UNIVERSAL
LIBRARY
OU_148732
UNIVERSAL
LIBRARY

Electricity AND Magnetism

Henry E. Duckworth

McMaster University

HOLT, RINEHART AND WINSTON
NEW YORK — CHICAGO — SAN FRANCISCO

January, 1964

Copyright © 1960
by Holt, Rinehart and Winston, Inc.
Library of Congress Catalog Card Number: 59-8709

21941-0110

Printed in the United States of America

Preface

What is the justification for *another* book on Electricity and Magnetism when several excellent texts already exist? The primary justification is my own class. The students in it are not prepared to tackle an advanced text (say, senior or first year graduate level) but do require a somewhat sterner treatment than existing lower-level texts provide. Presumably, other classes are conducted at this same intermediate level.

The text is written for students who have studied general physics and the calculus and who are in the act of acquiring an elementary knowledge of differential equations. Vector language is utilized in the text, but is developed as needed. The method of presentation of the subject is governed by two considerations. First, of course, the subject is presented for its own sake, as fundamental to much of science and engineering. Second, cognizance is taken of the fact that the study of electricity and magnetism provides many students with their first real opportunity to apply the calculus to physical problems. With this consideration in mind, a number of problems are included (particularly in the early chapters) for the express purpose of providing practice in the setting-up and subsequent solution of integrals.

The reader who peruses the text for evidences of originality will be disappointed. Many of the ideas were acquired years ago from sources now forgotten, while others bear the unmistakable stamp of one or other of the standard texts on electricity and magnetism. The author's contribution, if he has made one, is to present a concise, modern introduction to the subject. Every effort has been made to write directly and economically. In keeping with the modern point of view, many natural excursions are made into important areas of atomic, nuclear, and solid state physics. These excursions are brief but should serve as suitable introductions to a number of topics that lecturers may wish to discuss in greater detail. The quantitative treatment of such topics is simplified by the consistent use of mks units. The cgse and cgsm systems of units are sufficiently explained to enable students to comprehend texts employing them.

The subject of electronics is not treated as such and, indeed, is only occasionally touched upon. Since it is the usual practice now to study this subject per se, it was decided not to add to the bulk of the present text by including a superficial (and, necessarily, indecisive) treatment of it.

A few topics, usually more difficult ones, are developed in "small type" sections. The material in these sections is in no case prerequisite to any of the normal type sections. Problems based on small type sections are so indicated.

The relative difficulty of each of the problems is indicated. Those marked **E** (for "easy") should be done by all students, even those intending to fail the course. Those marked **M** (for "medium") and **H** (for "hard") are for better students only. If the text is sufficiently well received to justify a second edition I should be most happy to receive from users of it supplementary problems to be included therein. Suitable acknowledgment would, of course, be made.

The manuscript has profited enormously from the criticisms of Dr. C. F. Mate of McMaster University, Professor E. R. Pounder of McGill University, Professor D. B. Scott of the University of Alberta, and (especially) Dr. R. G. Stoner of Pennsylvania State University. Assistance, far beyond the call of duty, has been given by Mrs. Marion Tweedley in the preparation of the manuscript. Many of the rough diagrams were prepared by my son, while Paul van Rookhuyzen has assisted in the preparation of the index. At every stage the author has enjoyed the fullest cooperation of the publishers, Holt, Rinehart and Winston, Inc. and the Macmillan Company of Canada. Their judicious prodding of the author and expeditious editing of the manuscript has greatly hastened publication. Finally, and by no means least, the affectionate encouragement of my own family has been an indispensable aid in the writing of the manuscript.

I hope that readers will draw to my attention any errors that are discovered. In addition, with an eye on the future, I should appreciate receiving suggestions for modification of the text.

H. E. D.

Hamilton, Ontario
January 1960

The second printing contains a number of corrections which have been drawn to my attention by Dr. Alan Bateson of the University of Virginia, Dr. M. H. Edwards of the Royal Military College, Dr. Peter Fong of Syracuse University, Dr. Mario Iona of the University of Denver and Dr. Dennison Bancroft of Colby College. These corrections have been much appreciated.

H. E. D.

June 1961

Contents

CHAPTER 1

Electric Forces and Electric Fields 1

1.1. Electrification by friction, *1*
1.2. The electrical nature of matter, *3*
1.3. Coulomb's law, *5*
1.4. Electric intensity or field strength, *8*
1.5. Dependence of the strength of the electric field of a point charge upon the nature of the surrounding medium, *11*
1.6. Electric field of an infinitely long charged wire, *12*
1.7. Electric field of a large charged sheet, *14*
1.8. Gauss' theorem, *16*
1.9. Some applications of Gauss' theorem, *20*
1.10. Electrical stress on the surface of a charged conductor, *24*
1.11. Energy in an electric field, *25*
1.12. Value of the elementary charge: The oil drop experiment, *27*
1.13. *Problems*, *30*

CHAPTER 2

Electric Potential 33

2.1. Some elementary vector analysis, *33*
2.2. Gauss' theorem in vector form, *36*
2.3. General concept of potential and difference in potential, *38*
2.4. Relation between potential and field, *38*

2.5. Potential arising from a point charge and from a system of point charges, *43*
2.6. Potential arising from spherically and cylindrically symmetric charge distributions, *45*
2.7. Distribution of charge on an irregularly shaped conductor, *46*
2.8. Equipotential surfaces and lines of force, *47*
2.9. Acceleration of charged particles, *49*
2.10. Electrostatic lenses, *52*
2.11. Electrical images, *54*
2.12. Use of Laplace's equation in the solution of electrostatic problems, *59*
2.13. *Problems*, *61*

CHAPTER 3
Capacitance
66

3.1. Concept and definition of capacitance, *66*
3.2. Units of capacitance, *67*
3.3. Energy of a charged conductor, *67*
3.4. Principle of the capacitor, *68*
3.5. Capacitors in series and parallel, *70*
3.6. Uses of capacitors, *72*
3.7. The parallel-plate capacitor, *74*
3.8. Spherical capacitor, *77*
3.9. Cylindrical capacitor, *78*
3.10. Comparison of dielectric constants, *79*
3.11. Electrostatic voltmeters, *80*
3.12. *Problems*, *83*

CHAPTER 4
Dielectric Theory and Behavior
86

4.1. Electric dipoles: Permanent and induced, *86*
4.2. Potential and field arising from an electric dipole, *88*
4.3. Quadrupole and higher moments, *89*
4.4. Electric polarization and electric susceptibility, *90*
4.5. Electric displacement, *92*
4.6. Maxwell's displacement current, *94*
4.7. Boundary conditions at the interface between two dielectric materials, *96*

4.8. Refraction of lines of force in passing from one dielectric medium to another, *98*
4.9. Relation between the dielectric constant and the index of refraction, *100*
4.10. Field acting upon a molecule within a dielectric: The "inner field," *101*
4.11. Effect of temperature on molecular polarizability, *105*
4.12. Molecular polarizability and molecular structure, *109*
4.13. Ferroelectricity, *111*
4.14. Piezoelectricity and pyroelectricity, *113*
4.15. *Problems*, *113*

CHAPTER 5

Direct Currents

116

5.1. Electric current and current density, *116*
5.2. Electrical conduction in a metal: The electron gas, *119*
5.3. Motion of conduction electrons under an applied field: Ohm's law, *119*
5.4. Refraction of lines of current flow at a boundary between conductors of different resistivity, *122*
5.5. Laplace's equation for a steady current, *123*
5.6. Resistances in series and parallel, *124*
5.7. Energy dissipation resulting from DC current flow, *125*
5.8. Electromotive force, *127*
5.9. Condition for maximum power transfer, *131*
5.10. Advantages of various cell arrangements within a battery, *132*
5.11. Kirchhoff's laws, *134*
5.12. Solution of circuits by method of cyclic currents, *136*
5.13. Some electrical circuit theorems, *138*
5.14. Faraday's laws of electrolysis, *140*
5.15. Determination of current by electrolytic deposition: The voltameter, *142*
5.16. *Problems*, *143*

