{R0}} = \frac{V_c}{i \cdot R_0}$$

In practice, V_{R0} is maintained constant. Then, if a variable capacitor C is used and the oscillator is set to a given frequency $f(= 1/2\pi\sqrt{LC})$, C can be adjusted to give resonance. In this way, inductance L can be calculated from the following relation since:
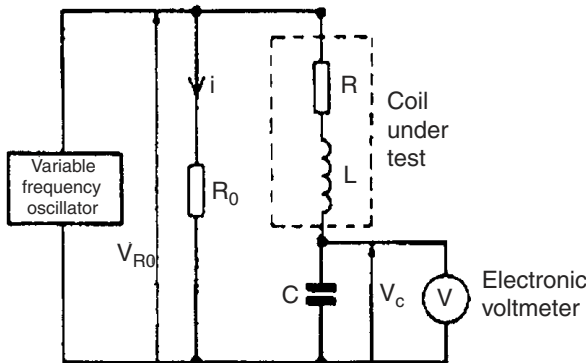
FIGURE 19.48   Q-meter circuit.

$$Q = \frac{2\pi \cdot f \cdot L}{R}$$

8. Wattmeter.   A wattmeter is an instrument for measuring electrical power in a circuit. The product of potential difference V and current i gives power P in an electrical circuit.

Fig. 19.49 shows typical connections of a wattmeter used for measuring power supplied to a load. The instrument has measuring power supplied to a load. The instrument has two coils: a) a current coil, which is connected in series with the load, like an ammeter, and b) a voltage coil, which is connected in parallel with the load, like a voltmeter.

As electrical power $P = V \cdot i = i^2 \cdot R = \dfrac{V^2}{R}$ (watts) the electrical energy (power $\times$ time) is expressed in watt $\times$ seconds (jouls). If the power is measured in kilowatts and the time in hours, then the unit of energy is kilowatt-hours. The "electricity meter" in the home records the number of kilowatt-hours used.

EXAMPLE   A piece of electrical equipment in a business office takes a current of 13 A from 220 V supply. If the equipment is used for 40 hours each week and 1 kWh of energy costs 10 cents, the cost per week of electricity is calculated as:

Power $= V \cdot i = 220 \times 13 = 2860$ W $= 2.86$ kW

Energy used per week $=$ power $\times$ time $= (2.86$ kW$) \times (40$ hr$) = 114.4$ kWh

Cost at 10 cents per kWh $= 114.4 \times 10 = 1144$ cents

Hence the weekly cost of electricity is $11.44.

Note: *Electronic measuring instruments* have advantages over instruments such as the moving-coil meters, etc., in that they have a much higher input resistance (some as high as 1,000 MΩ) and can handle a much wider range of frequency.

Errors are always introduced when using instruments to measure electrical quantities. The errors most likely to occur in measurements are those due to: a) the limitation of the instrument, b) the operator/user, c) the instrument disturbing the circuit. The calibration accuracy of an instrument depends on the precision with
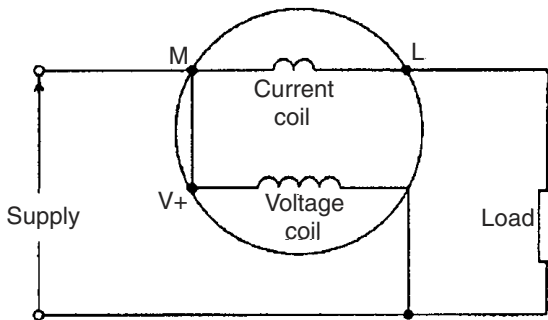
FIGURE 19.49   Wattmeter coils.

which it is constructed. Every instrument has a margin of error that is expressed as a percentage of the instrument's full scale deflection. Many industrial grade instruments have an accuracy of $\pm 2\%$.

OTHER MEASUREMENTS

1. pH Measurement.   pH is a parameter that quantifies the level of acidity or akalinity in a chemical solution. It defines the concentration of hydrogen atoms in the solution in grams/litre and is expressed as:

$$pH = \log_{10}[1/H^+]$$

where H^+ is the hydrogen ion concentration in the solution. The value of pH can range from 0, which describes extreme (pure) acidity, to 14, which describes extreme akalinity. Pure water has a pH of 7.

Litmus Paper.   The most universally known method of measuring pH is to use litmus paper that changes color according to the pH value. Unfortunately, this method gives only a very approximate inducation of pH unless used under highly controlled laboratory conditions.

The Gas Electrode.   The device known as the gas electrode is by far the most common on-line sensor used for pH measurements. The electrode consists of a glass probe containing two electrodes, a measuring one and a reference one, separated by a solid glass partition. Neither of the electrodes is in fact glass. The reference electrode is a screened electrode, immersed in a buffer solution, which provides a stable reference e.m.f. That is usually 0.0 V. The tip of the measuring electrode is surrounded by a pH-sensitive glass membrane at the end of the probe, which permits the diffusion of ions according to the hydrogen ion concentration in the fluid outside the probe. The measuring electrode therefore generates an e.m.f. proportional to pH that is amplified and fed to a display meter.

A Fiber-Optic pH Sensor.   This is another available device for pH measurement. In the solution around the probe tip the pH level is indicated by the intensity of light reflected from the tip of a probe coated in a chemical indicator whose color changes with pH (Fig. 19.50). Unfortunately, this device only has the capability to measure over a very small range of pH (typically 2 pH).

2. Sound Measurements.   Noise can arise from many sources in both industrial and non-industrial environments. Sound waves are a vibratory phenomenon. They are characterized by a pressure energy flux per unit area and per unit time as the acoustic wave moves through the medium.

The sound is measured in terms of the sound pressure level, S_P, defined as:

$$S_p = 20 \log_{10} \cdot \left(\frac{P}{0.0002}\right) \quad (dB)$$

where P is the r.m.s. sound pressure in μbar (1 bar $= 10^5$ N/m^2). dB is the sign for decibels. The quietest sound that the average human ear can detect is a tone at
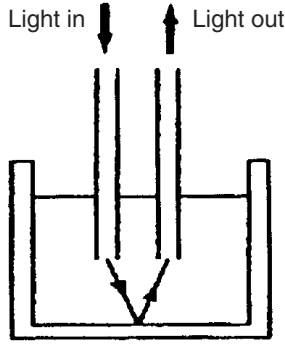
FIGURE 19.50   Sketch of the pH sensor.

a frequency of 1 kHz and the sound pressure level of 0. dB (2×10^{-4} μbar). At the upper end, sound pressure levels of 114 dB (3.45 mbar) cause physical pain. See Table 19.6 for approximate values of loudness of sound.

Sound is usually measured with a sound meter. This essentially processes the signal collected by a microphone, as shown in Fig. 19.51. The microphone is a

TABLE 19.6   Loudness of Sounds

Source	Intensity (dB)	Source	Intensity (dB)
Threshold of hearing	0	Loud conversation	70
Virtual silence	10	Door slamming	80
Quiet room	20	Riveting gun	90
Average home	30	Loud motor horn	100
Motor car	40	Thunder	110
Ordinary conversation	50	Aero engine	120
Street traffic	60	Threshold of pain	130
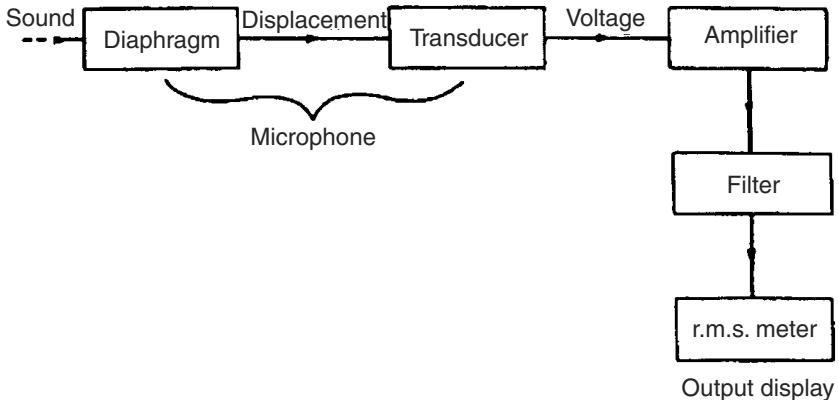
FIGURE 19.51   Sound meter.

diaphragm-type pressure-measuring device that converts sound pressure into a displacement. The displacement is applied to a displacement transducer (normally capacitive, inductive or piezoelectric type), which produces a low magnitude voltage output. This is amplified, filtered and finally given an output display on an r.m.s. meter. The filtering process has a frequency response approximating that of the human ear, so that the sound meter ''hears'' sounds in the same way as a human ear. In other words, the meter selectively attenuates frequencies according to the sensitivity of the human ear at each frequency, so that the sound level measurement output accurately reflects the sound level heard by humans.

If sound level meters are being used to measure sound to predict vibration levels in machinery, then they are used without filters, so that the actual rather than the human-perceived sound level is measured.

3. Nuclear Radiation Measurements.   Nuclear radiation is detected through an interaction of the radiation with the detecting device, which produces an ionization process. The degree of ionization may be measured with appopriate electronic circuitry. Two types of detection operations are normally performed: 1) a measurement of the number of interactions of nuclear radiation with the detector, and 2) a measurement of the total effect of the radiation. The first type of operation is a counting process, while the second operation may be characterized as a mean-level measurement. The counting operation frequently ignores the energy level of the radiation, while the mean-level measurement is used for determining the energy level of the radiation.

The Geiger-Muller Counter.   The Geiger-Muller counter is typically used for nuclear counting operations. A typical cylindrical-tube arrangement for a Geiger-Muller tube is shown in Fig. 19.52. The anode is a tungsten or platinum wire, while the cylindrical tube forms the cathode for the circuit. The tube is filled with argon and a small concentration of alcohol or some other hydrocarbon gas. The gas pressure is slightly below atmospheric. The ionizing particle or radiation is transmitted through the cathode material, or some window material installed therein, and through interaction with the gas molecules produces an ionization of the gas. If the voltage E is sufficiently high, each particle will produce a voltage pulse. The counting rate of the G-M tube is of the order of 10^4 counts/s.

Ionization Chambers.   Ionization chambers and photographic plates are used for energy-level measurements. An ionization chamber may be constructed in basically the same way as the Geiger-Muller counter (Fig. 19.52), except that the tube is
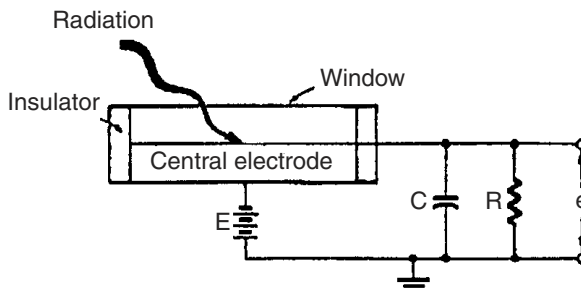
FIGURE 19.52   Typical cylindrical-tube arrangement for a Geiger-Muller counter.

operated at a much lower voltage. The arrangement may be modified to accommodate specific applications.

Photographic-Detection Methods.   When certain types of photographic films are exposed to nuclear radiation and subsequently developed, the opacity of the print may be taken as an indication of the total amount of radiation incident on the film during the time of exposure. Many specialized films are available for nuclear-radiation measurement, and this method is the primary one used for measuring the total radiation exposure for workers in atomic-energy installations. The photographic film badge may be used for detecting α and β particles, γ rays, and neutrons. In order to use a single badge for measurement of the exposure rates to the different types of radiation, several openings or windows may be constructed in a single badge. A different type of filter is placed over each window so that only one type of radiation is permitted to strike the photographic emulsion under each window. Thus, the opacity of the developed film under each of the windows gives an indication of the different radiation-exposure rates.

4. Gas Sensing and Composition Measurement.   Gas sensing and analysis is required in many applications (combustion, air-pollution, etc.).

For air pollution, measurements are important to establish acceptable levels of contamination of the atmosphere. Quantities of pollutants may be expressed on either a volumetric or a mass basis. For the mass basis, the appropriate unit would be grams per cubic meter. The volumetric unit normally employed is parts per million (ppm), which is defined as:

$$1 \text{ ppm} = \frac{1 \text{ volume of gaseous pollutant}}{10^6 \text{ volumes (air + pollutant)}}$$

or 0.0001 percent by volume = 1 ppm.

The analysis of products of combustion is especialy important not only for air-pollution-control applications but also for maintenance of most efficient burning rates and energy utilization.

Gas Chromatography.   Standard technique for measuring composition of gas samples is gas chromatography. A schematic of a gas chromatograph is shown in Fig. 19.53.
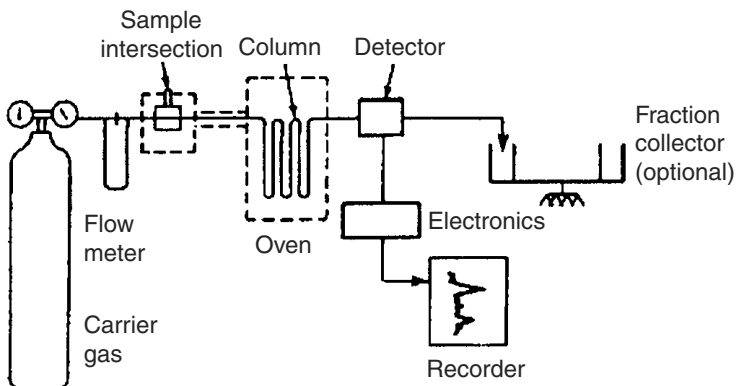
FIGURE 19.53   Schematic of gas chromatograph.

The sample gas is typically collected in a small glass syringe and introduced into the chromatographic column. The column can vary in length from 1 to 20 m and in diameter from 0.25 to 50 mm. Inside the column, either a solid absorbant, like granular silica, alumina, or carbon, or a liquid absorbant is contained to retard the flow of different constituents of the gas sample. The liquid can either coat the wall when the column is a thin capillary or be distributed over an inert granular material like diatomaceous earth in larger packed columns. Hundreds of liquids are available as absorbants, depending on the particular application.

Once the gas sample is impressed on the column, an inert carrier gas (typically helium) transports the sample through the column. Different components will be retained in the column for different lengths of time, so that they will appear in the discharge stream from the column at different times. The device may be calibrated to predict the retention time for various gas-liquid combinations. The components in the output flowstream may be detected in various ways. A thermal-conductivity detector is most common, but ultrasonic and electron-capture methods may also be used. Special detectors are available for specific elements.

The output of the chromatographic detector can be processed electronically and displayed on a chart recorder. The display takes the form of the serious spikes. The time of appearance of the spike indicates the particular component, while the height of the spike indicates the quantity present. The chromatographic column may require heating or cooling to achieve the proper separation time, so that a temperature control system is required. A flow regulator is also employed to maintain a constant flow rate of the carrier gas.

Orsat Apparatus. The simple Orsat apparatus (Fig. 19.54) is employed to analyze products of combustion. It consists of a measuring burette of three reagent pipettes, which are used to successively absorb carbon dioxide, oxygen, and carbon monoxide from the mixture. First, a sample of the combustion products is put into the measuring burette. Next, the sampling manifold is shut off and the sample is forced into the first reagent pipette, where carbon dioxide is absorbed. The sample is then brought back into the measuring burette, and the reduction in volume is noted. The procedure is then repeated for the other two pipettes, which absorb the O_2 and CO successively. In the measuring process, the combustion products are contained over water in the burette and remain saturated with water vapor. As a result, the volumetric proportions of the products are obtained on a so-called "dry basis", i.e. exclusive of the water vapor that is present. Potassium hydroxide is normally used
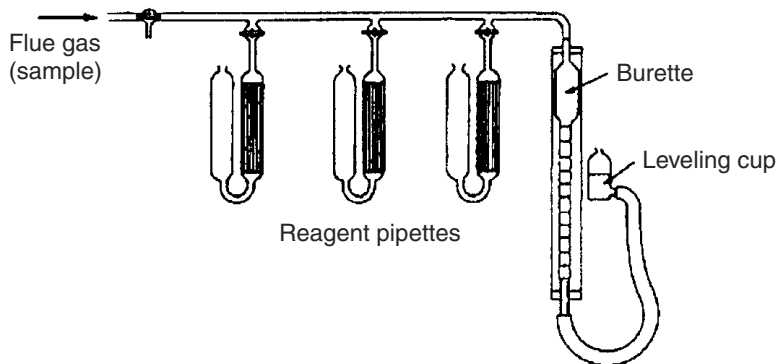
FIGURE 19.54 Schematic of Orsat apparatus.

as the reagent to absorb carbon dioxide. A mixture of pyrogallic acid and solution of potassium hydroxide is employed as the reagent for absorption of oxygen. Cuprous chloride is used to absorb carbon monoxide. Precautions must be taken to ensure that the reagents are fresh in order to minimize errors. The cuprous chloride must be changed after 10 absorptions or so if the carbon monoxide content is above 1 percent.

5. Humidity and Moisture Measurements.   The water-vapor content of air is an important parameter in many processes. The specific humidity, or humidity ratio, is the mass of water vapor per unit mass of dry air. The following reference temperatures are used for practical calculations:

- *The dry-bulb temperature*: the temperature of the air-water vapor mixture as measured by a thermometer exposed to the mixture.
- *The wet-bulb temprature*: the temperature indicated by a thermometer covered with a wick-like material saturated with liquid after the arrangement has been allowed to reach evaporation equilibrium with the mixture (Fig. 19.55).
- *The dew point of the mixture*: the temperature at which vapor starts to condense when the mixture is cooled at constant pressure.

The relative humidity (assuming vapor behaves like an ideal gas) is defined as the ratio of the actual mass of vapor to the mass of vapor required to produce saturated mixture at the same temperature:

$$\omega' = \frac{m_v}{m_{\text{sat}}} \approx \frac{p_v}{p_g}$$

where p_v is the actual partial pressure of the vapor and p_g is the saturated pressure of the vapor at the temperature of the mixture.

The specific humidity is (assuming also ideal-gas behavior):
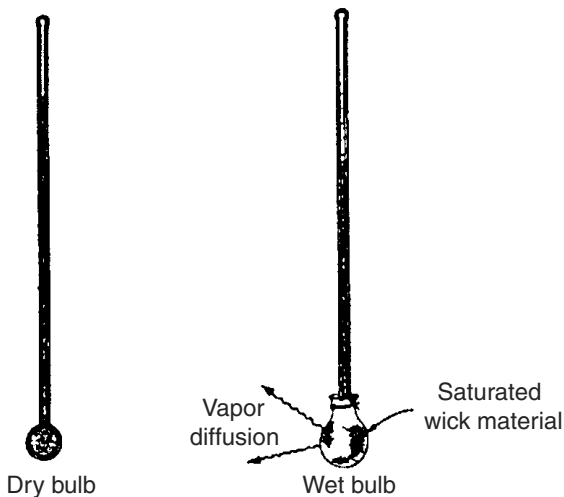
FIGURE 19.55   Measurement of dry- and wet-bulb temperature.

$$\omega'' = \frac{m_v}{m_a} \approx 0.622 \frac{p_v}{p_a}$$

where p_a is the partial pressure of the air.

To determine ω' i ω'' from the above equations p_v is calculated as[2]

$$p_v \approx p_{gw} - \frac{(p - p_{gw}) \cdot (t_{DB} - t_{WB})}{K - t_{WB}}$$

where p_{gw} = saturated pressure corresponding to wet-bulb temperature
$\quad\quad p$ = total pressure of the mixture
$\quad\quad t_{DB}$ = dry-bulb temperature, °C
$\quad\quad t_{WB}$ = wet-bulb temperature, °C
$\quad\quad K$ = 1537.8°C

Gravimetric Method. The most fundamental method for measuring humidity is a gravimetric procedure. A sample of the air-water-vapor mixture is exposed to suitable chemicals until the water is absorbed. The chemicals are then weighed to determine the amount of vapor absorbed. Uncertainties as low as 0.1 percent can be obtained with this method.

Industrial Moisture Measurement Techniques. Industrial methods for measuring moisture are based on the variation of some physical property of the material with moisture content. Many different properties can be used, and, therefore the range of available techniques is large.

Electrical Method. Microwaves at wavelengths between 1 mm and 1 m are absorbed to a much greater extent by water than most other materials. Measuring the amount of absorption of microwave energy beamed through the material is the most common technique for measuring moisture content.

Optical Methods. The refractometer is a well-established instrument that is used for measuring the water content of mixture. It measures the refractive index of the liquid, which changes according to the moisture content. Moisture-related energy absorption of near-infra-red light can be used for measuring the moisture content of solids, liquids, and gases.

Ultrasonic Methods. The presence of water changes the speed of propagation of ultrasonic waves through liquids. The moisture content of liquids can therefore be determined by measuring the transmission speed of ultrasound.

5. Viscosity Measurement. Viscosity measurement is important in many process industries. Most instruments for measuring viscosity work on one of two physical principles:

viscous friction exerted on a rotating body;

rate of flow of the liquid through a tube.

Rotating Concentric-Cylinder Viscosimeter. The rotating concentric cylinders apparatus is shown on Fig. 19.56. The inner cylinder is stationary and attached to an appropriate torque-measuring device, while the outer cylinder is driven at a constant
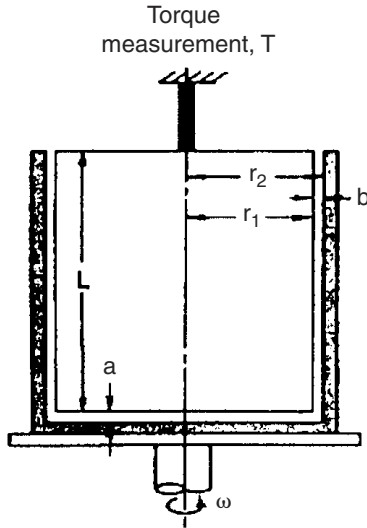
FIGURE 19.56   Rotating concentric-cylinder apparatus.

angular velocity ω. If the annular space b is sufficiently small in comparison with the radius of the inner cylinder, then the rotating-cylinder arrangement approximates the parallel- plate situation, and the velocity profile in the gap space may be assumed to be linear. Then

$$\frac{du}{dy} = \frac{r_2 \cdot \omega}{b} \tag{19.1}$$

where
$$\tau = \mu \cdot \frac{du}{dy} \quad \text{and} \quad b << r_1. \tag{19.2}$$

Now if the torque is measured, the fluid shear stress is expressed by

$$\tau = T/2\pi \cdot r_1^2 \cdot L \tag{19.3}$$

where L is the length of the cylinder (Fig. 19.56). Combining Eqs. (19.1)–(19.3)

$$\mu = T \cdot b/2\pi \cdot r_1^2 \cdot r_2 \cdot L \cdot \omega \tag{19.4}$$

The torque on the bottom disk is:

$$T_b = \frac{\mu \cdot \pi \cdot \omega}{2a} \cdot r_1^4 \tag{19.5}$$

where a is the gap spacing. Combining the last two relations

$$T = \mu \cdot \pi \cdot \omega \cdot r_1^2 \cdot \left(\frac{r_1^2}{2a} + \frac{2L \cdot r_2}{b} \right) \tag{19.6}$$

Once the torque and angular velocity are measured, then for the apparatus of specified dimensions value of viscocity μ is calculated from the last equation.

The Capillary-Tube Viscosimeter   The capillary-tube viscosimeter (the Saybolt apparatus) is an industrial device that uses the capillary-tube principle for measurement of viscosities of liquids (Fig. 19.57).

A cylinder is filled to the top with the liquid and enclosed in a constant-temperature bath to ensure uniformity of temperature during the measurements. The liquid is then allowed to drain from the bottom through the short capillary tube. The time necessary for 60 ml to drain is recorded, and this time is taken as indicative of the viscosity of the liquid. Since the capillary tube is short, a fully developed laminar-velocity profile is not established, and it is necessary to apply a correction to account for the actual profile. If the velocity profile were fully developed, the kinematic viscosity would vary directly with the time for drainage; i.e.

$$v = \frac{\mu}{\rho} = c_1 \cdot t$$

To correct for the non-uniform velocity profile, another term is added, to give:

$$v = \frac{\mu}{\rho} = c_1 \cdot t + \frac{c_2}{t}$$

With the constants inserted:

$$\frac{\mu}{\rho} = v = \left(0.00022 \cdot t - \frac{0.179}{t}\right) \cdot 10^{-3} \quad \left(\frac{\text{m}^2}{\text{s}}\right)$$

The symbol t designates the drainage time in seconds for 60 ml of liquid. ρ is the density of the test liquid.

Note: The wide range of sensors and instruments that are used for measuring physical variables and many engineering parameters are given in various references.[3,4,6,9]
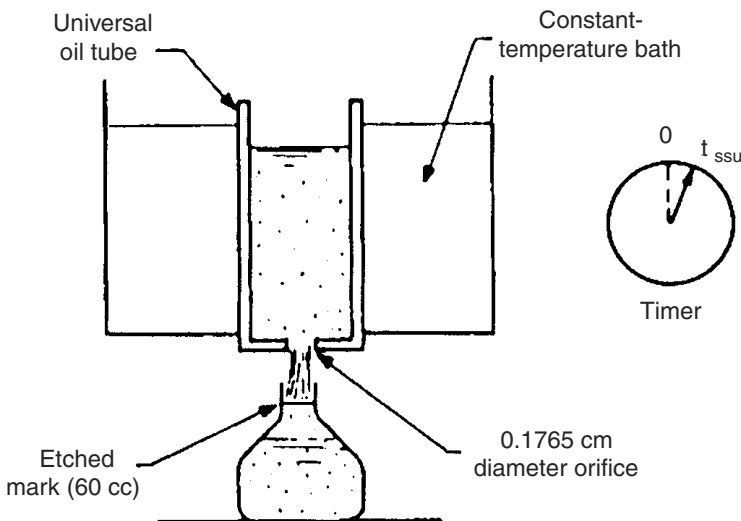
FIGURE 19.57   Schematic of the Saybolt viscosimeter.

NOMENCLATURE

Symbols = Definition, SI Units (U.S. Customary Units)
i = electrical current, A
R = resistance, Ω
V = voltage, V
B = magnetic field intensity (Wb webers/m^2)
E = electromotive force* (measured in volts)
f = frequency, Hz
t = time, s
p = pressure, N/m^2, Pa (lb$_f$/ft^2)
ω = angular velocity, s^{-1}, rad/s
τ = shear stress, N/m^2, kg$_f$/cm^2 (lb$_f$/in^2)
T = torque, N $\cdot$ m (lb$_f$ ft)
μ = viscosity (dynamic viscosity), Pa $\cdot$ s (N $\cdot$ s/m^2)
ρ = density, kg/m^3 (lb/ft^3).

REFERENCES†

1. R. A. Walsh, McGraw-Hill, *Machining and Metalworking Handbook*, 2d ed., McGraw-Hill, 1999.

2. E. R. G. Eckert and R. J. Goldstein, *Measurements in Heat Transfer,* 2d ed., Hemisphere Pub. Co., 1976.

3. R. W. Miller, *Flow Measurement Engineering Handbook,* 2d ed., McGraw-Hill, 1989.

4. R. J. Goldstein, *Fluid Mechanics Measurements,* Hemisphere Pub. Co., 1983.

5. *Fluid Meters, Their Theory and Application,* 5th ed. ASME, 1959.

6. A. S. Morris, *Measurements and Instrumentation Principles*, Butterworth-Heinemann, 2001.

7. J. P. Holman, *Experimental Methods for Engineers*, 3d ed., McGraw-Hill, 1978.

8. H. A. Rothbart, *Mechanical Design and Systems Handbook*, McGraw-Hill, 1985.

9. S. Soloman, *Sensors Handbook*, McGraw-Hill, 1999.

10. T. Baumeister, E. A. Avallone, T. Baumester III, *Mark's Standard Handbook for Mechanical Engineers,* 8th ed., McGraw-Hill, 1978. (Sections 15 and 16)

11. M. Smith, J. Hiley, K. Brown, *Hughes Electrical and Electronic Technology*, 8th ed., Prentice Hall, 2002.

12. H. Schenck, *Theories of Engineering Experimentation*, 2d. ed., McGraw-Hill, 1968.

13. E. O. Doebelin, *Measurement Systems: Application and Design*, 2d ed., McGraw-Hill, 1975.

---

*Provided by a source of energy such as a battery or generator is measured in volts.

†Most figures presented in this chapter are redrawn and upgraded from similar figures and sketches found in listed references. Figures 19.12, 19.27, 19.30, 19.31, 19.33, 19.56 are related to similar ones found in reference 7. Figures 19.32 and 19.41 are similar to those from reference 10. Figures 19.5, 19.18 and 19.26 are based on reference 16. Figures 19.29, 19.34 and 19.48 are based on reference 15. Data collected in Tables 19.1–19.6 are from references 1, 7, 8, and 16. Those references listed here but not cited in the text were used for comparison, clarification, clarity of presentation, and, most important, reader's convenience when further interest in the subject exists.

14. D. N. Keast, *Measurements in Mechanical Dynamics*, McGraw-Hill, 1967.

15. J. Bird, *Newnes Engineering Science*, Newnes, 2001.

16. J. Carvill, *Mechanical Engineer's Data Handbook*, Butterworth-Heinemann, 2001.

17. C. G. Peters and W. B. Emerson, *J. Res. Natl. Bur. Std.*, Vol. 44, p. 427, 1950.

18. J. G. Anderson, *J. Measurements and Control*, Vol. 22, p. 82, 1989.

19. G. W. Ewing, *Instrumental Methods of Chemical Analysis*, 4th. ed., McGraw-Hill, 1975.

20. J. F. McPartland and B. J. McPartland, McGraw-Hill, *National Electrical Code Handbook*, 23d ed., McGraw-Hill, 1999.

21. J. Whitaker and B. Benson, *Standard Handbook of Video and Television Engineering*, 3d. ed., McGraw-Hill, 2002.

22. F. S. Merritt, M. K. Loftin and J. T. Ricketts, *Standard Handbook for Civil Engineers*, 4th ed., McGraw-Hill, 1996.

23. R. A. Corbitt, *Standard Handbook of Environmental Engineering*, 2d ed., McGraw-Hill, 1999.

24. D. Christiansen, D. G. Fink, R. K. Jurgen and E. A. Torrero, *Electronics Engineers' Handbook*, 4th ed., McGraw-Hill, 1997.

25. K. S. Lion, *Instrumentation in Scientific Research,* McGraw-Hill, 1959.

CHAPTER 20
ENGINEERING ECONOMY, PATENTS, AND COPYRIGHTS

ENGINEERING ECONOMY*

Engineering economy is a study of the time value of money in an engineering environment. Thus the engineer might study—from an economic standpoint—the investment differences of different types of energy supplies, the relative cost of two different types of heat insulation, or the relative costs of two highway materials.