CHAPTER 6
Resistance
146

6.1. The band theory of solids, *146*
6.2. The Fermi-Dirac distribution, *146*
6.3. Insulators, conductors, and semiconductors, *148*
6.4. Resistivity and its use in the calculation of the resistance of ohmic materials, *151*
6.5. Variation of resistance with temperature, *153*
6.6. Superconductivity, *156*
6.7. Nonohmic or nonlinear resistances, *157*
6.8. Voltmeters and ammeters: Their use in the determination of resistance, *160*
6.9. Wheatstone bridge, *163*
6.10. The Kelvin double bridge, *168*
6.11. Platinum resistance thermometer, 169
6.12. Resistance standards, *170*
6.13. The potentiometer, *171*
6.14. Standard cells, *172*
6.15. Direct-reading potentiometers, *173*
6.16. Some applications of potentiometers, *175*
6.17. Relation between electrical and thermal conductivity: The Wiedemann-Franz law, *176*
6.18. *Problems*, *177*

CHAPTER 7
Thermoelectricity
180

7.1. Seebeck effect, *180*
7.2. Peltier effect, *181*
7.3. Thomson effect, *182*
7.4. Seebeck emf is the algebraic sum of Peltier and Thomson emf's, *184*
7.5. Thermoelectric power, *186*
7.6. Effect of other metals in a thermocouple circuit: Law of intermediate metals, *186*
7.7. Thermocouple thermometers, *188*
7.8. Thermopiles, *190*
7.9. Thermocouple meters, *190*
7.10. Thermoelectric generators and refrigerators, *191*
7.11. Source of thermal emf, *192*
7.12. *Problems*, *193*

CHAPTER 8

Moving Charges and Magnetic Fields

194

8.1. Force between electric currents, *194*
8.2. Force exerted by a magnetic field upon a charge in motion, *196*
8.3. Magnetic "fields," *198*
8.4. The magnetic induction of a current element: Ampère's law, *198*
8.5. Magnetic induction at the center of a circular current: The electromagnetic system of units, *200*
8.6. Magnetic induction along axis of a circular current, *202*
8.7. Magnetic induction along axis of a solenoid, *204*
8.8. Magnetic induction of a straight wire carrying a current, *205*
8.9. Line integral of the induction: Ampère's law in circuital form, *207*
8.10. Maxwell's second equation, *210*
8.11. Magnetic vector potential, *211*
8.12. Further discussion of the magnetic force exerted upon an electric current or a moving charge, *215*
8.13. Definition of the coulomb and the ampere, *218*
8.14. The mass spectroscope, *219*
8.15. The cyclotron, *221*
8.16. The Hall effect, *224*
8.17. Focusing of charged particles by magnetic lenses, *225*
8.18. *Problems*, *228*

CHAPTER 9

Electromagnetic Induction

230

9.1. Lenz's law, *230*
9.2. Magnitude of the induced emf: Faraday's law of electromagnetic induction, *231*
9.3. Faraday's law derived from Ampère's law, *233*
9.4. Alternative deduction of the direction of an induced current in a conductor moving in a magnetic field, *235*
9.5. Maxwell's third equation, *236*
9.6. Inductively coupled circuits: Mutual inductance and self-inductance, *236*
9.7. Inductances in series, *239*

9.8. Energy stored in a coil through which a current is flowing, *240*
9.9. Energy in a magnetic field, *241*
9.10. Calculation of self-inductance and mutual inductance, *243*
9.11. Eddy currents, *248*
9.12. The betatron—an induction accelerator, *250*
9.13. *Problems*, *252*

CHAPTER 10

The Magnetic Properties of Matter

254

10.1. History of magnetism, *254*
10.2. Force between magnetic poles: Coulomb's law for magnets, *255*
10.3. The electromagnetic system of units, *257*
10.4. The field of a magnetic dipole and the torque experienced by a dipole in a magnetic field, *259*
10.5. Magnetic dipole equivalence of a current loop, *261*
10.6. The magnetic shell and its closed current equivalent, *262*
10.7. Magnetic moment of an orbital electron: The Bohr magneton, *264*
10.8. Potential energy associated with the orientation of a magnetic dipole or a closed current in a uniform magnetic field, *267*
10.9. Zeeman effect, *268*
10.10. Force on a dipole in a nonuniform magnetic field: The Stern-Gerlach experiment, *270*
10.11. Gauss' theorem for magnetism, *272*
10.12. Refraction of $\boldsymbol{B}$ and $\boldsymbol{H}$ in passing from one magnetic medium to another, *273*
10.13. Magnetic polarization and magnetic susceptibility, *275*
10.14. Diamagnetism, paramagnetism, and ferromagnetism, *277*
10.15. Diamagnetism, *278*
10.16. Paramagnetism and its variation with temperature, *279*
10.17. Ferromagnetism, *281*
10.18. The magnetization curve, *287*

10.19. Hysteresis and the hysteresis loop, *289*
10.20. Experimental tracing of hysteresis loops, *292*
10.21. The magnetic circuit, *296*
10.22. Electromagnets, *298*
10.23. *Problems*, *299*

CHAPTER 11

Galvanometers, Ammeters, Voltmeters, and Related Instruments

302

11.1. The moving coil or d'Arsonval galvanometer, *302*
11.2. Sensitivity of a galvanometer, *304*
11.3. Galvanometer damping, *307*
11.4. The ballistic galvanometer, *316*
11.5. Meters operating on the moving-coil principle, *321*
11.6. Electrodynamometer action, *322*
11.7. Moving-iron instruments, *325*
11.8. Rectifier instruments, *325*
11.9. Thermocouple and hot-wire instruments, *326*
11.10. Electrometers, *327*
11.11. *Problems*, *330*

CHAPTER 12

Transient Currents

331

12.1. Discharge of a capacitor through a resistor, *331*
12.2. Determination of the value of a high resistance by the rate of discharge of a capacitor, *333*
12.3. Charging a capacitor through a resistor, *333*
12.4. Relaxation oscillator, *334*
12.5. Circuits containing inductance and resistance, *335*
12.6. Discharge of a capacitor through a resistor and inductor, *336*
12.7. Charge of a capacitor through a resistor and inductor, *338*
12.8. The Tesla oscillator, *339*
12.9. *Problems*, *340*

CHAPTER 13

Alternating Currents

342

13.1. The simple AC generator, *342*
13.2. Graphical representation of alternating emf's and currents, *344*
13.3. Definition of unit AC current, *346*
13.4. Effective or RMS value of an AC current, *346*
13.5. AC current and voltage in a series LR circuit, *348*
13.6. AC current and voltage in a series RC circuit, *354*
13.7. AC current and voltage in a series LCR circuit, *356*
13.8. Power dissipation in an AC circuit, *359*
13.9. Resonance in series circuits, *360*
13.10. Potential differences at resonance, *361*
13.11. Sharpness of resonance: Quality of a circuit, *362*
13.12. Parallel circuits, *363*
13.13. Complex algebra, *365*
13.14. Application of complex algebra to series circuits, *368*
13.15. Application of complex algebra to parallel circuits, *370*
13.16. Parallel resonance, *371*
13.17. Admittance, conductance, and susceptance, *372*
13.18. Kirchhoff's laws in complex form, *373*
13.19. AC bridges, *374*
13.20. Condition for maximum power transfer in AC circuits, *376*
13.21. The transformer, *377*
13.22. AC voltages and currents as rotating vectors, *381*
13.23. *Problems*, *382*

CHAPTER 14

Maxwell's Equations and Electromagnetic Waves

386

14.1. Plane electromagnetic waves in free space, *387*
14.2. Use of the exponential form to represent plane harmonic waves, *393*
14.3. Reflection and refraction of plane electromagnetic waves, *395*
14.4. Amplitudes and phases of plane

waves reflected and refracted at interfaces between different dielectric media, *398*

14.5. Electromagnetic waves in a conducting medium, *403*

14.6. Conclusion, *404*

14.7. *Problems*, *405*

APPENDIX

Some Useful Vector Relationships

406

INDEX

411

ELECTRICITY AND MAGNETISM

Chapter 1

Electric Forces and Electric Fields

The study of electricity and magnetism falls into two main divisions, one concerned with static electrical phenomena and the other with charges in motion. The former, known as electrostatics, provides the natural setting in which to develop the concepts of electric field and potential. The latter, known as electrodynamics, includes both current electricity and electromagnetism.