In any engineering economic study, the following factors will enter: (1) the first cost of each alternative way of accomplishing a given task or reaching an end result, (2) the cost of borrowed money, (3) the time required to recover the investment, (4) operating and maintenance costs, if any, of each alternative, and (5) any other relative costs associated with the project or task being considered.

Since "it always comes down to money," engineering economy is finding much wider application today than ever before. And with the ready availability of computers of all sizes to do the "numbers crunching," almost every small and medium-sized project is subjected to a rigorous economic analysis *before* final approval to move ahead. For this reason, every engineer must feel comfortable with the fundamentals of engineering economics. And, of course, *every* major engineering project is given a searching economic study—all the more reason why the engineer must know the fundamentals of the subject. They are presented in this section of this handbook.

In somewhat different terms, engineering economics is the search for and recognition of alternatives which are then compared and evaluated in order to come up with the most practical design and creation. The primary objective of this chapter is to provide the principles, concepts, techniques, and methods by which alternatives within a project† can be compared and evaluated for the best monetary return. If economic criteria are not considered properly when profit is the ultimate objective, the result is bad engineering. Many technologically brilliant projects have been destroyed as a result of unsound economic analysis.

BASIC CONCEPTS

1. Objectives of Economic Analysis.   The primary objective of any economic analysis is to identify and evaluate the probable economic outcome of a proposed project so that available funds assigned to it may be used to optimum advantage.

The analysis is always made from the viewpoint of the owner of the project and usually involves a comparison of alternatives on a monetary basis. It should be recognized that an action always involves at least two possible courses: doing it or not doing it. If the analysis is to yield results, the criteria by which alternatives are evaluated should have the following objectives:

1. Profit maximization
2. Cost minimization
3. Maximization of social benefit
4. Minimization of risk of loss
5. Maximization of safety, quality, and public image

The preceding list is by no means complete. This chapter will deal primarily with the first three; however, when applicable, the others will not be neglected.

2. Procedure for Economic Analysis.   An economic analysis revolves around three processes: *preparation, analysis,* and *evaluation.* The process of preparation can be broken down into three steps: understanding the project, defining the objective, and collecting data. Similarly, analysis involves analysis of data, interpretation of results, and formulation of alternative solutions. Finally, there are two steps to evaluation: evaluation of the alternatives and identification of the best alternative(s).

Analysis and evaluation can be handled by computer. In fact, nearly all medium- and large-size firms in the United States have this type of capability, yielding quick feedback for decision making. Decision making can be considered to be the fourth process involved in economic analysis. The eight steps that make up preparation, analysis, and evaluation will now be briefly discussed.

1. *Understanding the project:* One cannot and should not attempt to perform an economic analysis without clearly understanding the project. This is often a substantial difference between what one thinks the project is and what the project really is.

2. *Defining the objectives:* Failure to clarify the objectives of a project will often create dissatisfaction. The objectives must be clearly stated and compatible with each other. Some examples are

- Meet or exceed a specified minimum rate of return on an investment
- Return the investment (break even) within a specified time period
- Obtain a specific share of a market

The firm's objectives and criteria must be specified and defined quantitatively. That is, for example, if return of the investment within a specified time period is sought, the piece of equipment or operation being investigated should pay for itself within the specified time period, based on the potential savings or profits realized through its use.

3. *Collecting data:* This begins with a complete review of published literature, if available, or of historical data in the firm's files. Often, the wheel gets reinvented only because someone did not bother to search out existing information. If not otherwise available, data can be obtained through private sources or roughly estimated through assumptions.

4. *Analysis of data:* This is the process of converting developed data into something meaningful and useful. The use of a computer is highly beneficial. Often the computer can generate maximum amounts of information at a minimum cost.

5. *Interpretation of results:* Interpretation of the results usually occurs upon completion of the analysis. The results must be well organized, stored properly, and then carefully adapted and utilized in the evaluation phase.

6. *Formulation of alternative solutions:* Different avenues leading to the same final objective are to be investigated and proposed to management as alternative methods. Hence, in the case of a new product, for example, different methods of manufacturing would be proposed or different levels of automation would be suggested; then their effect on final project profitability would be examined.

7. *Evaluation of alternatives:* In evaluating the alternatives, it is important to use uniform criteria for all, not different ones for each alternative. If the return-on-investment concept is to be used on one alternative, then each alternative is to be evaluated by the return-on-investment concept. Furthermore, the method of calculation of the return on investment (ROI) must be uniform as well.

8. *Identification of the best alternative(s):* In most firms, this task is performed by top management; accordingly, the analyst must narrow the choice down to the two or three best possibilities. Top management can then identify the one that comes closest to meeting the objectives.

As soon as the analysis, evaluation, selection, and approval of the best alternative have been completed, the implementation process begins. The engineer or technologist will design, procure, and then install the equipment or assets called for in the project. He or she will often be present during the start-up of the operation and make revisions if necessary.

3. Capital Expenditure Policies.   Wear and tear of productive facilities necessitates their eventual replacement. Industrial and consumer demands for more goods and services necessitate an increase in supply and hence an increase in productive facilities. Firms respond to these pressures through capital expenditures by investing new plants, equipment, and products. Often, capital spending can be minimized by sacrificing some output capabilities or through productivity improvement or cost reduction opportunities.

Basically, capital expenditures can be classified into five general groups:

1. Maintenance of productive facilities
2. Optimization of existing productive capacity
3. Mechanization or automation of existing facilities
4. Expansion of product lines or productive capacity
5. Necessities due to governmental regulations

For whichever purpose the expenditure is made, except in the case of number 5, the final criterion is the profit or savings to be realized through the modification.

4. Basic Concepts and Assumptions.   In order to make an economic evaluation of a project, certain basic information must be established. More detailed techniques will be considered later.

Revenue R refers to any increase in the owner's equity resulting from sales or services of business.

Gross profit G (also referred to as *gross income*) is the yearly earning from a venture throughout its operating life. It is equal to revenue minus raw materials cost, operating expenses including overhead, maintenance, labor, and social security and unemployment taxes. It does not include deductions for depreciation and income tax. It can be expressed as $G = R - C$, where C is the various costs listed above.

Breakeven is a situation at which gross profit G is equal to zero. Stated differently, breakeven occurs where revenue R from sales or services just equals the costs C associated with doing business, the ones mentioned in the explanation of gross profit. Hence, $G = R - C = 0$, or $R = C$.

Fixed capital investment I consists of the investment in facilities and equipment.

Working capital I_w is money tied up in raw materials, intermediate- and finished-goods inventories, and accounts receivable, as well as cash needed to operate a given project.

Income tax rate t is a government tool for controlling inflation. A high rate decreases the supply of money available for business investment and spending. The federal income tax rate can be taken at 48 percent, while the state income tax rate can be taken at an average of 5 percent.

Depreciation d consists of a fixed annual charge on the facility or equipment investment which will result in recovery of the initial investment at the end of the useful life of the item. If the actual life of the facility or equipment is known, an exact rate of depreciation can be established where the sum of the rates will just equal the investment.

Interest i is the rental charge for the use of borrowed money. It is another inflation controller. High interest rates discourage borrowing, making new investments less desirable.

Net profit P is equal to gross profit minus depreciation, interest, and income tax. It can be expressed as

$$P = G - i(I + I_w) - t(G - dI) \tag{20.1}$$

where I = capital investment in facilities or equipment (fixed capital)
I_w = working capital
G = gross profit
i = interest rate
d = depreciation rate
t = income tax rate (federal and state)

Rate of return on investment (ROI) is the annual rate of return on the original investment and can be expressed as

$$\text{ROI} = \frac{\text{net profit}}{\text{total investment}} = \frac{P}{I + I_w} \tag{20.2}$$

This ratio, often multiplied by 100 to yield a percentage, is the simplest and perhaps the most widely used index for measuring the attractiveness of a venture.

Payout, or payback, time (PT) is another form of measuring the attractiveness of a venture. It is the ratio of capital investment to yearly net profit and can be expressed as

$$\text{PT} = \frac{I}{P} \tag{20.3}$$

where PT is given in years. It should be noted that payout, or payback, time is based only on I rather than on total capital investment, that is, $I + I_w$.

Discounted rate of return is the rate at which the sum of future profits equals the total capital investment (fixed plus working). It can be expressed mathematically as follows:

$$(I + I_w) = \frac{P_1}{(1 + i)^1} + \frac{P_2}{(1 + i)^2} + \cdots + \frac{P_n}{(1 + i)^n} \tag{20.4}$$

where I = fixed capital investment
I_w = working capital
P_1 = net profit or saving at the end of first year
P_2 = net profit or saving at the end of second year
P_n = net profit or saving at the end of year n
i = after-tax interest rate, which is found by trial-and-error

Discounting is done because of the fact that future profits generated by a project will decrease in value over time due to inflation. This topic is discussed in detail later, pp. 20.29.

Minimum return rate is the minimum acceptable rate of annual return on investment, or minimum return rate ROI_m set by the firm. New projects, no matter how technologically sound, must show at least that rate of return, after taxes, before they can be considered. The minimum return rate is established based on the following variables: (1) the cost of borrowed money or the interest the firm must pay for the use of someone else's money, (2) 5-year average return on shareholders' equity,* (3) potential risk of failure associated with given projects. In general, most firms employ (1) plus (2) if a project or series of projects is relatively familiar to the firm. They add on (3) if the project is a totally new venture for the firm. Hence, if the going interest rate is 7 percent per year, and if the firm's average return on shareholders' equity is 13 percent, then the minimum acceptable return rate is set at 20 percent. On the other hand, if there is a substantial risk involved with a project, the ROI_m might be raised to 30 or 40 percent.

---

*Average return on shareholders' equity can be defined as the net income divided by the average shareholders' investment (stocks, bonds, etc.). Mathematically it can be expressed as

$$\frac{\text{Net profit}}{\text{Average shareholders' investment}} \times 100$$

5. Engineering Economics and Social Values. If engineers or technologists are to be more than just technicians, they must look beyond the profitability of a venture. They must look forward to the possible social implications of the operation.

These social and ethical concerns, compared to technical matters, are not so readily formalized or calculated. Nevertheless, they should not be ignored, since the firm's reputation and image may be at stake. Degree of automation, for instance, should be carefully analyzed, for it may have an important bearing on labor problems. Pollution abatement, on the other hand, is just as essential as monetary return on investment and should always be incorporated when economic studies are performed: to neglect this is to neglect true engineering economics.

COST ESTIMATING

1. Types of Estimates and Their Costs. There are a number of methods available for cost estimation, ranging in accuracy from a rough estimate to a detailed estimate derived from drawings, blueprints, and specifications. The choice of method depends on the purpose of the estimate. In general, three types of estimate are employed by most industries:

1. Order-of-magnitude
2. Semidetailed (budget-authorization estimate)
3. Detailed (firm estimate)

These estimates are used for feasibility studies, selection among alternative investments, appropriation-of-funds requests, capital budgeting, and presentation of fixed-price bids, to name but a few.

Regardless of the estimating method being employed, it is important to recognize that the level of detail carries a price tag directly proportional to its level of accuracy and to the time required to prepare the data as input for the estimate. Simultaneously, the level of detail and level of accuracy are directly proportional to the quality and quantity of the output which is the final estimate.

Table 20.1 depicts the relationship between the cost of the estimate as a percentage of total project cost and the probable accuracy of the estimate, based on total project costs of $500,000, $1,000,000, $5,000,000, $10,000,000, $15,000,000, and $20,000,000. As can be seen, estimates on larger projects require a lower expenditure per project dollar, while smaller projects require a higher percentage. Hence, from Table 20.1, an estimate accurate to within ± 10 percent* of a $500,000 project would be expected to be around 1.6 percent or $8000. On a $20,000,000 project, an estimate accurate to within ± 10 percent would be expected to be around 0.30 percent of the total project cost, or $60,000. Similarly, on a $5,000,000 project, a ± 10 percent accurate estimate would be about 0.46 percent of the total project, or $23,000.

2. Order-of-Magnitude Estimates. The order-of-magnitude cost estimates usually have an average accuracy level of ± 50 percent, often varying from ± 30 to ± 70 percent, depending on the size of the project. Such estimates, as mentioned

---

*The plus (+) means that the estimate is below the actual costs, while the minus (−) means that the estimate is above those costs.

TABLE 20.1 Estimation of the Cost of Cost Estimating as a Percentage of Total Project Cost

Level of accuracy, %	Total cost of the project					
	$500,000	$1,000,000	$5,000,000	$10,000,000	$15,000,000	$20,000,000
5	4.00	3.70	1.00	0.80	0.70	0.65
10	1.60	1.50	0.46	0.40	0.34	0.30
15	0.76	0.70	0.21	0.19	0.17	0.15
20	0.44	0.37	0.13	0.11	0.10	0.08
25	0.28	0.24	0.08	0.07	0.07	0.05
30	0.21	0.17	0.06	0.05	0.05	0.04
35	0.16	0.13	0.04	0.04	0.03	0.03
40	0.12	0.10	0.03	0.03	0.02	0.02
45	0.10	0.08	0.03	0.02	0.02	0.02
50	0.08	0.06	0.02	0.02	0.01	0.01

Used by permission of NL Industries, Inc.

previously, require much less detail than firm estimates, making them least costly to prepare but most risky in terms of over- or underexpenditure. Nevertheless, estimates of this type are extremely important for determining if a proposed project should be given further consideration or for screening a large number of alternative projects in a short period of time. With order-of-magnitude cost estimates, precision is sacrificed for speed; information quickly becomes available to show whether expected profit is sufficient to justify the risk of investment or, simply, whether the project warrants further consideration.

Order-of-magnitude estimates are usually derived from cost indexes, cost ratios, historical data, experience, or physical dimensions.

Cost Indexes. Most cost data which are available for immediate use in order-of-magnitude estimates are historical, that is, based on conditions at some time in the past. Since the value of money depreciates continuously as a function of time, this means that all published cost data are out of date. Some method must be used for converting past costs to present costs. This can be done by the use of a cost index which gives the relative cost of an item in terms of the cost at some particular base period.

If the cost at some time in the past is known, the equivalent cost at the present time can be obtained by multiplying the original cost by the ratio of the present index value to the index value at the time of original cost. Mathematically, this can be expressed as follows:

$$\text{Present cost} = (\text{original cost}) \frac{(\text{index value at present})}{(\text{index value at time of original cost})} \quad (20.5)$$

Expressed differently,

$$\text{Cost in year } B = \frac{(\text{index value at year } B)(\text{cost in year } A)}{(\text{index value at year } A)}$$

Cost indexes can be used to give a general estimate, but no index can account for all economic factors, such as changes in labor productivity or local conditions. The common indexes permit fairly accurate estimates ±10 percent at best, if the

period involved is less than 10 years. For periods greater than 10 years, the accuracy falls off rapidly.

There are many types of cost indexes* covering every area of interest: equipment, cost, labor, construction, raw materials, etc. The most common of these indexes are Marshall and Stevens equipment cost indexes, and *Chemical Engineering* plant cost indexes.†

Cost-Capacity Relationship.   Cost estimates can be rapidly approximated where cost data are available for similar projects of different capacity. In general, costs do not rise linearly; that is, if the size doubles, the cost will not necessarily increase twofold. The reason for this is that the fabrication of a large piece of equipment usually involves the same operations as a smaller piece, but each operation does not take twice as long; further, the amount of metal used on a piece of equipment is more closely related to its area than to its volume. Accordingly, the relationship can be expressed mathematically as

$$C_B = C_A \times \left(\frac{Q_B}{Q_A}\right)^X \tag{20.6}$$

where C_B = cost at capacity B
C_A = cost at capacity A
Q_B = quantity or capacity B
Q_A = quantity or capacity A
X = cost-capacity factor

The component X in the preceding equation, the *cost-capacity factor,* varies according to the type of project being considered. The range is from 0.2 to 1.00, the average, however, being 0.6 to 0.8. Steam electric generating plants, for example, have a factor of about 0.8. Waste-treatment plants usually range between 0.7 and 0.8. Large public housing projects also average about 0.8. On the other hand, steel storage tanks have a cost-capacity factor as low as 0.4 or as high as 0.8, depending on their shape. In the absence of other information, a factor of 0.75 can be used.