Before it was realized that all magnetic effects stem from the motion of charges, magnetism was itself treated as a static affair (magnetostatics). As this point of view is still a useful one, and as the method of treatment closely follows that of electrostatics, many of the concepts that are derived in the first few chapters will be directly applicable in the later study of electromagnetism.

1.1 Electrification by Friction

The history of static electricity is a long and venerable one. It began, at least by tradition, in the sixth century B.C., with the observation by the Greek natural philosopher, Thales of Miletus, that amber when rubbed attracts small objects to itself. The real founder of the science, however, may be considered to be William Gilbert (c1540–1603), physician to Elizabeth I, who carefully repeated the observations of the ancients and systematically extended them.

His *De Magnete* (1600) listed many materials, including not a few quaint ones, that possess the same property as amber, and a second group of substances that are not excited by friction.

The discovery that two types of electricity exist was made by Dufay (1699–1739), superintendent of the French Royal Botanical Gardens. He was responsible for giving the name *vitreous* electricity to that produced when glass, wool, and certain other substances are excited, and *resinous* electricity to that produced by exciting amber, silk, etc. These two types of electricity were distinguished from one another by the fact that a body with vitreous electricity attracts bodies with resinous electricity but repels other bodies with vitreous electricity, and vice versa.

Benjamin Franklin (1706–1790), American statesman, printer, scientist, and writer, demonstrated that electricity is not created during electrification. Rather, the "electrical fire" is transferred from one body to another, the body that acquires more than its usual amount becoming *positively* charged (vitreous electricity) while the other exhibits an equal *negative* (resinous) charge.

It was subsequently observed by John Canton (1718–1772), a diligent English investigator of electrical effects, that a given body may exhibit positive electricity when rubbed with one material, but negative electricity when rubbed with a second. Indeed, substances can be arranged in a *triboelectric* series, shown below, in which they are listed according to their relative ability to assume a positive charge during friction. A substance will be electrified positively on being rubbed by any substance below it in such a series, but will be electrified negatively if the rubbing substance lies above it.

Rabbit's fur
Glass
Wool
Cat's fur
Silk
Cotton
Amber
Sulfur

Thus, long before the nature of electricity was understood, it had been established experimentally that:

a. Like charges of electricity repel, unlike charges attract;
b. Positive and negative charges appear in equal amounts during electrification.

Also, as we shall see, the magnitude of the force between charged bodies had been studied carefully, as well as the effect upon it of the medium in which the charged bodies are located.

1.2 The Electrical Nature of Matter

It had been suspected from classical times that all matter is composed of aggregations of minute particles or *atoms*. This suspicion was confirmed by the work of nineteenth-century chemists and physicists, although John Dalton's (1766–1844) belief that atoms are indivisible is no longer held.

Atoms are thought to consist of negatively charged particles called *electrons*, positively charged particles called *protons*, and neutral particles called *neutrons*. The electronic and protonic charges are equal in magnitude, although opposite in sign. The protons and neutrons are approximately equal in mass ($\sim 1.67 \times 10^{-27}$ kg) and are located in a small central region, or *nucleus*, which is consequently positively charged. The mass of an electron ($\sim 9.11 \times 10^{-31}$ kg) is $\sim 1/1840$ that of a proton. The electrons move around the nucleus in orbits which are large by comparison with the dimensions of the nucleus itself. In accordance with the quantum theory, only certain electron orbits are permitted,[1] and a definite energy is associated with each. The jumping of an electron from one permitted orbit to another occurs when energy is radiated or absorbed. Inner orbits correspond to lower energy states. In a normal atom the number of extranuclear electrons is equal to the number of nuclear protons, with the result that the atom is electrically neutral.

The chemical properties of the atom are predominantly determined by the number of nuclear protons, a quantity known as the *atomic number* and usually designated by Z. Each different Z corresponds to a different *element*, of which there are roughly one hundred now known. Hydrogen, with a single proton in the nucleus, has $Z = 1$, whereas the atomic number of uranium, the heaviest naturally occurring element, is 92.

The simplest atom, that of light hydrogen (H^1), is represented in Figure 1.1 on the left, while that of heavy hydrogen (or deuterium) is shown on the right. These sketches are, of course, not drawn to scale.

Atoms such as these, which differ in mass but possess similar chemical properties, are called *isotopes*. This name is derived from two Greek words, *isos*, meaning "the same," and *topos*, meaning "place." Thus, in a table in which the elements are arranged according to their chemical properties (the Periodic Table), isotopes of the same element occupy the *same place*. Normal

[1] It was suggested in 1913 by the Danish theoretical physicist, Niels Bohr (1885–), that a permitted orbit is one for which the angular momentum of the electron is an integral multiple of $h/2\pi$, where h is an important constant, known as *Planck's constant*, having the value 6.62×10^{-34} joule-sec. This quantum condition is termed *Bohr's First Postulate* or the *Bohr Angular Momentum Principle*. The symbolic statement of the principle is $mvr = nh/2\pi$; where $n = 1, 2, 3, \cdots$, m is the electronic mass, v is the velocity of the electron as it traverses its orbit, and r is the radius of the orbit.

hydrogen, as found in nature, contains both the lighter isotope, H^1, and the heavier, D^2, the abundance of the former being ~6600 times that of the latter. A third isotope of hydrogen, H^3 (or *tritium*, T^3), has been produced artificially but is not stable, decaying to He^3 with a half life of 12.46 years.

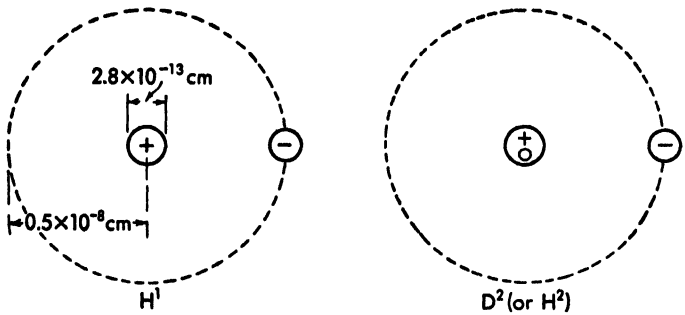


Fig. 1.1

The superscripts 1, 2, and 3 which have been used to distinguish the three isotopes of hydrogen are called *mass numbers* and refer to the numbers of nuclear constituents or *nucleons*, as the protons and neutrons are collectively termed. Because of the approximate equality of the proton and neutron masses, and the relative insignificance of that of the electron, the mass number gives a useful approximate indication of the mass of the isotope concerned. Some representative masses are $H^1 = 1.00814$ atomic mass units, $Cu^{63} = 62.94961$ amu, and $U^{238} = 238.124$ amu, and in all other cases there is a similar approximate equality of the mass number and the actual mass. One *atomic mass unit*, on the physical scale, is defined as one-sixteenth the mass of O^{16}. This unit of mass differs slightly, but significantly, from the unit of *chemical atomic weight*, which is defined as one-sixteenth the *average* mass of the oxygen atoms. Because of the presence in nature of O^{17} (0.037 percent) and O^{18} (0.204 percent), in addition to the abundant O^{16}, the chemical unit of mass is ~1.000275 times heavier than the physical one.[a]

Most of the naturally occurring elements form stable isotopes. The world's record is held by the element tin, which possesses a total of ten: Sn^{112}, Sn^{114}, Sn^{115}, Sn^{116}, Sn^{117}, Sn^{118}, Sn^{119}, Sn^{120}, Sn^{122}, and Sn^{124}. Twenty-five to thirty of the elements are *anisotopic*, that is, they exist in only one stable form. The remainder of the naturally occurring elements possess isotopes ranging in number from two to eight. Altogether, there are about three hundred stable atomic species.

[a] It appears probable, at the time of writing, that the physical and chemical scales of mass will be unified by basing each upon the mass of the abundant isotope of carbon, C^{12}. If so, the atomic mass unit will be defined as one-twelfth the mass of C^{12}.

In the modern picture of the atom the electrons are the vulnerable entities and can often be removed from the atom without much difficulty. The protons, on the other hand, are located in an almost impregnable position. The act of *charging* a substance consists of adding or removing electrons—a process not unlike that pictured by Franklin, although it must be admitted that he visualized only one type of electricity. The process of adding electrons to or subtracting electrons from an individual atom is known as *ionization*.

1.3 Coulomb's Law

The first systematic quantitative study of the forces between electrically charged objects was conducted in 1795 by Charles Coulomb (1736–1806), a French engineer. On the basis of measurements made with a sensitive torsion balance, Coulomb was led to the conclusion that the force between point charges varies directly as the magnitude of each charge and inversely as the square of the distance between them. This law of force was actually discovered somewhat earlier by Henry Cavendish (1731–1810), an English scientific recluse, but lay buried in his personal papers until 1879. Perhaps no better example exists to support the dictum that a scientific investigation is not complete until its results have been published.

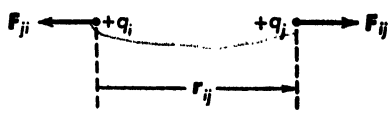


Fig. 1.2

Let us consider the charge configuration shown in Figure 1.2. Two point charges of magnitude q_i and q_j are located in a vacuum and are separated by the distance r_{ij}. The symbolic statement of Coulomb's law is, therefore,

$$F_{ij} \propto \frac{q_i q_j}{r_{ij}^2} \qquad \text{or} \qquad F_{ij} = \frac{q_i q_j}{C r_{ij}^2} \tag{1-1}$$

The value of the factor of proportionality, C, depends upon the system of units employed. This matter is discussed further below.

When it is desired explicitly to express the vector character of the electrostatic force, which is not done in equation [1-1], the symbol for the force will be printed in bold face, as $\boldsymbol{F}_{ij}$ or $\boldsymbol{F}_{ji}$. Other vectors will be similarly designated. The vector representation corresponding to equation [1-1] is

$$\boldsymbol{F}_{ij} = \frac{q_i q_j}{C r_{ij}^3} \boldsymbol{r}_{ij} \tag{1-2}$$

where $\boldsymbol{r}_{ij}$, the vector distance from q_i to q_j, gives the direction of $\boldsymbol{F}_{ij}$, the force exerted by q_i upon q_j. The inclusion of $\boldsymbol{r}_{ij}$ in the numerator creates the necessity

for an additional r_{ij} in the denominator in order to make equation [1-2] equal in magnitude to equation [1-1]. Evidently,

$$\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji} \qquad [1\text{-}3]$$

The total force exerted upon a point charge, q_j, by n point charges is given by

$$\boldsymbol{F}_j = \sum_{i=1}^{i=n} \boldsymbol{F}_{ij} = \sum_{i=1}^{i=n} \frac{q_i q_j}{C r_{ij}^3} \boldsymbol{r}_{ij} = \frac{q_j}{C} \sum_{i=1}^{i=n} \frac{q_i}{r_{ij}^3} \boldsymbol{r}_{ij} \quad (i \neq j) \qquad [1\text{-}4]$$

This inverse-square law of electrical force is closely analogous to the law of universal gravitation proposed by Isaac Newton (1642–1727) a century earlier to describe gravitational attraction between objects. We shall see, in Chapter 10, that a similar expression describes the force between magnetic *poles*.

The Electrostatic Unit of Charge

The force of attraction or repulsion, as given in equation [1-1], was used by Coulomb as the basis for the definition of the first unit of charge, the so-called *electrostatic unit*. This unit was chosen to make $C = 1$ when the force F_{ij} is measured in dynes and the separation r_{ij} is measured in centimeters. Thus, the electrostatic unit of charge is defined as *that point charge which, when placed 1 cm from a similar point charge in free space (a vacuum), is repelled by a force of 1 dyne.* The electrostatic unit of charge may be abbreviated to *cgse* unit of charge (cm-gm-sec-electrostatic), and is also frequently called the *statcoulomb*. The statement of Coulomb's law in the cgse system is then

$$F_{ij} = \frac{q_i q_j}{r_{ij}^2} \text{ dynes} \qquad [1\text{-}5]$$

or, in vector form, the force of repulsion experienced by q_j is

$$\boldsymbol{F}_{ij} = \frac{q_i q_j}{r_{ij}^3} \boldsymbol{r}_{ij} \text{ dynes} \qquad [1\text{-}6]$$

If other units are employed for the charge, force, and distance, the factor of proportionality in equation [1-1] is, of course, not necessarily equal to unity, as we shall shortly have occasion to observe.

The mks System of Units

For many years three different systems of units have been employed in the study of electricity and magnetism. The first of these is the electrostatic system, referred to in the preceding paragraph, which has traditionally been employed in the study of electrostatic phenomena. Current electricity, on the other hand, has been discussed in terms of *practical* units—the *ampere*, *volt*, *ohm*, *coulomb*, etc.

Finally, magnetic effects have been described in *electromagnetic* units, which are discussed briefly in Chapters 8 and 10.

This three-unit usage has a natural historic basis, but it is needlessly complicated and is now being replaced, following official action in 1950 of the International Electrochemical Commission, by the *rationalized mks* system of units. This new system is derived from mechanics, but in electricity it coincides with the old practical system.

Three of the primary units in the mks system are, as its name suggests, the *meter* (abbreviated m), *kilogram* (kg), and *second* (sec). The unit of force is the *newton* (1 n = 10^5 dynes), defined (with reference to Newton's second law) as that force which will give to a mass of 1 kg an acceleration of 1 meter per sec^2. The unit of work is the *joule* (1 joule = 10^7 ergs), defined as the work done when a force of 1 newton acts through a distance of 1 meter.

The fourth primary unit is that of charge, the *coulomb* (coul). This is specifically indicated by designating the system as the *mksc* system.[8] In practice, the "c" portion of "mksc" is usually understood, rather than written. The coulomb and the corresponding unit of current, the *ampere*, are defined in § 8.13. The coulomb, incidentally, is very large by comparison with the statcoulomb mentioned above; in fact 1 coulomb = 3×10^9 statcoulombs.