In general, the cost-capacity concept should not be used beyond a tenfold range of capacity, and care must be taken to make certain that the two capacities or equipments are similar with regard to construction, materials of construction, location, and other pertinent variables such as time reference. In industry it is a standard procedure to limit scaling to capacity ratios of 2:1 and in some extreme cases 3:1. Should two different time references be used, convert the cost of the previous project to a current basis, using an appropriate index to correct historical costs for time differential.

3. Semidetailed Estimates.   Semidetailed, or budget, estimates are on average accurate to within ±15 percent, ranging between ±10 and ±20 percent. For most projects this level of accuracy is quite adequate for decision making, giving the potential investor enough information to decide whether or not to proceed. During selection from among alternative investments, if order-of-magnitude estimates still yield two or more alternatives, semidetailed estimates are then employed for further screening.

---

*For a detailed summary of various cost indexes, see *Engineering News-Record,* 180: 77–88, 1968.
†Published in *Chemical Engineering,* a McGraw-Hill publication.

Semidetailed estimates, which are most frequently applied for preparing definitive estimates, require more information than do order-of-magnitude estimates. Instead of using mathematical relationships, historical costs, or project similarity, the project must be considered on its own. Actual quotations are to be obtained on major equipment and major related items. Equipment installation labor is evaluated as a percentage of the delivered equipment costs. Preliminary design data are usually necessary along with some drawings from which costs for concrete, steel, piping, instrumentation, etc. are obtained. Unit costs are then applied to the measured units. For example, schedule 40 1-in 316 stainless steel piping costs $5.16 per linear foot. A percentage of delivered equipment cost is often used instead of measured units to achieve the same goal. Table 20.2 shows an example of this.

Estimating by percentage of delivered-equipment cost is commonly used for preliminary and budget estimates. The method yields highly accurate results when applied to projects similar in nature to *recently* completed ones.

It should be noted that in the preceding analysis a 10 percent contingency allowance has been added as a buffer to reflect possible inaccuracy or inflation. In general, for a project with engineering substantially completed and with major pieces of equipment priced (vessels, pumps, conveyors, heat exchangers, processing equipment, instrumentation and controls, etc.) a contingency allowance of 5 percent is added. For projects for which engineering is 15 to 25 percent completed and for which major items of equipment have been estimated with 50 percent covered by firm quotes, a contingency allowance of 10 percent is added. For projects for which engineering is less than 10 percent completed and for which major items of equipment have been estimated with quotes for less than 50 percent, a contingency allowance of 15 percent is added. Finally, for projects that have been scoped, for which only preliminary engineering has been completed, and for which major equipment items have been specified but substantially no design has been undertaken, a contingency allowance of 20 percent is added.

4. Detailed Estimates.   Detailed cost estimates should have an accuracy of between 5 and 10 percent. They thus require careful determination of each individual item in the project or detailed itemizing of each component making up the cost. Facilities, equipment, and materials needs are determined from complete engineering drawings and specifications and are priced either from up-to-date cost data or, preferably, from firm-delivered quotations. Installation costs are computed from up-

TABLE 20.2   Percent of Delivered Cost Tabulation

Components	Percent of delivered equipment cost
Purchased (major) equipment installation	34
Piping, installed	20
Instruments, installed	5
Electrical, installed	4
Buildings	5
Utilities, installed	5
Raw materials storage, installed	2
Engineering, overhead, etc.	15
Contingencies	10
Total	100

to-date labor rates, efficiencies, and worker-hour calculations. These estimates, however, as seen from Table 20.1, are time-consuming and costly to prepare and should be used only when absolutely necessary. In fact, they are almost exclusively prepared by contractors bidding for a given job; however, these bids are often accurate to within ± 5 percent or better. If the bid is too high, the job may not be awarded. If the bid is too low, the job, although awarded, will produce a loss to the contractor. Therefore, it is absolutely necessary to be as accurate as possible.

5. Capital-Investment Cost Estimation.   The fixed capital requirements of a new project can be broken down into three components for estimating purposes: (1) depreciable fixed investment, (2) expensed or amortized* investment, and (3) nondepreciable fixed investment. *Depreciable fixed investment* can be further broken down into buildings and services; equipment, including installation; and other items such as transportation and shipping and receiving facilities. The *amortized investment* consists of research and development, engineering and supervision, startup costs, and other things, including franchises, designs, and drawings. The *nondepreciable fixed investment* also has identifiable parts. The two parts are land and working capital.

Most of the preceding components are self-explanatory. For some, further elaboration follows.

Research and Development.   This usually pertains to new manufacturing plants, where expenditures for such activities are usually associated with process-improvement and cost-reduction efforts aimed at increasing productivity and profits. Research and development costs average 3 to 4 percent of sales or services associated with the project. For large pharmaceutical companies, these costs may be as high as 8 to 10 percent.

Engineering and Supervision.   This is the cost for construction design and engineering, drafting, purchasing, accounting, travel, reproduction, communications, and various office expenses directly related to the project. This cost, since it cannot be directly charged to equipment, materials, or labor, is typically considered as an indirect cost ranging from 30 to 40 percent of the purchased-equipment cost or 10 to 15 percent of the total direct costs of the project.

Startup Cost.   After a project has been completed, a number of changes usually have to be made before the project can operate at an optimum level. These changes cost money for equipment, materials, labor, and overhead. They result in loss of income while the project is not producing or is operating at only partial capacity. These costs may be as high as 12 percent of the fixed-capital investment, although they usually stay under 10 percent. In general, an allowance of 10 percent for startup cost is quite satisfactory.

Land.   This is the cost for land and the accompanying surveys and fees, which usually amounts to 4 to 8 percent of the purchased-equipment cost or 1 to 2 percent of the fixed-capital investment. Because the value of land usually appreciates with

---

Amortization is a form of depreciation applicable to intangible assets such as patents, copyrights, franchises, etc. Generally, straight-line depreciation methods must be used, and only certain items that are amortized can be deducted as expenditures for federal income tax purposes.

time, this cost is not included in the fixed-capital investment when estimating certain operating costs, such as depreciation.

Working Capital.   This consists of the total amount of money invested in raw materials; intermediate and finished-goods inventories; accounts receivable; cash kept on hand for monthly payment of operating expenses such as salaries, wages, and raw materials purchases; accounts payable; and taxes payable.

The raw materials inventory usually amounts to a 1-month supply of the raw materials valued at delivered prices. Finished products in stock and intermediate products have a value approximately equal to the total manufacturing cost for 1 month's production or service. Credit terms extended to customers and from suppliers are usually based on an allowable 30- to 45-day payment period (accounts receivable and accounts payable, respectively). The cost of working capital varies from 10 to 25 percent of fixed-capital investment and may increase to as much as 50 percent or more for firms producing products of seasonal demand because of large inventories that must be carried for long periods of time.

6. Operating-Cost Estimation.   Determination of the necessary capital investment (fixed and working) for a given project is only one part of a complete cost estimate. Another equally important part is the operating-cost estimation. *Operating cost, production cost,* or *manufacturing cost* is the cost of running a project, a manufacturing operation, or a service. In this section and throughout this chapter the three costs are considered together.

Accuracy is as important in estimating operating cost as it is in estimating capital investment. The largest cause of error in operating-cost estimation is overlooking elements that make up the cost. Accordingly, it is very. useful to break down operating cost into its elements, as shown below. This breakdown then becomes a valuable checklist to preclude omissions.

1. Direct (variable) operating costs
   - *a.* Raw materials
   - *b.* Operating labor
   - *c.* Operating supervision
   - *d.* Power and utilities
   - *e.* Maintenance and repairs
   - *f.* Operating supplies
   - *g.* Others: laboratory charges, royalties, etc.

2. Indirect (fixed) operating costs
   - *a.* Depreciation
   - *b.* Taxes (property)
   - *c.* Insurance
   - *d.* Rent
   - *e.* Other: interest

3. General overhead costs
   - *a.* Payroll overhead
   - *b.* Recreation

c. Restaurant or cafeteria

d. Management

e. Storage facilities

4. Administrative costs

a. Executive salaries

b. Clerical wages

c. Engineering and legal costs

d. Office maintenance

e. Communications

5. Distribution and marketing costs

a. Sales office

b. Sales staff expenses

c. Shipping

d. Advertising

e. Technical sales service

As can be seen, operating costs fall into two major classifications: direct and indirect. *Direct costs* (also called *variable costs*) tend to be proportional to production or service output. *Indirect costs* (also called *fixed costs*) tend to be independent of production or service output. Some costs are neither fixed nor directly proportional to output* and are known as *semivariable costs.* Direct and indirect costs are usually estimated on a basis of cost per unit of output or service and can generally be regarded as linear over a wide range of production or service volume.

The other costs—overhead, administrative, and distribution and marketing—are expressed on a time basis, since they are related to the level of investment rather than to the level of output. The period is usually 1 year because (1) the effect of seasonal variation is evened out, (2) this permits rapid calculations at less than full capacity, and (3) the calculations are more directly usable in profitability analysis.

The best source of information for an operating-cost estimation is data from similar or identical projects. Most firms have extensive records of their operations, permitting quick estimation from existing data. Adjustments for increased costs resulting from inflation must be made, and differences in size of operation and geographical location must be considered.

Methods for estimating operating cost in the absence of specific information are discussed below.

Direct (Variable) Costs

Raw Materials. The amount of raw materials required per unit of product can usually be determined from literature, experiments, or process material balances. Credit is usually given for byproducts and salvageable scrap. In many cases, certain materials act only as an agent of production and may be recovered to some degree.

---

*For example, one supervisor is able to oversee the workers in a department up to a certain level of production. If additional workers are added to the department, the point is eventually reached where it is necessary to hire another supervisor.

Accordingly, the cost should be based only on the amount of raw materials actually consumed.

Direct price quotations for raw materials from prospective suppliers are preferable to published market prices. For an order-of-magnitude cost estimate, however, market prices are often sufficient.

Freight or transportation charges should be included in the raw materials costs, and these charges should be relevant to where they are to be used. For example, if raw materials purchasing is centralized and then the materials are dispersed to various locations, the added freight cost from central point to final destination also must be included.

Operating Labor. The average rate for labor in different industries at various locations can be obtained from the U.S. Department of Labor, Bureau of Labor Statistics, *Monthly Labor Review.* Depending on the industry, operating labor may vary from 5 to 25 percent of the total operating cost.

The most accurate way to establish operating labor requirements is to use a complete manning table, but shortcut methods are available and are quite satisfactory for most cost estimates. One technique suggests that labor requirements vary to about the 0.20 to 0.25 power of the capacity ratio when plant capacities are scaled up or down. Hence, Eq. (20.6) (capacity relationship) can be employed. The equation recognizes the improvement in labor productivity as plants increase in output and the lowering of labor productivity as output decreases, and it can be used to extrapolate known worker hours or cost of labor from one operation to another of a different capacity. The operations must, however, be similar in nature.

Operating Supervision. The cost for direct supervision of labor is generally estimated at 15 to 20 percent of the cost of operating labor.

Power and Utilities. The cost for utilities, such as steam, electricity, natural gas, fuel oil, cooling water, and compressed air, varies widely, depending on the amount of consumption, the location, and the source. In Niagara Falls, New York, for instance, electric power is relatively cheap compared to other locations. Natural gas is relatively cheap in the Gulf Coast states.

As a rough approximation, power and utility costs for a manufacturing facility amount to 10 to 20 percent of the operating cost. Utility consumption does not vary directly with output level, and variation to the 0.9 power of the capacity ratio is a good relationship.

Maintenance and Repairs. Maintenance and repairs are necessary if a plant, an office, a warehouse, a manufacturing facility, etc., is to be kept in efficient operating condition. These costs include the cost for labor, materials, and maintenance supervision.

Records for the firm's existing operations are the only reliable source of maintenance cost, but with experience it can be estimated as a function of investment. Maintenance cost as a percentage of fixed-capital investment ranges between 3 and 15 percent, with the average between 8 and 10 percent.

For operating rates of less than full capacity (100 percent), the following is generally true: For a 75 percent operating rate, the maintenance and repairs cost is about 85 percent of its full-capacity cost; for a 50 percent operating rate, the maintenance and repair cost is about 75 percent.

Operating Supplies. These are supplies such as charts, janitorial supplies, lubricants, etc., which are needed to keep a project functioning efficiently. Since these cannot be considered as raw materials or maintenance and repairs materials, they are classified as operating supplies. The annual cost for operating supplies is approximately equal to 2 percent of the total investment.

Others.   Charges for laboratory facilities, patents, royalties, rentals (for copying machines, typewriters, and machinery), etc., can range between 2 and 10 percent of operating labor, selling price, or operating cost, depending on the particular situation.

Indirect (*Fixed*) Costs

Depreciation.   Equipment, buildings, and other material objects require an initial investment, which must be written off as an operating expense. In order to write off this cost, a decrease in value is assumed to occur throughout the useful life of the material assets. The decrease in value is termed *depreciation.*

The annual depreciation rate on a straight-line basis for machinery and equipment is generally about 5 percent of the fixed-capital investment; for buildings the rate is about 2.5 percent, while for equipment used for research and development the rate is 10 percent. For pollution-abatement equipment the rate is about 20 percent.

Property Taxes.   The amount of local property taxes is a function of the location of the operation and the regional laws. For highly populated areas, the range on an annual basis is 2 to 5 percent of the fixed-capital investment, while for less-populated areas local property taxes are about 1 to 2 percent.

Insurance.   Insurance rates depend on the type of operation and the extent of available protection facilities. Generally, these rates annually amount to 1 to 2 percent of the fixed-capital investment.

Rent.   Annual costs for rented land and buildings are about 8 to 12 percent of the value of the rented property.

Interest.   Since borrowed capital (fixed and working) is usually used to finance a project, interest must be paid for its use. Fluctuations in interest rates are a function of the state of the economy and can range between 7 and 12 percent, depending on the borrower (size and type of firm), with the average being 8 percent.

Overhead Costs

General Overhead.   General overhead costs usually include payroll, recreation, restaurant or cafeteria, management, and storage facilities as a lump sum. The costs range between 40 and 70 percent, with the average being 50 percent, of the total cost for operating labor, operating supervision, and maintenance and repairs.

Administrative Costs.   The salaries for top management, such as the director of manufacturing, the technical director, the vice president for operations, etc., are not a direct manufacturing cost. Still, they must be charged to administrative costs along with clerical wages, engineering and legal costs, office maintenance expenses, and communications costs (telephone, teletype), which are all part of the operating cost. These costs may vary markedly from operation to operation and depend somewhat on whether the operation is a new one or an addition to an existing one. In the absence of accurate cost figures from records, or for a quick cost estimate, the administrative costs may be approximated as 40 to 60 percent of the operating labor, with the average about 45 percent.