Coulomb's Law in the mks System of Units

Equation [1-5], which expresses Coulomb's law in the cgse system, is not terse by accident. Instead, its simplicity ensues from the fact that the cgse unit of charge is defined in terms of Coulomb's law. The freedom to adopt a simplifying definition does not exist in the mks system. Here the unit of charge (the coulomb), the unit of force (the newton), and the unit of distance (the meter) have each been defined already in some other context. Hence, the factor of proportionality in the mks statement of Coulomb's law is an experimentally determined quantity.

It is known from experiment that if a point charge of 1 coulomb were to be placed 1 meter from a similar point charge in free space it would be repelled with a force of $1/(1.113 \times 10^{-10})$ newtons. Thus, in the mks system, the factor of proportionality in equation [1-1] has the value

$$\begin{aligned} C &= 1.113 \times 10^{-10} \text{ coul}^2/\text{n-m}^2 \\ &= 4\pi(8.85 \times 10^{-12} \text{ coul}^2/\text{n-m}^2) \\ &= 4\pi\epsilon_0 \end{aligned} \qquad [1\text{-}7]$$

[8] Inasmuch as a current of one ampere represents the passage of one coulomb of charge per second, and it is the ampere, not the coulomb, which is operationally determined, many physicists speak of the mks*a*, rather than the mksc system.

The quantity 4π, although cumbersome, is introduced in order to simplify equations that will be developed later. The quantity $\epsilon_0 = 8.85 \times 10^{-12}$ coul^2/n-m^2 is known as the *permittivity* of free space. The mks statement of Coulomb's law, is therefore

$$F_{ij} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}^2} \text{ newtons}^4 \qquad [1\text{-}8]$$

or, in vector form,

$$\boldsymbol{F}_{ij} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}^3} \boldsymbol{r}_{ij} \text{ newtons} \qquad [1\text{-}9]$$

In keeping with the modern trend in physics and the recommendation of the International Electrotechnical Commission, we shall employ mks units almost exclusively. The brief reference made above to electrostatic units should enable the student to correlate any results derived herein with those appearing in older texts.

It is now known that the basic electrical charge is that of the electron or proton. The method by which the magnitude of this elementary charge was determined will be discussed in § 1.12. In the meantime we shall anticipate this discussion by noting that the value of the elementary charge (designated by e) is 1.6×10^{-19} coulombs. Thus, although electricity is granular, it is pretty fine-grained. According to convention, the protonic charge is $+e$ and the electronic charge is $-e$.

1.4 Electric Intensity or Field Strength

Surrounding any charge or group of charges is a region known as the *electric field*, in which the effect of the presence of these charges is felt. Let us suppose that an additional small or "test" charge, q, is brought into the region and, at a certain point in it, experiences an electrostatic force, $\boldsymbol{F}$. The electric *intensity* or *field strength*, designated by $\boldsymbol{E}$, at the point in question is a vector defined by the equation

$$\boldsymbol{E} = \frac{\boldsymbol{F}}{q} \qquad [1\text{-}10]$$

This definition assumes that the test charge is small enough to ensure that its presence in the region does not disturb the existing charge configuration. The importance of this proviso may be underlined by rewriting equation [1-10] in the form

$$\boldsymbol{E} = \lim_{q \to 0} \frac{\boldsymbol{F}}{q} \qquad [1\text{-}11]$$

---

[4] It should be noted, for convenience in calculation, that the numerical value of $1/4\pi\epsilon_0$ is very closely 9×10^9.

Thus, although the field strength has units of force per unit charge, it is not generally true that its value for an existing field can be determined by actually bringing a unit charge to the point in question and measuring the force which it experiences. Certainly one would be chary of doing so if he were using the coulomb ($\sim 10^{19}$ elementary charges) as the unit of charge. In many cases the presence of a test charge of 1 coulomb would give rise to induced charges whose effects could mask completely those of the original charge configuration. In this way a value for the field strength could be found, but it would be the value with the test charge present and not, as desired, its value prior to the importation of the test charge.

The force $\boldsymbol{F}_k$ that is experienced by a charge q_k at any point in an electric field is simply the product of the field strength and the (algebraic) value of the charge:

$$\boldsymbol{F}_k = \boldsymbol{E}q_k \qquad [1\text{-}12]$$

As suggested above, although this relation is always true, it may sometimes happen, particularly if q_k is large, that the value of $\boldsymbol{E}$ to be used in equation [1-12] is not the value that prevailed prior to the arrival of q_k.

Field of an Isolated Point Charge

Let us consider, with the aid of Figure 1.3, the strength of the field arising from an isolated point charge, q_i. In this case, since q_i is far removed from any other object, the intensity to which it gives rise may be ascertained with a test charge of *any* magnitude. As indicated in the figure, we shall employ a test charge, q.

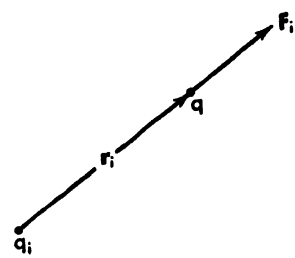


Fig.1.3

From Coulomb's law (in the mks system) the force $\boldsymbol{F}_i$, exerted upon the test charge located in a vacuum at a vector distance $\boldsymbol{r}_i$ from the point charge q_i, is

$$\boldsymbol{F}_i = q\left[\frac{q_i}{4\pi\epsilon_0 r_i^3}\boldsymbol{r}_i\right] \qquad [1\text{-}13]$$

From this, the intensity or field strength $\boldsymbol{E}_i$, a distance $\boldsymbol{r}_i$ from the isolated point charge q_i, is seen to be

$$\boldsymbol{E}_i = \frac{\boldsymbol{F}_i}{q} = \frac{q_i}{4\pi\epsilon_0 r_i^3}\boldsymbol{r}_i \qquad [1\text{-}14]$$

or

$$E_i = \frac{q_i}{4\pi\epsilon_0 r_i^2} \qquad [1\text{-}15]$$

Equation [1-14] is of fundamental importance inasmuch as every charge configuration consists ultimately of an assemblage of point charges. Hence, the total field, $\boldsymbol{E}$, arising from an assemblage may be computed by adding vectorially the fields of the individual charges:

$$\boldsymbol{E} = \sum_1^n \boldsymbol{E}_i = \frac{1}{4\pi\epsilon_0} \sum_1^n \frac{q_i}{r_i^3} \boldsymbol{r}_i \qquad [1\text{-}16]$$

This principle is applied in § 1.6 and thereafter.

Relation Between Field Strength and Lines of Force

Lines of force have been used since the time of Michael Faraday (1791–1867) to represent the directions and magnitudes of electric (and magnetic) fields. Each line of force is so drawn that at every point its tangent gives the direction of the electric field vector while the density of lines of force (that is, the number per unit cross-sectional area [6]) is proportional to the field strength. Each line of force is pictured as originating on a positive charge and terminating on a negative one. The lines of force associated with three common charge configurations are shown in Figure 1.4.