Distribution and Marketing Costs.   The costs of selling a product or a service provided by a successful operation are charged to distribution and marketing costs. Included in this category are salaries, wages, supplies, and other expenses of sales offices; salaries, commissions, and travel expenses for salespeople; and expenses

for shipping, containers, advertising, and technical sales service. As with adminis-trative costs, for a quick estimate, the distribution and marketing cost can be ap-proximated as 10 percent of *sales.*

7. Shortcut Method for Operating-Cost Estimation. Given the information in the previous section, it is possible to assign average values to all the components that make up operating cost. For some components, an average value has been given already; for the remaining ones, a value is given in the following outline. The use of average values should be discouraged, however, since the results of such an estimate lack individuality. Averages make no allowance for differences between situations and do not challenge the true capability of the estimator. The outline is useful, nevertheless, in instances where a high level of precision is neither necessary nor possible, and sometimes no data at all are available. Any output may be con-sidered better than no output, and the most expedient figures are often necessary.

Average Data for Order-of-Magnitude Operating-Cost Estimate

1. Direct operating costs
   a. Raw materials—estimate from price lists
   b. Byproduct and salvage value—estimate from price lists
   c. Operating labor—from literature or from similar operations
   d. Operating supervision—17.5 percent of operating labor
   e. Utilities—15 percent of operating cost or from similar operations
   f. Maintenance and repairs—9 percent of fixed capital investment
   g. Operating supplies—2 percent of total investment (fixed plus working)
   h. Others
      (1) Laboratory—10 percent of operating labor
      (2) Royalties—2.5 percent of sales or service charges
      (3) Contingencies—10 percent of direct operating cost
2. Indirect operating costs
   a. Depreciation—7 percent of fixed capital investment
   b. Property taxes—2.5 percent of fixed capital investment
   c. Insurance—1.5 percent of fixed capital investment
   d. Interest—8 percent of total investment
   e. Rent—10 percent of the rented property
3. Overhead cost—50 percent of operating labor, supervision, and maintenance and materials
4. Administrative cost—45 percent of operating labor
5. Distribution and marketing costs—10 percent of sales

BREAKEVEN ANALYSIS

The relationship of sales revenue, costs, and volume to profit and loss is funda-mental to every business, and a basic understanding of these relationships is nec-

essary before any project is carried out. Payback time, return on investment, or discounted-cash-flow analysis is not always a sufficient tool to demonstrate what happens to profit as changes occur in sales revenue, costs, and volume. Break-even analysis, particularly breakeven charts, is useful in this regard, by exhibiting the relationship among the preceding variables and the degree of effect each has on the final profit.

This section deals with the cost-volume-profit relationship, development of breakeven equations and charts, estimation of cost-volume relationships, and their application for effective decision making.

1. Terminology. In breakeven analysis, the elements considered are total revenue from sales and total costs incurred; the latter is broken down into fixed cost, semi-variable cost, and variable cost.

Total revenue from sales is an estimate of the dollars to be realized from the sales of products or services. It is the first figure to be established and is the most basic. One approach may be multiplying the number of units expected to be sold by the unit selling price to get the revenue figure. Another approach may be adjusting last year's dollar total upward or downward as the economy indicates. Total revenue from sales, however, does not include fixed income or nonoperating income.

Fixed costs are indirect costs. They tend to remain constant in total dollar amount regardless of volume, or output. At zero volume and at 100 percent volume, the total dollar amounts are the same. Fixed costs may include rent, interest on investment, property taxes, property insurance, executive salaries, allowance for depreciation, and sums spent for advertising.

Variable costs are those which vary directly with the level of output. Direct labor, materials, and certain supplies used are considered variable. If volume is halved, variable costs will be halved; if volume is doubled, variable costs will double. For example, if the production of one desk requires 10 worker hours at $5 per hour, then the production of two desks requires 20 worker hours ($100), or a variable cost per desk of $50.

Total cost is the sum of all fixed and variable costs incurred over a fiscal year.

Gross profit is total revenue less total costs, given that the former is greater than the latter. Gross profit is computed before income taxes.

Loss is total costs less total revenue, given that the former is greater than the latter. On balance sheets and other tables, loss is usually enclosed in brackets or preceded by a minus sign. Brackets will be used in this book.

The *breakeven point* (*BE*) is the volume of output at which neither profit nor loss occurs, or where total revenue from sales is equal to total cost (fixed plus variable). Often, the breakeven point is expressed as a percent of production or service capacity instead of as sales volume.

2. Mathematical Analysis. The point in the operation of a project at which revenues and incurred costs are equal to each other is the *breakeven point.* At this particular output or level of operation, a project will realize neither a profit nor a loss. The breakeven point can be computed mathematically or can be ascertained graphically by presenting the relationship of revenue, costs, and volume of a productive capacity. Graphical analysis is presented later.

If b is the variable cost per unit of output and C_v is the total variable cost for the year, then for an annual output of x units, the variable cost is

$$C_v = bx \tag{20.7}$$

Let C_f equal the fixed cost per year, which remains relatively constant, and let C_t be the total annual cost; then for an annual output of x units, the total cost is

$$C_t = C_f + bx = C_f + C_v \tag{20.8}$$

C_t is a linear relationship, since b is assumed to be independent of x, and C_f is constant.

If p is the selling price per unit of output and G is the gross profit, then using Eq. (20.8),

$$G = px - C_t = px - (C_f + bx) \tag{20.9}$$

Since no profit or loss occurs at the breakeven point, Eq. (20.9) at $G = 0$ becomes

$$G = 0 = px - (C_f + bx)$$

or

$$px = C_f + bx$$

Finally,

$$\text{Breakeven output or volume (BEV)} = \frac{C_f}{p - b} \tag{20.10}$$

In terms of breakeven sales dollars (BES), the relationship is

$$\text{BES} = \frac{C_f}{1 - b/p} \tag{20.11}$$

3. Graphical Analysis.   The mathematical analysis for breakeven points is relatively simple. Nevertheless, the use of a breakeven chart provides a clearer idea of the firm's position vis-à-vis its breakeven point by enabling a person to see several important cost relationships that would otherwise be difficult to visualize. As with the mathematical analysis, most breakeven charts work with the idea that fixed costs do not change when sales volume increases or decreases but that variable or direct costs rise or fall proportionately with sales.

Figure 20.1 shows the essential features of a breakeven chart.

In Fig. 20.1, the vertical axis (or y axis) represents both sales revenue and costs, and the horizontal axis (or x axis) shows both production volume in units and sales volume in dollars. The vertical units of measurement on a breakeven chart are always dollars, while the horizontal units of measurement can be units of production, dollars, percent of capacity, hours, etc.

At zero sales, revenue is also zero. Therefore, the total revenue line is a straight line passing through the origin, representing the sales revenue increasing as volume sold increases ($y = px$). If volume (the x axis) is expressed in terms of sales volume (dollars), the revenue line on the breakeven chart is simply the straight line $y = x$.
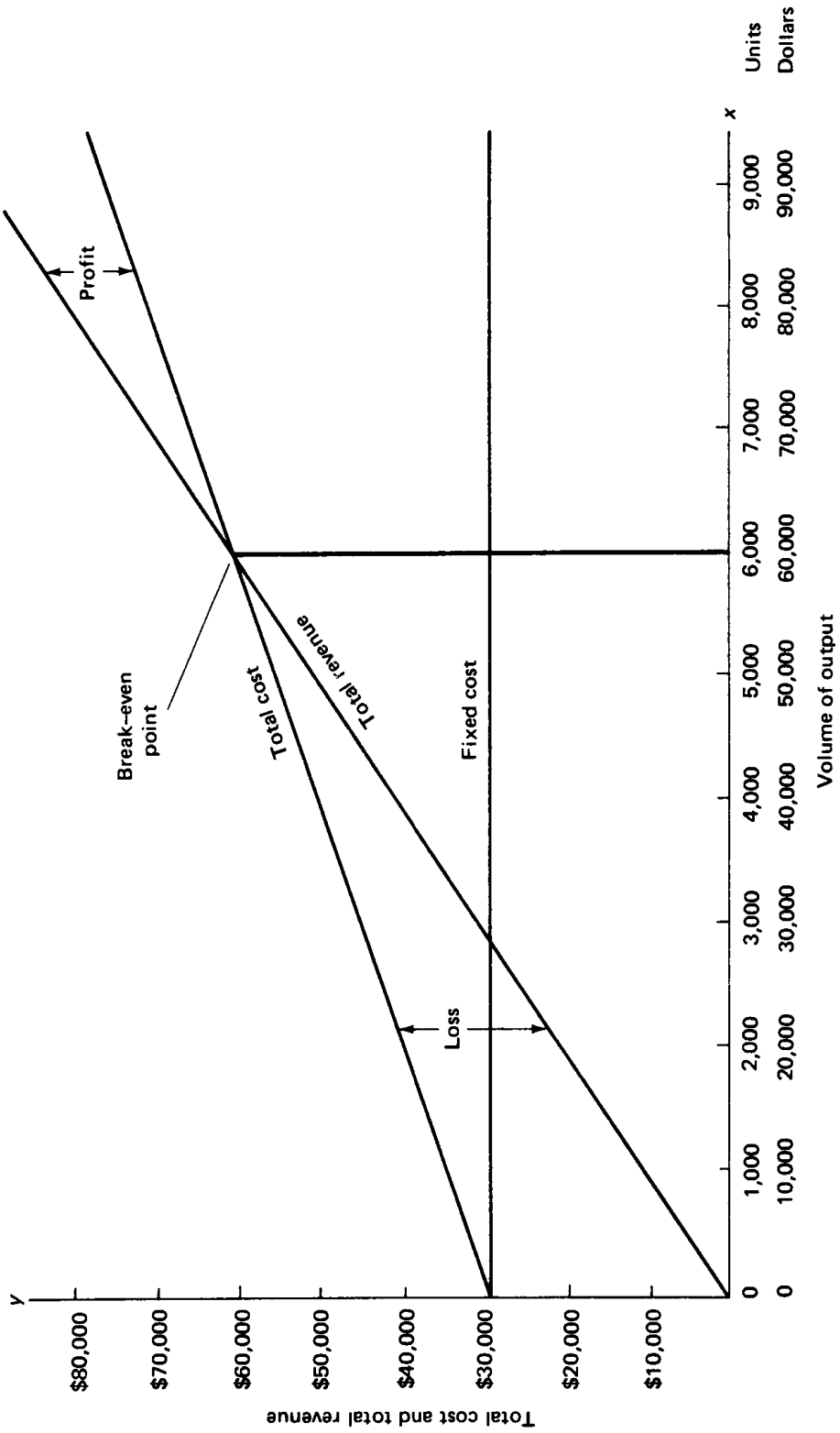
FIGURE 20.1 Essential features of a breakeven chart.

20.18

Next, the total amount of fixed costs is plotted on the graph. Fixed costs in this case total $30,000 throughout the range of sales shown on the breakeven chart. Variable costs are then added to fixed costs to arrive at total costs, which can be expressed in terms of y. Thus $y = C_t + bx$, with the slope of the variable cost line equal to b, or in this case $5 per unit. These variable costs are plotted above the fixed cost line, thereby summing the total cost of operations for any given level of sales or output.

The point at which the total cost line intersects the total income line, 6000 units or $60,000 in this example, is the breakeven point. To the left of this point, the vertical distance between the total income and the total cost lines indicates a net loss; to the right, it represents the net profit. Hence, at sales volume of 4000 units, a net loss of $50,000 − $40,000 = $10,000 occurs. At a sales volume of 8000 units, $80,000 − $70,000 = $10,000 in gross profits is realized.

Effect of Changes in the Various Components.   The revenue-cost-volume relationships suggest that there are three ways in which the profit of a project can be increased:

1. Increase the selling price per unit (p).
2. Decrease the variable cost per unit (b).
3. Decrease the fixed cost (C_f).

The separate effects of each of these possibilities are shown in Fig. 20.2. Each starts from the current situation (Fig. 20.1: $b = $5/unit, $p = $10/unit, $C_f = $30,000, BE = 6000 units = $60,000).
The effect of a 10 percent change in each factor is calculated:

1. A 10 percent increase in selling price would decrease the breakeven point to 5000 from 6000 units. At a sales volume of 8000 units, the gross profit becomes $18,000, an increase in gross profit of $8000 ($18,000 − $10,000), while at 4000 units, the loss in profit becomes $6000 instead of the original $10,000.
2. A 10 percent decrease in variable cost would shift the breakeven point from 6000 to 5455 units. At a sales volume of 8000 units, the gross profit becomes $14,000, an increase of $4000 ($14,000 − $10,000), while at 4000 units, the loss in profit becomes $8000 instead of the original $10,000.
3. A 10 percent decrease in fixed cost would shift the breakeven point from 6000 to 5400 units. At a sales volume of 8000 units, the gross profit becomes $13,000, an increase of $3000 ($13,000 − $10,000), while at 4000 units, the loss in profit becomes $7000 instead of the original $10,000. Figure 20.3 shows the three changes simultaneously.

If we look more closely at some of the relationships, we can calculate, for example, that a 10 percent increase in fixed cost could be offset by a 3.75 percent increase in selling price, a 7.5 percent decrease in variable cost, or a 7.5 percent increase in volume sold. This clearly shows that if one wanted to increase the profit of a given project, the sequence of actions should be (1) increase selling price, (2) sell more units, (3) decrease variable cost, and (4) decrease fixed costs.

Another important calculation made from the breakeven chart is the *margin of safety.* This is the amount or ratio by which the current or operating volume exceeds the breakeven volume. Assuming the current volume is 8000 units, and the breakeven point in our illustrative situation is 6000 units, the margin of safety then is
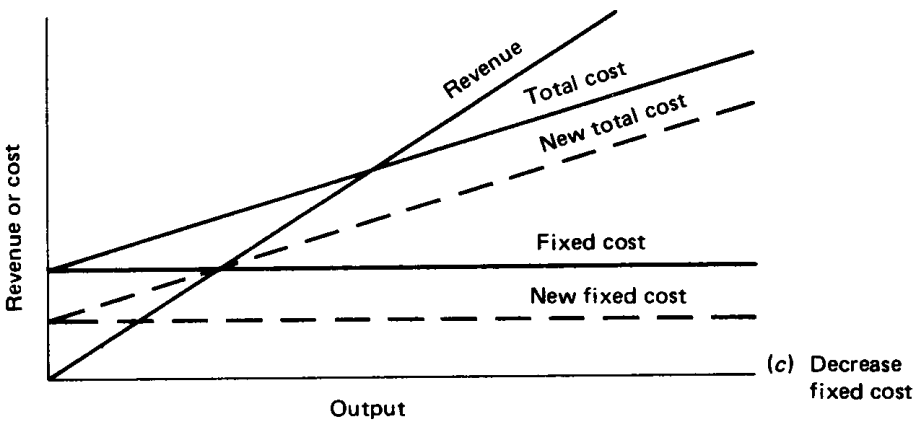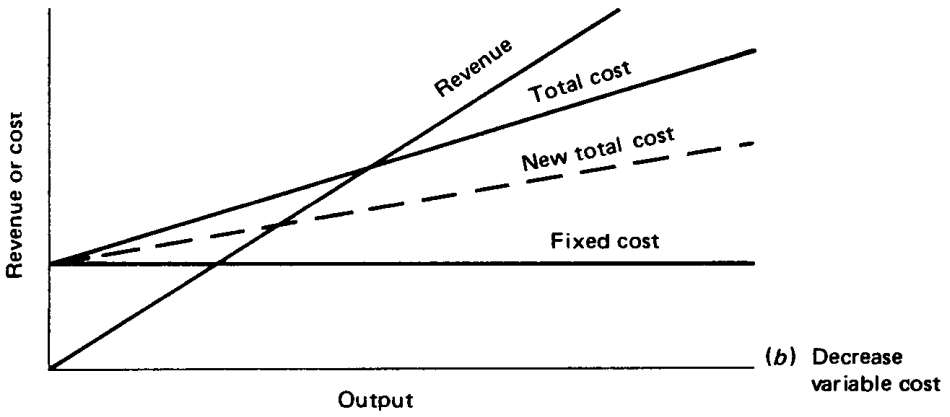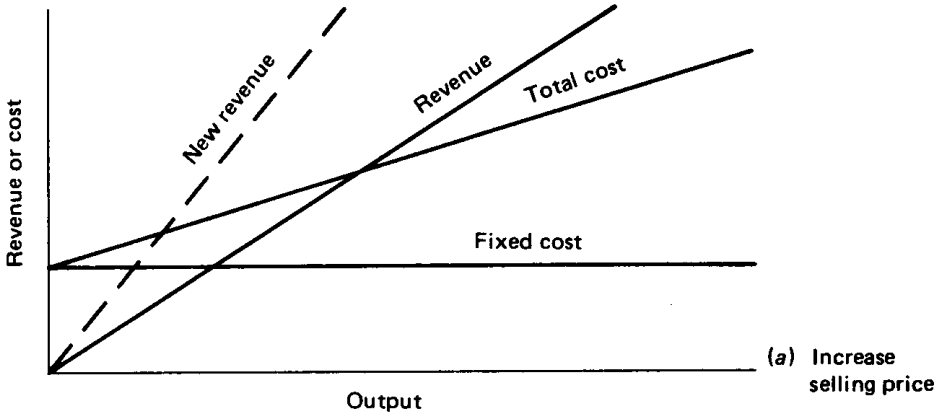
FIGURE 20.2 Effect of increased selling price (*a*), decreased variable cost (*b*), decreased fixed cost (*c*).
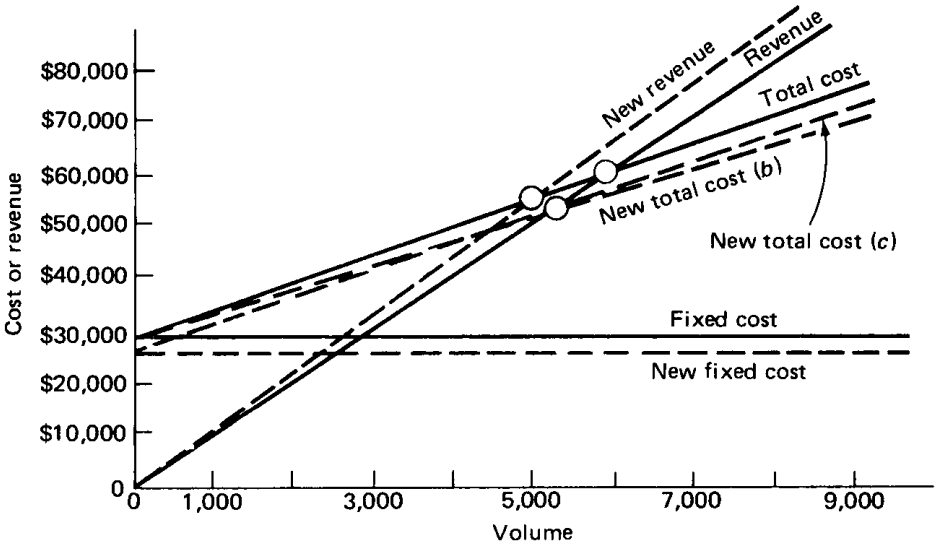
FIGURE 20.3   Effects of three changes.

33.33 percent. Sales volume can decrease by 25 percent before a loss is incurred, given that other factors remain constant.

4. Estimation of the Cost-Volume Relationship.   In many practical situations, costs are expected to vary with volume or output in the straight-line relationship shown in Fig. 20.4. The formula for this line of expected costs can be estimated by any of the following methods.
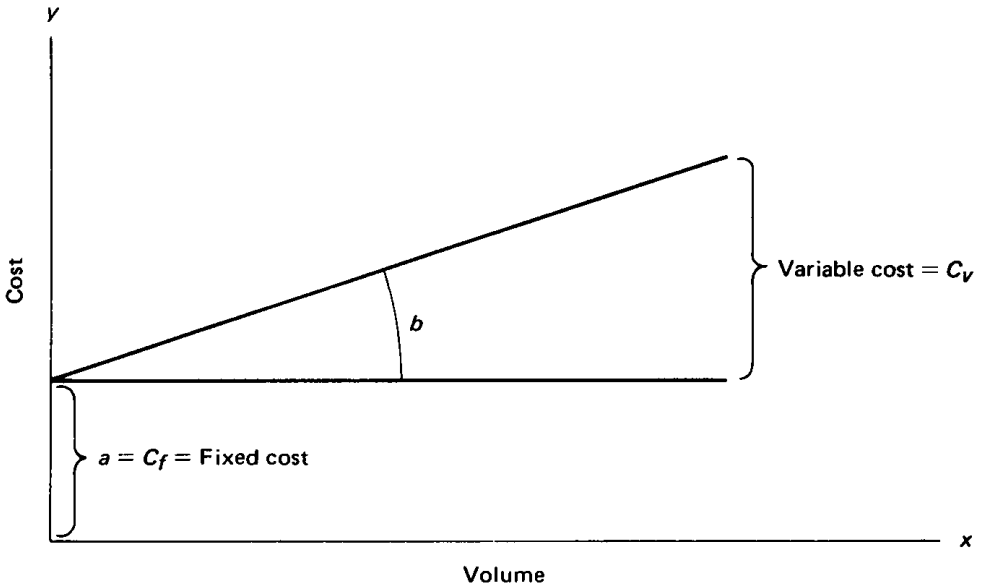
FIGURE 20.4   Straight-line relationship for cost vs volume.

High-Low Method. Designating cost as *y*, volume as *x*, and the variable component as *b*, and letting fixed cost component $C_f = a$, the cost at any volume can be found from the formula $y = a + bx$, which is simply the general formula for a straight line.

If the values of *a* and *b* for a given line are unknown, they can be calculated, provided that total costs are known for any two points, or volume levels, on the line. Let

$$C_{TL} = \text{total cost at the lower volume}$$

$$C_{TH} = \text{total cost at the higher volume}$$

$$x_L = \text{lower volume}$$

$$x_H = \text{higher volume}$$

The variable cost component *b* is then

$$b = \frac{C_{TH} - C_{TL}}{x_H - x_L} \tag{20.12}$$

and the fixed cost component *a* is

$$a = y_H - bx_H = C_{TH} - bx_H \tag{20.13}$$

Scatter-Diagram Method. Another way to estimate *a* and *b* is to plot actual costs recorded in the past periods against the volume levels in those periods, such as illustrated in Fig. 20.5, and then draw a straight line through the points so that the vertical deviation of the points above and below the line are exactly equal. In the event the points in the scatter diagram are numerous or widely scattered, the average values of the groups of data should be plotted to serve as guidepoints in drawing the line.

First, the data should be divided into several groups according to values of *x*, with each group having the same number of data points. An average is then taken of each group with respect to both *x* and *y*, these averages are then replotted, and
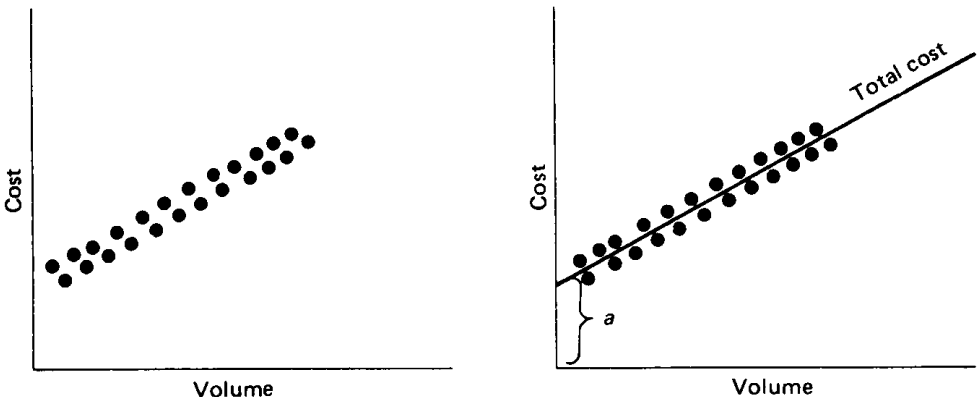
FIGURE 20.5 Scatter-diagram for cost vs volume.

a straight line is fitted as described above. The criterion of goodness of fit in the above method is a visual one and rather subjective.

EVALUATING INVESTMENTS

This section concerns itself with some of the different methods being employed to evaluate venture decisions. Years ago, intuition was an adequate tool for making such decisions. Today, however, owing to pressures of competition, quantitative techniques have replaced intuition, reducing the risk of failure. A decision to undertake a venture involves careful analysis of capital requirements and other resources for profit maximization within the framework of social responsibilities. Right timing of the venture is critical as well, since a product or process remains profitable for a limited time only, the average life cycle being 5 years.

Because of these life cycles, investments and costs must be carefully controlled. Among factors to be evaluated in an investment analysis are the uncertainties of the state of the economy, the possibility of operating failures, and technological changes. With quantitative evaluations, making an investment decision is a matter of weighing anticipated profits against the minimum profitability standard set by a firm. The company must make profitable investments and refrain from making unprofitable ones.

Investment evaluation must be objective, realistic, appropriate to the situation, and easily understood by management. This responsibility lies in the hands of the evaluating engineer or technologist.

There are many methods in use for evaluating investments. Some incorporate the time value of money; that is, they consider the fact that profits to be realized from a given project decrease in value in the future as a result of inflation. Or, stated differently, the dollar will have a lower purchasing power next year and even lower the year after, and so on. Other methods neglect the time value of money, which makes them inefficient if a project has a long useful life. The basic and most important methods involving the time value of money are

1. Payback time (PT)
2. Return on investment (ROI)
3. Benefit/cost analysis (B/C)

The most widely used is payback time; however, return on investment is the most logical and theoretically acceptable means of determining investment feasibility. Since no single method is best for all cases, the engineer or technologist should understand the basic concepts involved in each method and be able to choose the best one suited to the situation.

Before proceeding to describe each of the methods, there are investments to which none of them can be applied for measure of worth yet such investments are carried out regardless. Such investments are classified as *investments due to necessity.*

1. Investment Due to Necessity.   Investments of this type are not very popular in view of the fact that they have no direct effect on cost reduction (saving) or on sales or profit increase but are necessary for uninterrupted operations, to satisfy

social or legal requirements, or to satisfy intangible but very important goals. The following are typical investments due to necessity:

1. Replacement of worn-down equipment
2. Investments made to conform to governmental regulations, such as antipollution measures
3. Replacement of facilities after disasters
4. Investment in research and development in order to remain competitive

The preceding investments, although not necessarily directly exhibiting tangible savings or profits, nevertheless do influence future profits, since shutting down a given operation may mean that no profits or savings will be realized from the operation. The following example will illustrate.

2. Payback Time.   For studies involving the design of components (equipment) for processing plants, it is often convenient to evaluate capital expenditures in terms of *payback time* (other equivalent terms are *payback period, payout time, payoff period,* and *payoff time*). It is the number of years over which capital expenditure (not including working capital) will be recovered, or paid back, from profits made possible by the investment; that is, the project or piece of equipment will "pay for itself" in this number of years. It is a quick and convenient but crude method of identifying projects that are apt to be either highly profitable or unprofitable during their early years. If the payback time is equal to or only slightly less than the estimated life of the project or piece of equipment, then the proposal is obviously a poor one. If, on the other hand, the payback time is considerably less than the estimated life, then the proposal begins to look attractive.

Another attractive feature of payback time is its usefulness in selecting acceptable proposals out of several investment alternatives having similar characteristics.

The major disadvantage of payback time is its failure to consider profits after the investment has been recovered.

There are several ways to determine payback time. The three most widely used are *payback time based on average yearly gross profit, payback time based on average yearly net profit,* and *payback time based on average yearly cash flow.*

Payback Time Based on Average Yearly Gross Profit

$$\text{PT} = \frac{\text{capital invested}}{\text{average yearly gross profit}} = \frac{I}{G} \qquad (20.14)$$

where PT is expressed in years. It should be noted that the invested capital does not include working capital.

Payback Time Based on Average Yearly Net Profit

$$\text{PT} = \frac{\text{captial invested}}{\text{average yearly net profit}} = \frac{I}{P} \qquad (20.15)$$

where $P = G - i(I + I_w) - t(G - dI)$
$\quad G$ = gross profit
$\quad I$ = capital invested in equipment or facilities, or fixed capital invested

I_w = working capital
i = interest rate on all borrowed capital
d = depreciation rate on fixed capital invested
t = income tax rate (federal and state)

As can be seen, Eq. (20.15) includes the interest charged on borrowed capital.

Payback Time Based on Average Yearly Cash Flow

$$\text{PT} = \frac{\text{capital invested}}{\text{average yearly cash flow}} = \frac{I}{P + dI} \qquad (20.16)$$

Cash flow, which is the total amount of money generated by an investment, is found by adding the annual depreciation charge to net profit.

As in Eq. (20.15), the preceding expression includes the interest charged on borrowed capital, fixed and working.

The following set of examples will illustrate the use and application of the preceding concepts.

Maximum Payback Time.
This is the maximum acceptable payback time (PT_M), which is set by the firm, based on minimum acceptable return rate. It is expressed as

$$\text{PT}_M = \frac{1}{d + [\text{ROI}_m/(1 - t)]} \qquad (20.17)$$

where d = depreciation rate
ROI_m = minimum acceptable return rate
t = income tax rate (federal and state)

Equation (20.17) establishes a parameter for payback time; that is, projects must "pay for themselves" within a period that is equal to or less than the time set by the firm, using a minimum acceptable return rate.

3. Return on Investment. The return-on-investment approach relates the project's anticipated net profit to the total amount of capital invested. The total capital invested consists of that actually expended for facilities or equipment as well as working capital.

To obtain reliable estimates of investment returns, it is necessary to make accurate calculations of net profits at the required capital expenditure. To compute net profit, estimates must be made of direct production costs, fixed expenses such as interest, depreciation, and overhead, and general expenses.

The main disadvantage of the return-on-investment approach is its complexity as compared with payback time, but this disadvantage allows increased precision and thoroughness.

The two most widely used approaches in determining return on investment are *return on original investment* and *return on average investment*.

Return on Original Investment.
This is the percentage relationship of the average annual net profit to the original investment (which includes nondepreciable items such as working capital). It can be expressed as

$$\text{ROI} = \frac{\text{net profit} \times 100}{\text{capital invested} + \text{working capital}} = \frac{P}{I + I_w} \times 100 \qquad (20.18)$$

4. Benefit / Cost Analysis. There are capital expenditure projects to which payback time or ROI cannot be applied as measures of project worth. Governmental expenditures for public works in the areas of flood control, environmental protection, conservation, highways, public health, and urban renewal are a few examples of such projects. Since no cash flows are realized (that is, no cash receipts are available throughout the life of such projects), there is no basis for an economic evaluation. Nevertheless, since nearly all federal projects are financed from a common pool of tax funds, a project's worth must be assessed by some means and weighed against an estimate of the project's cost to make sure that whatever limited funds are available are allocated and spent wisely. This is where benefit/cost (B/C) analysis comes in.

The term *benefit* in benefit/cost analysis refers to the savings from projects such as public works. Examples of such benefits are reduced destruction of property from flooding, reduced demand for medical and hospital services due to a cleaner environment, reduced accident costs by elimination of hazardous intersections or improvement of guard rails, reduced vehicle wear and tear due to improved road surfaces, reduced travel time due to shorter routes, higher speeds, or elimination of unnecessary stops, and reduced crime and looting due to urban renewal.