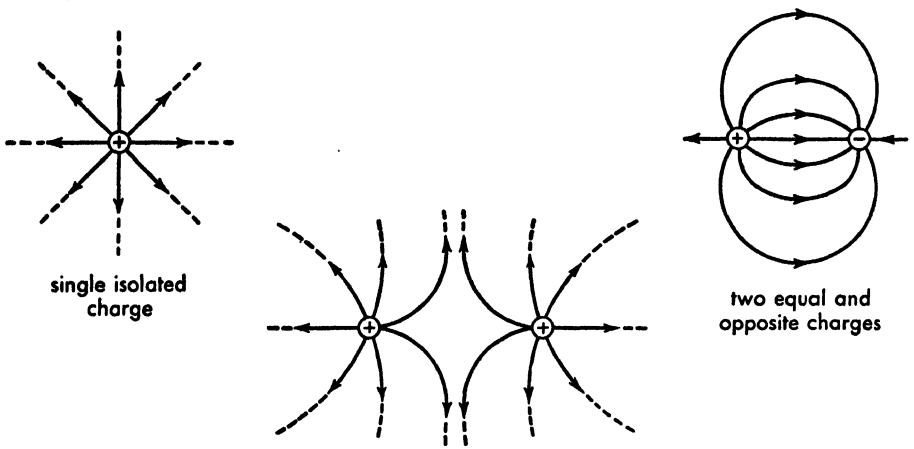


Fig. 1.4

Faraday pictured lines of force as repelling one another laterally while at the same time attempting to contract longitudinally. These characteristics are evident in Figure 1.4. The chief merit of the line-of-force language is the advantage which it offers in graphical representation.

[6] The cross-sectional area is chosen, of course, to be perpendicular to the lines of force.

1.5 Dependence of the Strength of the Electric Field of a Point Charge upon the Nature of the Surrounding Medium

In media other than free space the force between charges is less than its free space value. If the dielectric medium is homogeneous and electrically isotropic (and, of course, uncharged), the force between point charges is written as

$$\boldsymbol{F}_{ij} = \frac{q_i q_j}{4\pi\epsilon r_{ij}^3}\boldsymbol{r}_{ij} \qquad [1\text{-}17]$$

where ϵ, the permittivity of the medium concerned, is known (usually, and in this text) as the *dielectric constant* and (occasionally) as the *capacitivity*. For all material media $\epsilon > \epsilon_0$. Reference is often made to the dimensionless ratio $k_e = \epsilon/\epsilon_0$, known variously as the *relative capacitivity* and (as in this text) as the *dielectric coefficient*.[6]

In equations [1-14] and [1-15], giving the field of a point charge, ϵ_0 must be replaced by ϵ in the presence of a dielectric. From now on, in the interests of generality, we shall use ϵ rather than ϵ_0, with the understanding that $\epsilon = \epsilon_0$ in free space.[7]

The field at a distance r_i from a point charge q_i is then written

$$E_i = \frac{q_i}{4\pi\epsilon r_i^2} \quad \text{or} \quad \boldsymbol{E}_i = \frac{q_i}{4\pi\epsilon r_i^3}\boldsymbol{r}_i \qquad [1\text{-}18]$$

At first sight this reduction in the electric force in the presence of a dielectric may seem incongruous. A charged object contains the same number

unsuitable

[6] In a number of texts the dielectric constant is defined as the ratio of the force between point charges measured in a vacuum to the measured value when the vacuum is replaced by the dielectric. In this definition ϵ is identical with the dimensionless ratio, k_e, which we have defined above. In terms of this usage, Coulomb's law is written

$$\boldsymbol{F}_{ij} = \frac{q_i q_j}{4\pi\epsilon\epsilon_0 r_{ij}^3}\boldsymbol{r}_{ij}$$

and expressions stemming from it are similarly modified.

The usage that we are employing, $\epsilon = k_e\epsilon_0$, is recommended by the Commission on Symbols, Units and Nomenclature of the International Union of Pure and Applied Physics, and is used in many recent texts.

[7] The simple replacement of ϵ_0 by ϵ in the presence of a homogeneous, isotropic dielectric is valid only if the dielectric completely fills the region occupied by the electric field. This condition is satisfied in the ideal cases which we shall study in the first three chapters. In Chapter 4 we shall consider some cases in which cavities exist in the dielectric. We shall then find that the value of $\boldsymbol{E}$ is not simply given by $\boldsymbol{E}_0/k_e$, but depends strongly upon the shape of the cavity. See also "Report of Coulomb's Law Committee of the American Association of Physics Teachers," *American Journal of Physics*, 18, p. 6 ff. (1950).

of charges whether located in a vacuum or surrounded by an insulating material medium, and, consequently, the same number of lines of force emerge from it in either case. And yet the force which is exerted on a nearby charge falls off when the material medium is introduced.

The explanation of this effect lies in the electrical nature of the dielectric medium. In the molecules of many dielectrics the constituent positive and negative charges, although equal in number, are not distributed uniformly with respect to one another. As a result, to the outside observer, the molecule is seen to be *polarized*, that is, its positive charge is slightly displaced from its negative charge. Electrically, the molecule is equivalent to a positive charge separated slightly from an equal and opposite charge. Molecules of this kind are known as *polar molecules*. When placed near a free charge—say, a positive one—these polar molecules orient themselves, to the extent that circumstances warrant, with their negative ends pointing towards the positive charge. This effectively neutralizes a portion of the positive charge and diminishes the electrical field to which it was giving rise.

Further, in dielectrics in which polar molecules do not normally exist, temporary polar molecules are induced under the action of an electric field. In this case the electric field causes a shifting of the orbital electrons. The induced polar molecule so produced has the same effect upon the field as though it were a permanent polar molecule. Induced and permanent polar molecules are further discussed in § 4.1. Both the rotation of permanent polar molecules and the induction of temporary ones represent a *displacement* of charges in the dielectric.

1.6 Electric Field of an Infinitely Long Charged Wire

We shall now derive the expression for the strength of the external electric field of an infinitely long charged wire, such as that represented in Figure 1.5. This result is of particular significance because of the widespread use of wires in electrical apparatus.

Let us assume that the density of charge on the wire is λ per unit length. Then the charge contained in the element of wire of length dy is $\lambda\, dy$.[8] This may be considered a point charge, in which case, as in equation [1-18], it gives

[8] This type of statement, in which reference is made to the charge contained in an infinitesimal element of material, implies the existence of a continuous charge distribution. In actual fact, at the microscopic level the charge distribution is discontinuous. In most cases that we shall consider in this text, however, the number of electric charges involved will be sufficiently large as to constitute a continuous distribution at the macroscopic level. We are, therefore, justified in speaking of the charge in an element of length, area, or volume, as the case may be. We shall employ the symbols λ, σ, and ρ for the *linear charge density*, the *superficial charge density*, and the *volume charge density*, respectively.

rise to a field at the point P of magnitude

$$dE = \frac{\lambda\, dy}{4\pi\epsilon l^2} \qquad [1\text{-}19]$$

and possessing the direction shown in the figure. The total field strength at P arising from the entire wire is the sum of the many elemental contributions of this type.

Before proceeding with this integration, however, we should note that there exists in the wire a second element, designated in the figure as dy', which is also located at a distance l from P. This gives rise to a field dE' which is equal in magnitude to dE, although differing in direction. In fact, the vertical components of dE and dE' cancel one another, and will do so for each such pair of elements. For this reason, the total field at P is the sum of the horizontal components of the elemental fields, that is,

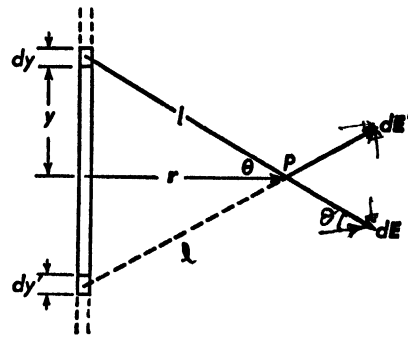


Fig. 1.5

$$E = \frac{\lambda}{4\pi\epsilon}\int_{-\infty}^{+\infty} \frac{\cos\theta\, dy}{l^2} \qquad [1\text{-}20]$$

Acting on the principle that an integration with respect to an angle is generally simpler than with respect to a linear dimension, we shall express the variables l and dy in terms of θ:

$$y = r\tan\theta$$

$$dy = r\sec^2\theta\, d\theta \qquad \text{and} \qquad l = r\sec\theta \qquad [1\text{-}21]$$