It should be mentioned, however, that the positive benefits mentioned above often yield negative benefits simultaneously. These negatives must be incorporated into the analysis as well; that is, the negative benefits must be subtracted from the positive benefits to yield a net positive benefit. An example of a negative benefit would be longer distances to travel because a highway improvement proposed for safety reasons will restrict access to a highway. Accordingly, additional wear and tear on a car or bus will be realized as well as an increase in energy consumption. The costs considered are similar to those of private enterprises, that is, costs for engineering, construction, and maintenance, plus less obvious costs for the survey; design inspection of the construction; bid evaluation; transportation; contract negotiation, award, and management; supervision; personnel; accounting; and other applicable services.

The preceding definitions of benefit/cost permit measurable benefits to be weighed against measured costs. For a project to be economically acceptable, it must yield user benefits which meet or exceed the cost of providing those benefits; that is, the ratio of benefit to cost must equal or exceed 1. Expressed mathematically,

$$\frac{\text{Positive user benefits minus negative user benefits}}{\text{Initial investment plus annual operating costs}} = \frac{\text{B}}{\text{C}} \geqq 1 \qquad (20.19)$$

5. Incremental Analysis. *Incremental analysis* is the evaluation of the profitability of a project and its alternatives on the basis of the effects specifically caused by each project. Thus incremental cash flows, incremental operating costs, or incremental investments are those which occur (or do not occur) as a direct result of a particular project or course of action.

For example, consider a project to replace a 10,000-L chemical reactor with an identical 25,000-L unit. The only relevant factors which can be considered are productivity improvement per unit of time and costs that will be reduced or in-

curred; such as reduction in energy consumption per unit of output, or increase or decrease in maintenance costs. Reduction in labor cost would only be counted if an operator could be laid off because of the slack time generated by the productivity improvement. The analysis would not include such plant costs as overhead, supervision, quality control, floorspace occupancy, etc., since these would not increase or decrease as a result of this project.

Deciding Between Alternatives with Incremental Analysis. The concept of incremental analysis finds its greatest use and application in decision making between alternative investments. Investments and profits of mutually exclusive projects are compared incrementally (occurrence of one excludes the occurrence of the others); then, depending on the resultant ROIs, a decision is reached as to which of the alternatives is the best.

    When comparing one alternative with another, the first task is to determine the incremental profit representing the difference between the two profits, that is, between projects *A* and *B*, *A* and *C*, *B* and *C*, etc. The second task is to determine the difference between the two investments, for *A* and *B*, *A* and *C*, *B* and *C*, etc. The incremental investment is considered desirable if it yields an ROI greater than or equal to the minimum rate of return (ROI_m) as set by the firm. In simple terms, the incremental ROI can be expressed mathematically as follows:

$$\text{ROI}_{\text{incremental}} = \text{ROI}_{B-A} = \frac{P_B - P_A}{(I + I_w)_B - (I + I_w)_A} \qquad (20.20)$$

    To apply the rate of return on an incremental basis for a group of independent projects, it is first necessary to rank the projects in ascending order of their total investment, fixed plus working capital ($I + I_w$). Then Eq. (20.20) is employed to yield the incremental rates of return.

EVALUATING INVESTMENTS USING THE TIME VALUE OF MONEY

Since many economic feasibility studies extend over a long period of time or deal with projects having long useful lives, it is necessary to recognize that future cash flows or profits generated by these projects will decrease in value over time because of inflation. As stated earlier, today's dollar will have less purchasing power next year and even less the year after, and so on.

    This section concerns itself with methods of comparing cash flows or profits at various points in time. These take into account the time value of money, which is interest. The basic and most important methods are

1. Maximum payback time (PT_M)
2. Discounted-cash-flow analysis (DCF)
3. Benefit/cost analysis (B/C)

    The most widely used of these methods is DCF; however, since no single method is best for all cases, the engineer or technologist should understand the basic concepts involved in each method and be able to choose the one best suited to the needs of the particular situation.

Before proceeding to the methods listed above, it is necessary to define *interest* and to present the mathematical relationships which permit conversion of money at a given point in time to an equivalent amount at some other point in time.

1. Interest and the Time Value of Money. *Interest* is defined as the compensation paid for the use of borrowed money. Since most firms have to borrow capital in order to expand, interest expense plays a significant role in an economic analysis. The rate at which the interest is to be repaid is usually determined at the time the capital is borrowed, along with a scheduled time of repayment. To the lender, interest represents compensation for not being able to use the money elsewhere right now. That is, interest in most respects is the compensation for the decrease in value of the money between now and when the loan is repaid, this decrease being due to inflation. The borrower, on the other hand, must invest the borrowed capital in an activity that will yield a return higher than the penalty (interest) for borrowing the capital.

2. Interest Formulas and Interest Tables. In economic terms, *principal* is defined as capital on which interest is paid, while *interest rate* is defined as the cost per unit of time of borrowing a unit of principal. The time unit most commonly taken is 1 year.

Interest Formula Symbols

$$PW = \text{principal, or present worth}$$

$$i = \text{interest rate per period}$$

$$n = \text{number of interest periods}$$

$$m = \text{interest periods per year}$$

$$S = \text{principal plus interest due, or future worth}$$

$$R = \text{uniform periodic payment}$$

Compound Interest. *Compound interest* can be defined as interest earned on interest, that is, interest is earned, not only on the principal, but also on all previously accumulated interest. If a payment is not made, the interest due is added to the principal and interest is charged on this converted principal during the following year.

Annuities. An *annuity* is a series of equal payments occurring at equal time intervals. Payments of this type are used to pay off debt or depreciation. In the case of depreciation, the decrease in the value of equipment with time is accounted for by an annuity plan. For the uniform periodic payments made during n discrete periods at i percent interest to accumulate an amount S,

$$S = R \left[\frac{(1 + i)^n - 1}{i} \right] \tag{20.21}$$

Equation (20.21) is termed the *annuity compound-amount factor,* and $(1 + i)^n / i$ is the *compound-amount factor.*

The reciprocal of Eq. (20.21) is known as the *annuity sinking-fund factor* and can be used to determine yearly depreciation cost. It is expressed as follows:

$$R = S \left[\frac{i}{(1 + i)^n - 1} \right] \tag{20.22}$$

where $i/[(1 + i)^n - 1]$ is the sinking-fund factor.

An analysis of Eq. (20.22) shows that equal amounts of R, when invested at i percent interest, will accumulate to some specified future amount S over a period of n years. In terms of depreciation expense, it shows that equal yearly depreciation costs invested at an i percent interest rate for n years will accumulate to an amount equal to the original cost of the equipment.

The *present worth* (PW) of an annuity is the principal which would have to be invested at the present time at compound interest i to yield a total amount at the end of the annuity term equal to the amount of the annuity. In other words, it is the present amount (PW) that can be paid off through equal annual payments of R over n years at i percent interest. Combining Eq. (20.24) with Eq. (20.21) gives

$$\text{PW}(1 + i)^n = S = R \left[\frac{(1 + i)^n - 1}{i} \right] \tag{20.23}$$

or

$$\text{PW} = R \left[\frac{(1 + i)^n - 1}{i(1 + i)^n} \right]$$

Equation (20.23) is known as the *annuity present-worth factor,* where $[(1 + i)^n - 1]/i(1 + i)^n$ is the *present-worth factor.*

The reciprocal of Eq. (20.23) is known as the *annuity capital-recovery factor.* It is the annual payment R required to pay off some present amount PW over n years at i percent interest. It can be expressed as follows:

$$R = \text{PW} \left[\frac{i(1 + i)^n}{(1 + i)^n - 1} \right] \tag{20.24}$$

3. Maximum Payback Time. This is the maximum acceptable payback time (PT_M) which is set by the firm based on minimum acceptable return on investment and on the time value of money. Accordingly,

$$\text{PT}_M = \frac{1 - t}{\left[\dfrac{i(1 + i)^n}{(1 + i)^n - 1} \right] + (\text{ROI}_m - i) - d} \tag{20.25}$$

where n = length of useful life of the project
$\quad \text{ROI}_m$ = minimum acceptable return on investment
$\quad\quad d$ = depreciation rate
$\quad\quad t$ = income tax rate (federal plus state)

4. Discounted-Cash-Flow Analysis. The most popular method for evaluating investment alternatives, taking into account the time value of money, is the *discounted-cash-flow* (DCF) method. It includes all cash flows over the entire life of

the project and adjusts them to one point fixed in time, usually the time of the original investment. The method requires a trial-and-error calculation to determine the compound interest rate at which the sum of all the time-adjusted cash outflows (investment) equals the sum of all the time-adjusted cash inflows (net profit plus depreciation). The main attractiveness of this technique is that it considers both the amount and the timing of all cash inflows and outflows.

The discounted-cash-flow rate of return is the *after-tax interest rate i* at which capital could be borrowed for the investment and just break even at the end of the useful life n of the project. To determine the DCF rate of return, the present worth PW of the project—or *net cash flow* (NCF) (net profit plus depreciation) compounded on the basis of end-of-year income—is expressed as

$$PW = NCF_1 \left[\frac{1}{(1 + i)^1} \right] + NCF_2 \left[\frac{1}{(1 + i)^2} \right] + \cdots + NCF_n \left[\frac{1}{(1 + i)_n} \right] \qquad (20.26)$$

where $1/(1 + i)^1$ = present worth at the end of year 1
$1/(1 + i)^2$ = present worth at the end of year 2
$1/(1 + i)^n$ = present worth at the end of the project

Then the present worth of the fixed investment I compounded at interest rate i plus working capital (I_w) is expressed as

$$PW = (I + I_w)_0 \left[\frac{1}{(1 + i)^0} \right] + (I + I_w)_1 \left[\frac{1}{(1 + I)^2} \right] + \cdots + (I + I_w)_n \left[\frac{1}{(1 + i)^n} \right]$$

$$(20.27)$$

where $1/(1 + i)^0$ is the present worth at the start of the project.

If only one time investment is made, at the beginning of the project, then Eq. (20.27) becomes

$$PW = (I + I_w)_0 \left[\frac{1}{(1 + i)^0} \right] = \frac{I + I_w}{1} = I + I_w$$

The present worth of the investment must be equal to the present worth of the cash flows; that is, investment (cash outflows, or O) must equal the sum of all the time-adjusted cash inflows. Therefore, we set Eq. (20.26) equal to Eq. (20.27):

$$I + I_w = \frac{NCF_1}{(1 + i)^1} + \frac{NCF_2}{(1 + i)^2} + \cdots + \frac{NCF_n}{(1 + i)^n} \qquad (20.28)$$

or expressed differently,

$$O = -(I + I_w) + \frac{NCF_1}{(1 + i)^1} + \frac{NCF_2}{(1 + i)^2} + \cdots + \frac{NCF_n}{(1 + i)^n} \qquad (20.29)$$

A trial-and-error calculation is required to determine the DCF rate of return i.

Since the discounted-cash-flow rates of return are only approximated by the process of interpolation, a slight error is introduced. To keep the error to a minimum, interpolation should only be attempted between adjacent tabled values, for example, between 10 and 11 percent. Interpolation between, say, 2 and 10 percent would increase the error greatly.

PATENTS, TRADEMARKS AND COPYRIGHTS*

PATENTS

1. United States. (All statutory references are to the Patent Act of 1952, 35 U.S.C., et seq.)

Types of Patents Granted in the United States. United States patents are classified as utility, design, and plant patents with the first category comprising the vast majority of patents granted in the United States. A design patent can be obtained for any ''new, original and ornamental design for an article of manufacture'' (35 U.S.C. § 171). The emphasis in design patents is on the ornamental, rather than functional, aspects of the article of manufacture. A plant patent may be obtained by ''whoever invents or discovers and asexually reproduces any distinct and new variety of plant'' (35 U.S.C. § 161), and gives that person the right to exclude others from asexually reproducing the plant or selling or using the plant so produced (35 U.S.C. § 163).

The Patent Grant. In return for a full disclosure of an invention to the public in the patent document, the United States will grant an exclusive monopoly to an inventor or her or his assignee for the term of the patent. As a result of U.S. ratification of the GATT treaty, the patent term for utility patents has changed from 17 years from the grant of the patent to 20 years from the initial filing date of the application. This new patent term begins on the date the patent is issued and ends 20 years from the date of initial filing. The new 20-year term from date of filing applies to all patents granted on applications filed on or after June 8, 1995. All patents based on applications filed prior to June 8, 1995, will have a term that is the greater of the 20-year term provided by the new law or 17 years from the grant. The 20-year term can also be extended for a maximum of 5 years for delays in grant of a patent due to interferences, secrecy orders, or successful appeals to the Patent and Trademark Office (PTO) or the courts. However, any extensions are subject to certain limitations. The term for a design patent is 14 years. After the term of the monopoly expires, the public is then free to practice the invention. Alternatively, an inventory may choose not to disclose his or her invention to the public but instead maintain it as a trade secret; but by doing so, an inventor cannot stop others from using his or her invention if someone else independently develops the same invention.

Determining What Is Patentable. The basic statutory requirements for patentability are utility of the invention (35 U.S.C. §101), novelty (35 U.S.C. §102), and nonobviousness (35 U.S.C. §103). To meet the utility requirement, an inventor need only disclose in his or her patent an invention which is new, useful, and functional in the field of applied technology as opposed to theoretical or abstract ideas. Also, developments in the nontechnical arts, social sciences, methods of doing business,

---

*This chapter is drawn from *Marks' STANDARD HANDBOOK FOR MECHANICAL ENGINEERS,* 10th ed., by Eugene A. Avallone (ed.) and Theordor Baumeister III (ed.). Copyright © 1996. Used by permission of McGraw-Hill, Inc. All rights reserved. Updated in 2003 by editors.

computer programs per se, and the like are not generally patentable. However, there have been recent developments in the law that have expanded the possibilities for patenting certain types of business systems and some aspects of computer programs. A patent professional should be consulted regarding these developments.

The novelty and nonobviousness requirements of the patent laws are assessed in light of the prior art which generally comprises all written materials antedating the inventor's application for a patent and/or his or her actual date of invention (see section on interferences below for more detail on what constitutes date of invention). Also included in prior art are prior public uses and sales by others of the same or similar processes or products as that of the inventor. A prior public disclosure or offer to sell by the inventor made more than 1 year before the inventor's patent application also constitutes prior art which can prevent the grant of a patent [35 U.S.C. §102(b)]. Thus, it is crucial that patent applications be promptly filed (within a year) after any effort at commercial exploitation of an invention by or on behalf of the inventor. In assessing an invention against the novelty requirement of 35 U.S.C. §102, the courts have held that if the invention is different—without regard to how different—from each item of prior art, it meets that requirement In other words, if no single piece of prior art discloses the invention, it is considered novel under 35 U.S.C. §102.

Since few inventions—at least not those for which patent protection is sought—are identical to any single item of prior art, the patentability of most patents must be assessed under 35 U.S.C. §103, i.e., whether "the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains." Such assessment requires a factual determination at the time of the invention of the prior art, differences between the prior art and the invention, and the level of skill in the art to which the invention pertains, and a further assessment whether those differences would have been obvious in light of the level of skill of the practitioners in the art. To assist in this determination, the Supreme Court has indicated that courts should look to important background facts such as long-felt need for the invention, failure of others to achieve the invention, and ultimately, any commercial success of the invention [*Graham v. John Deere Company of Kansas City,* 383 U.S. 1, 86 S.Ct. 684 (1966)].