After the obvious limits of integration are assigned, equation [1-20] takes the form

$$E = \frac{\lambda}{4\pi\epsilon r}\int_{-\pi/2}^{\pi/2} \cos\theta\, d\theta = \frac{\lambda}{4\pi\epsilon r}\sin\theta\Big]_{-\pi/2}^{\pi/2} \qquad [1\text{-}22]$$

Thus, we obtain $$E = \frac{\lambda}{2\pi\epsilon r} \qquad \text{or} \qquad E = \frac{\lambda}{2\pi\epsilon r^2}\, r \qquad [1\text{-}23]$$

This result, that the field strength varies inversely as the first power of the distance, might have been anticipated from the following line of reasoning. Because the wire is infinitely extensive in one dimension, the lines of force are always normal to it, as shown in the left-hand section of Figure 1.6, and do not diverge in this plane. In the plane normal to the wire, however, as shown in the right-hand section of the figure, the lines of force diverge steadily, their

density decreasing linearly with the distance from the wire. This is in contrast to a point charge, from which the lines of force diverge in *all* directions.

If the wire be of finite length, the point at which the field strength is to be evaluated may be symmetrically or asymmetrically located with respect to its ends. It is suggested that the student consider these cases *via* the media of Problems 1.5 and 1.7.

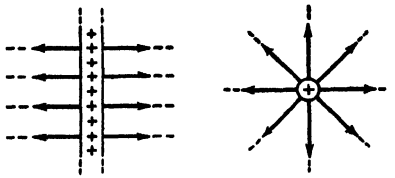


Fig. 1.6

1.7 Electric Field of a Large Charged Sheet

In preceding sections we have discussed the fields of a point charge and a long charged wire and have arrived at the result that they vary as r^{-2} and r^{-1}, respectively. On this basis one might guess that the field of a large charged sheet is independent of the distance from it (that is, it varies as r^0) and, indeed, by a consideration of lines of force, he can easily convince himself that such is the case.

As indicated in Figure 1.7, there is no basis for predicting that the lines of force emerging from a large charged sheet will diverge in *any* direction. Consequently, the field has a constant value regardless of distance from the sheet. This result will now be derived analytically by much the same method as was employed in the preceding section for the case of the long wire.

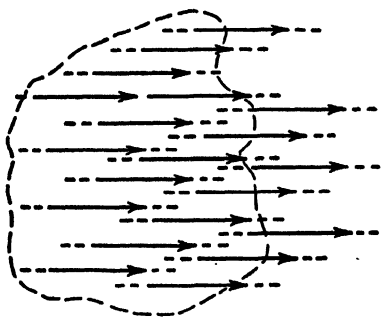


Fig. 1.7

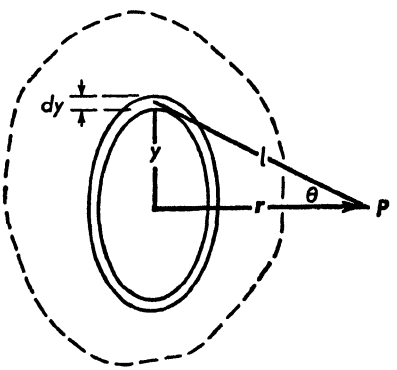


Fig. 1.8

Let us assume the density of the charge on the sheet to be σ per unit area. As suggested in Figure 1.8, a convenient element of the sheet to use, if one envisages subsequent integration with respect to an angle, is an annular ring of radius y and width dy. The total charge on this element is $\sigma 2\pi y\, dy$ units.

In calculating the field at P of this element of the sheet, one needs to reckon only with horizontal components, since the vertical component of the field of any portion of the element is canceled by that of its diametrically opposite counterpart. The field at P of this element is, therefore

$$dE = \frac{\sigma 2\pi y\, dy}{4\pi\epsilon l^2} \cos\theta \qquad [1\text{-}24]$$

which, when integrated, gives the total field at P of the entire sheet:

$$E = \frac{\sigma}{2\epsilon} \int \frac{\cos\theta\, y\, dy}{l^2} \qquad [1\text{-}25]$$

As in equations [1-21], the two variables y and l may be expressed in terms of the angle θ, which in this case, it will be noted, has the limits 0 and $\pi/2$. Equation [1-25] may now be written

$$E = \frac{\sigma}{2\epsilon} \int_0^{\pi/2} \sin\theta\, d\theta \qquad [1\text{-}26]$$

whence

$$E = \frac{\sigma}{2\epsilon} \qquad [1\text{-}27]$$

or

$$E = \frac{\sigma}{2\epsilon r} r \qquad [1\text{-}28]$$

As expected, this analytic result agrees with the earlier prediction that the field of a large charged sheet is independent of the distance from it.

Field Between Oppositely Charged Conducting Sheets

A common example of this configuration (Fig. 1.9) is the parallel-plate capacitor, which will be studied in Chapter 3. At this point we are concerned merely with the field strength between the two plates.

The charge on an isolated metal sheet resides equally on both surfaces of the sheet, and lines of force emerge indiscriminately from both front and back faces. If an equal and oppositely charged sheet is brought near however, as in Figure 1.9, the total charge on the first sheet is attracted to the side nearer the second.

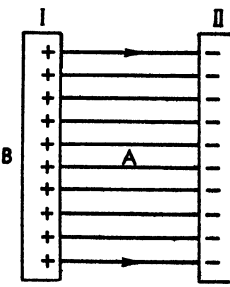


Fig. 1.9

The field in region B thus drops to zero while that in region A, between the two sheets, is twice its original value (equation [1-27]):

$$E = \frac{\sigma}{\epsilon} \qquad [1\text{-}29]$$

where σ is still, as it was in equation [1-27], the total charge density per unit cross-sectional area.

This same description applies if Plate II is grounded rather than given the opposite charge. The charge on Plate I then induces an equal and opposite charge on Plate II and, because of their mutual attraction, these two charges locate themselves on the inner surfaces of the two conducting sheets on which they reside.

To summarize, then, the entire field is confined to the space between the two oppositely charged sheets, where it has the value given by equation [1-29]. Elsewhere the field is zero. It should be added that if the two sheets are not infinitely large, these remarks do not apply to the portion of the field near the sheet edge. More will be said concerning this matter in § 3.7.

1.8 Gauss' Theorem

Karl Friedrich Gauss (1777–1855), director of the Göttingen Observatory, made outstanding contributions to astromomy, to mathematics, and to electricity and magnetism. The theorem that bears his name relates the surface integral of the electric field vector to the net charge enclosed within the surface, as follows:

$$\epsilon \int_s E_n \, ds = q \qquad [1\text{-}30]$$

Here E_n represents the normal component of the field at the position of the surface element ds. E_n is positive wherever $\boldsymbol{E}$ emerges from the surface and negative wherever $\boldsymbol{E}$ enters the surface. The symbol $\int_s$ refers to integration over *any* closed surface surrounding the charge q. The enclosed volume is either a vacuum ($\epsilon = \epsilon_0$) or is occupied by a homogeneous isotropic dielectric ($\epsilon = \epsilon$). Because of its generality this theorem is useful in the solution of many problems; in particular, the results of § 1.6 and § 1.7 can be readily deduced from it, as will be done in the following section.

Proof of Gauss' Theorem for a Spherical Surface Surrounding a Concentric Point Charge

Let us consider first, for simplicity, the case of a point charge q located at the center of a spherical surface, as shown in Figure 1.10. Here r is the radius of the sphere and ds is an element of the surface at the point P.