Documentation of an Invention.   In ongoing research, recordation of developments is essential to prove the date of invention. The date of invention may become important in a PTO proceeding to determine priority of invention as between two inventors. Such a proceeding is called an interference. In an interference, the priority of invention must usually be established by contemporaneous, witnessed documentation of the work underlying the invention. A notebook used to record, explain, and date research work is invaluable in an interference to prove an early invention date. The notebook should be signed by the researcher and witnessed by one or more persons who have direct knowledge of, and understand, what was being done during the making of the invention.

Prior to the recent ratification of the GATT and NAFTA treaties, inventive activity occurring in foreign countries could not be relied upon to establish a date of invention in the United States. As a result of GATT and NAFTA, 35 U.S.C. §104 has been amended to provide that evidence of inventive activity in the territory of a World Trade Organization (WTO) member country, Canada, or Mexico be treated the same as inventive activity in the United States. This change in law applies to

all applications filed on or after January 1, 1996. However, a date of invention may not be established in a WTO member country (other than the United States, Canada, or Mexico) that is earlier than January 1, 1996, except as provided by statute (35 U.S.C. §119 and §365). Thus, companies or individuals engaging in research in foreign countries may rely on that activity in an interference.

Preparing the Patent Application.   A recommended first step in preparing an application is to evaluate the prior art and its relation to the invention for which patent protection is to be sought, utilizing the obviousness guidelines of the Supreme Court noted above. The most comprehensive source of prior art is the search room at the PTO in Arlington, Virginia, where over 4 million patents and publications available to the public and are broken down into thousands of classes and subclasses to facilitate a search. Less comprehensive compilations of U.S. patents can also be found in the libraries of many large cities. Copies of U.S. patents can be obtained from these libraries or directly from the PTO. Search results should be evaluated by a competent patent attorney or agent who can advise whether the invention is patentable over the prior art uncovered and who can also focus on the differences between the invention and prior art when preparing a patent application covering the invention.

If the invention, after careful evaluation, is considered to be patentable, the next step is identification of the proper inventors. In the United States, patent applications are filed in the name of the individual(s) who contributed to all or part of the invention rather than the legal owner of the invention or application, e.g., the inventor's employer. Where two or more persons contributed to all or part of the invention (a joint invention), the respective efforts of these persons must have been made in collaboration with each other and the invention must be the product of their aggregate efforts.

The patent application comprises four essential parts: the specification or detailed description of the invention; claims, particularly pointing out the features which the inventor considers to be his or her invention; a drawing illustrating the invention, where applicable; and an oath or declaration of the inventor(s) stating a belief that he or she believes himself or herself to be the original and first inventor of the invention claimed in the application (35 U.S.C. §111–115).

In addition to the standard patent application, applicants may elect to file a provisional application. A provisional application only requires the filing of a specification complying with 35 US.C. §112, paragraph 1, and drawings where necessary for the understanding of the invention. The application must be made in the name of the inventors, include the appropriate filing fee and a suitable cover sheet A provisional application does not require any claims, nor does it require an oath or declaration. A provisional application is not examined by the PTO, cannot claim priority of an earlier application, and cannot issue as a patent In addition, provisional applications will automatically become abandoned 12 months after filing and can become abandoned prior to the expiration of that 12-month-period for failure to respond to a PTO requirement or to pay a required fee. A provisional application may be revived, but its pendency cannot extend beyond 12 months from its filing date. The provisional application is a form of domestic priority system which provides a quick and inexpensive method to establish an early effective filing date which establishes a constructive reduction to practice for any invention described in the application. However, this priority is only effective provided a suitable non-provisional application is filed prior to the expiration of the provisional application.

Proceedings In the U.S. Patent and Trademark Office. After the application is filed with the appropriate fee ($365 to $730 basic fee for utility patents; $150 to $300 for design patent; and $245 to $290 for plant patent), an examiner in the PTO reviews the application to ensure compliance with all statutory requirements as to the form of the application and to independently assess the patentability of the claims in the application in light of prior art. If found to be patentably distinct from the prior art, a patent will be granted upon payment of another fee ($605 to $1,210 for a utility patent; $210 to $420 for a design patent; $305 to $610 for a plant patent). If the examiner concludes the invention claimed in the application is not patentable over prior art, all such claims will be rejected. An administrative appeal to the Patent Office Board of Appeals is available to challenge such rejection and a further appeal to the federal courts is available from an adverse decision of the Board of Appeals (35 U.S.C. §§134, 141, 145).

To maintain a patent, the owner must also make periodic payments to the PTO before the fourth, eighth, and twelfth anniversaries of the patent grant. These maintenance payments increase from $480 to $960 to $1,450 to $2,900 as the patent ages. Since the above-mentioned fees change annually, all fees should be determined in advance to avoid the consequences of fee underpayment.

Rights of a Patent Owner. A patent gives the patent owner (either original applicant or her or his assignee) the right to prevent anyone else from infringing upon his or her patent rights by making, using, offering to sell, selling within the United States, or importing the patented invention into the United States. If someone else does infringe a patent, the patent owner can seek an injunction against further infringement and/or damages for such unauthorized use of his or her invention (35 U.S.C. §§28l–284). The federal courts have exclusive jurisdiction over suits relating to patents.

A patent owner may commercially exploit his or her patent by selling (assigning) or licensing it to another. Multiple licenses can be granted under a single patent giving each licensee nonexclusive rights under the patent. Geographic and use limitations in a license are also possible as a means of maximizing the patent owner's benefit from his or her patent.

Reissue and Reexamination of Patent. If through inadvertence, accident, or mistake, the original patent was defective or claimed too much or too little, the error may be repaired by a reissue proceeding provided the subject matter needed to correct the mistake appeared in the original patent (35 U.S.C. §251). The term of a reissued patent expires the day the original patent would have expired.

Also, if the patentability of a patent is put in question, either the patent owner or a third party may ask the PTO to reexamine and reevaluate the patent vis-à-vis the prior art. The PTO then proceeds to examine the patent much the same as it would with an original application.

2. Foreign Countries

International Convention for the Protection of Industrial Property. The important provision of the convention is that any person who has duly applied for a patent in one of the contracting states shall have a right of priority for 12 months in making application in the other states.

Patent Cooperation Treaty (PCT) and European Patent Convention (EPC).   To simplify the process of obtaining foreign patents, it is now possible to file a single application, e.g., in the United States, and have that application mature into a patent in each country worldwide which is a member of the Patent Cooperation Treaty. Similarly, if the focus of intended patent protection is in Europe, a single application in the European Patent Office located in Berlin, Munich, or the Hague can be issued as a patent in most of the countries on the European continent under the European Patent Convention.

Use of either the PCT or EPC provisions will generally be cost-effective versus separate filings in individual countries, since, for example, translation costs are usually minimized or eliminated. However, as in all such complex matters, knowledgeable patent professional counsel should be consulted.

TRADEMARKS

1. United States   [Section references (§) refer to Lanham Act of 1946, and 15 U.S.C., both section numbers being given.]

What Is a Trademark?   The term *trademark* includes any word, name, symbol, or device or any combination thereof adopted and used by manufacturers or merchants to identify their goods and distinguish them from those manufactured or sold by others (§45, §1127). A *service mark* is defined as a mark used in the sale or advertising of services to identify the services of one person and distinguish them from the services of others (§45, §1127).

How Rights in a Mark Are Obtained.   Unlike a patent, but somewhat like a copyright, ownership of a mark is obtained by use. If use occurs only within a state, a mark may be registered in that state. Forms for registering marks under state law are obtainable from the Secretary of State of each state. If use occurs only in interstate or foreign commerce, a mark may be registered in the U.S. Patent and Trademark Office (PTO) by filing an application in the form prescribed (37 C.F.R. Pt. IV).

After federal registration the symbol ® may be used with the mark. Prior to federal registration, the symbols TM for a trademark or SM for a service mark may be used to give notice of a claim of common law or state law tights.

A person who has a bona fide, good-faith intention to use a trademark in commerce, but is not actually using the mark, may apply for registration by filing an Intent To Use application in the PTO. Thus, a person who is not actually using a mark may still obtain federal registration provided that person commences using the mark in interstate commerce within a specified period after allowance of the Intent To Use application.

Tests for Registrability and Infringement.   A mark may be registrable on the Principal Register only if, when applied to the goods or used in connection with the services, it is not likely to cause confusion or mistake or to deceive. The same test is used to determine whether a mark infringes on an earlier mark. Marks that are descriptive, misdescriptive, geographical, or primarily merely a surname may not be registrable on the Principal Register, if it all, until after 5 years of substan-

tially exclusive and continuous use in interstate commerce (§2[f], §l052[f]). Marks not registrable on the Principal Register may be registrable on the Supplemental Register if they are capable of becoming distinctive as to the applicant's goods or services (§23, §1091).

Term of Registration. Federal registrations remain in force for 10 years unless canceled. However, unless an affidavit is filed during the sixth year after registration showing that the mark is still in use, the Commissioner of Patents and Trademarks will cancel the registration of the mark (§8, §1058). In addition, federal registrations only remain in force for marks which are actually in use. Nonuse of a registered mark can result in a finding that the mark has been abandoned by its owner. The law currently provides that nonuse of a mark for 2 years or longer is prima facie evidence that the mark has been abandoned. Within the sixth year after registration on the Principal Register a registered mark may become more secure from legal challenge, i.e, ''incontestable'' upon the filing of the required affidavit (§15[31], §1065[b]). During the last 6 months of the registration term an application for renewal may be filed (§9, §1059). The terms of state registrations vary.

Preliminary Search. Before adopting a mark, it is advisable to have a search made in the PTO to determine whether the mark under consideration would conflict with any pending application, registrations, or others' use of the mark in connection with any similar goods or services.

Cost of Registering a Mark. Government fees and time for payment are fixed by law (§31, §1113) but attorneys' fees vary.

Effect of Federal Registration. Federal registration of a mark affords nationwide protection, and once the certificate has been issued, no person can acquire any additional rights superior to those obtained by the federal registrant Federal registration of a mark establishes federal jurisdiction in an infringement action, can be the basis for treble damages, and is admissible as evidence of trademark rights. Registration on the Principal Register constitutes constructive notice, constitutes prima facie or conclusive evidence of the exclusive right to use the mark in interstate commerce, may become incontestable, and may be recorded with the United States Treasury Department to bar importation of goods bearing an infringing trademark.

Assignment of a Mark. A mark can be assigned only in conjunction with the goodwill of the business or that portion thereof with which the mark is associated. Assignments of registered marks or of applications for registration should be recorded in the United States Patent and Trademark Office within 3 months after the date of assignment (§l0, §1060).

2. Foreign Countries

Registration. In general, most foreign countries require the registration of a mark in compliance with local requirements for trademark protection. In most countries registration is compulsory and provides the sole basis for protection of a trademark. A few countries afford common-law protection to unregistered marks. Because the laws and regulations in foreign countries regarding trademark registration vary so much, it is not practicable to summarize the requirements for registration in the

space allocated to this note. Anyone interested in foreign protection of a mark should consult an attorney.

COPYRIGHTS

1. United States [Section references (§) refer to Copyright Act of 1976, 17 U.S.C.]

Subject Matter of Copyright. Copyright protection subsists in works of authorship in literary works, musical works including any accompanying words, dramatic works, including any accompanying music, pantomimes and choreographic works, pictorial, graphic, and sculptural works, motion pictures and other audiovisual works, and sound recordings, but not in any idea, procedure, process, system, method of operation, concept, principle, or discovery [§102]. However, copyright protection may be obtained for computer programs, including both the source code and object code. In addition, owners of copyrights covering computer programs (including any tape, disk, or other medium embodying such program) generally have the right to authorize or prohibit commercial rental to the public of the originals or copies of their copyrighted work. Noncommercial transfers of lawfully made copies of computer programs by a nonprofit educational institution to another nonprofit educational institution or to faculty, staff, and students does not constitute rental, lease, or lending under the copyright laws.

Method of Registering Copyright. After publication (sale, placing on sale, public distribution), registration may be effected by an application to the Register of Copyrights along with the fee ($20.00) and two copies of the writing (§40l; 407; 409). The Copyright Office (Register of Copyrights, Library of Congress, Washington, DC 20540) supplies without charge the necessary forms for use when applying for registration of a claim to copyright.

Form of Copyright Notes. The notice of copyright required (§401) shall consist either of the word "copyright," the abbreviation "Copr.," or the symbol ©, accompanied by the name of the copyright owner and the year date of first publication.

The copyright notice for phonorecords of sound recordings should contain the symbol ℗, the name of the copyright owner, and the year date of first publication.

Who May Obtain Copyright. Only the author or those deriving their rights through the author can rightfully claim copyright protection. Where a copyrighted work is prepared by an employee within the scope of his or her employment (a "work made for hire"), the employer and not the employee is considered to be the author. The name of the author (creator of the work) is normally used in the copyright notice even though an assignee may file the application for registration.

Duration of Copyright. Copyright in a work created on or after January 1, 1978, subsists from its creation and, with certain exceptions, endures for a term consisting of the life of the author and 50 years after the author's death (§302). Copyright in the work created before January 1, 1978, but not theretofore in the public domain or copyrighted, subsists from January 1, 1978, and endures for the term provided by §302. In no case, however, shall the copyright term in such a work expire before

December 31, 2002; and, if the work is published on or before December 31, 2002, the term of copyright shall not expire before December 31, 2027 (§303). Any copyright, the first term of which is subsisting on January 1, 1978, shall endure for 28 years from the date it was originally secured, with certain exceptions (§304). All terms of copyright provided by §302–304 run to the end of the calendar year in which they would otherwise expire (§305).

Infringement of Copyright. A violation of any of the exclusive rights (§1) by a person not licensed could lead to liability (§101) for infringement. Actions for infringement of copyright are brought in the U.S. district courts.

Consulting Attorneys. An attorney should be consulted before publication unless the author or proprietor has previously registered copyrights. A mistake could cause the copyright to be lost if the work is published without proper notice.

2. Foreign Countries

Universal Copyright Convention. Under this convention a U.S. citizen may obtain a copyright in most countries of the world simply by publishing within the United States using the prescribed notice, namely, © accompanied by the name of the copyright proprietor and the year of first publication placed in such manner and location as to give reasonable notice of claim of copyright. While the term ''Copyright'' on the notice is adequate for copyright under U.S. law, only the © is recognized under the convention.

REFERENCES

Chisum, ''Patents,'' Matthew Bender. Rules of Practice of U.S. Patent and Trademark Office, 37 C.F.R. §1, et seq.

Callman, ''Unfair Competition, Trademarks and Monopolies,'' Callaghan and Co., Chicago. McCarthy, ''Trademarks and Unfair Competition,'' Lawyers Cooperative, Bancroft-Whitney. Gilson, ''Trademark Protection and Practice,'' Matthew Bender.

Nimmer, ''Nimmer on Copyright,'' Matthew Bender. Copyright Office Regulations, 37 C.F.R., Chapter 2.

INDEX

ABOUT THE EDITORS

Dr. Ejup N. Ganić received his masters degree in engineering from the University of Belgrade and attained his doctorate from M.I.T. Dr. Ganic continued his academic career at the University of Illinois at Chicago where he received a tenured position. He is the author of several major engineering handbooks and numerous journal articles. He is presently with the University of Sarajevo.

Tyler G. Hicks is a consulting engineer, a licensed Professional Engineer, and a successful author. He has worked in plant design and operation in a variety of industries, taught at several engineering schools, and lectured in the United States and abroad. Mr. Hicks holds a bachelor's degree in mechanical engineering from Cooper Union School of Engineering in New York.