As we have seen, for the usual homogeneous, isotropic dielectrics, the magnitude of the electric field at the point P is given by

$$E = \frac{q}{4\pi\epsilon r^2} \qquad [1\text{-}18]$$

Gauss' theorem refers specifically to the normal component of the field. In this case the entire field is normal to the surface and points out from it, that is, $E_n = +E$. Thus, the product $\epsilon E_n\, ds$ is

$$\epsilon E_n\, ds = \frac{q\, ds}{4\pi r^2} \qquad [1\text{-}31]$$

which becomes, when integrated over the entire spherical surface,

$$\epsilon \int_s E_n\, ds = \frac{q}{4\pi r^2} \int_s ds \qquad [1\text{-}32]$$

Because the integral $\int_s ds$ refers to an integration over the surface, it is simply the surface area of the sphere, namely, $4\pi r^2$. Thus,

$$\epsilon \int_s E_n\, ds = \frac{q}{4\pi r^2} (4\pi r^2) = q \qquad [1\text{-}33]$$

in agreement with Gauss' theorem.

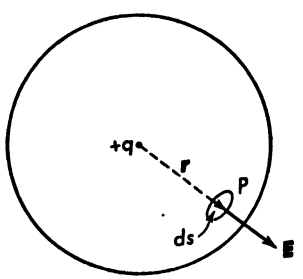


Fig. 1.10

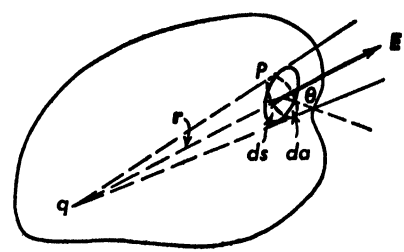


Fig. 1.11

Gauss' Theorem for Irregular Surfaces

When the charge is surrounded by an irregular surface, as in Figure 1.11, the proof of Gauss' theorem is somewhat more difficult.

The normal component of the field at the point P is $E \cos\theta$. The value of $\epsilon E_n\, ds$ is then

$$\epsilon E_n\, ds = \epsilon E\, ds \cos\theta \qquad [1\text{-}34]$$

But $ds \cos\theta$ is the projection of ds in the plane normal to $\boldsymbol{E}$, i.e., the plane normal to the radius vector $\boldsymbol{r}$. This projected area is shown in Figure 1.11 as da. Substituting $ds \cos\theta = da$ in equation [1-34], we have

$$\epsilon E_n \, ds = \epsilon E \, da$$

$$= \epsilon \left[\frac{q}{4\pi\epsilon r^2} \right] da \qquad [1\text{-}35]$$

and Gauss' surface integral becomes

$$\epsilon \int_s E_n \, ds = \frac{q}{4\pi} \int_s \frac{da}{r^2} \qquad [1\text{-}36]$$

Digression Concerning Solid Angles

As the expression under the integral sign in equation [1-36] is an element of *solid angle*, a few words of explanation concerning solid angles will be given for those who are not familiar with the concept.

The element of an *ordinary angle*, $d\theta$, may be written as

$$d\theta = \frac{dl}{r} \text{ radians} \qquad [1\text{-}37]$$

where dl is an element of arc of a circle of radius r. $d\theta$ is the angle subtended at the center of the circle by the arc dl. As both dl and r have dimensions of length, the angle itself is dimensionless. The *total* planar angle is

$$\theta = \oint d\theta = \frac{1}{r} \oint dl = \frac{2\pi r}{r} = 2\pi \text{ radians} \qquad [1\text{-}38]$$

The symbol $\oint$ indicates that the integration is carried around a closed path, in this case a circle.

In a similar manner an element of *solid angle*, $d\omega$, is defined as the angle subtended at the center of a sphere of radius r by an element, da, of the spherical surface (Fig. 1.12). Thus

$$d\omega = \frac{da}{r^2} \text{ steradians} \qquad [1\text{-}39]$$

Because of the presence of r^2 in the denominator, solid angles are also dimensionless. The *total* solid angle in space is found by integrating equation [1-39] over the surface of the sphere:

$$\omega = \int_s \frac{da}{r^2} = \frac{1}{r^2} \int_s da = \frac{4\pi r^2}{r^2} = 4\pi \text{ steradians} \qquad [1\text{-}40]$$

Although we have employed a spherical surface in order to evaluate the total solid angle, the result obtained is a general result.

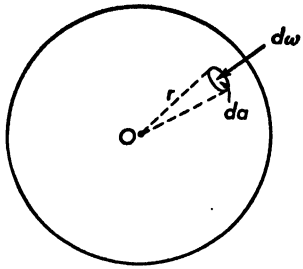


Fig. 1.12

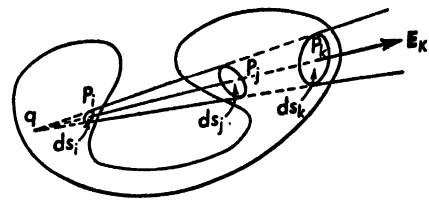


Fig. 1.13

Gauss' Theorem for Irregular Surfaces (continued)

Returning now to equation [1-36], we recognize that the integrand on the right is an element of solid angle, in which case

$$\epsilon \int_s E_n \, ds = \frac{q}{4\pi} \int_s d\omega = \frac{q}{4\pi} (4\pi) = q \qquad [1\text{-}41]$$

thereby establishing Gauss' theorem for an irregular surface of the type shown in Figure 1.11.

Let us also consider briefly the irregular surface shown in Figure 1.13, in which some of the field vectors cut the surface more than once, as at P_i, P_j and P_k. For each of the three surface elements shown, $\epsilon E_n \, ds$ has the form of equation [1-36]. The total value of $\epsilon E_n \, ds$ for the three elements is, therefore,

$$\frac{q}{4\pi} \left(\frac{da_i}{r_i^2} - \frac{da_j}{r_j^2} + \frac{da_k}{r_k^2} \right) \qquad [1\text{-}42]^9$$

where the negative sign preceding the second term in the parentheses arises from the fact that the vector $\boldsymbol{E}_j$, which passes through da_j, is directed *into* the surface, and is consequently entered in expression [1-42] as a negative quantity. Since da_i/r_i^2, da_j/r_j^2, and da_k/r_k^2 are identical solid angles, the last two terms in expression [1-42] cancel one another. The single remaining term corresponds exactly to equation [1-36]: when integrated over the entire surface it leads again to Gauss' theorem. Thus, Gauss' theorem is established for irregular surfaces of any shape, whether they involve the re-entry of field vectors or not.

As a corollary of this last result, we note that a charge located *outside* a closed surface gives rise, through the surface, to pairs of $\epsilon E_n \, ds$ contributions that cancel one another in the same manner as did those through ds_j and ds_k (Fig. 1.13). Hence, Gauss' theorem is proven for all closed surfaces, charge-containing or not.

[9] As in Figure 1.11, the da's are the projections of the ds's in planes normal to the radius vectors.

If several charges $q_1 \cdots q_i \cdots q_n$ are enclosed within the surface, Gauss' theorem is written as

$$\epsilon \int_s E_n \, ds = \sum_1^n q_i \qquad [1\text{-}43]$$

Finally, if there is a continuous distribution of charge within the enclosing surface,

$$\epsilon \int_s E_n \, ds = \int_v \rho \, dv \qquad [1\text{-}44]$$

where ρ is the charge density in the volume element dv and the symbol $\int_v$ indicates that the integrand is summed over the entire volume enclosed by the surface.

1.9 Some Applications of Gauss' Theorem

Field Strength at a Point External to a Spherical Charge Distribution

Let us employ Gauss' theorem to deduce the external field of a conducting spherical shell of radius a (Fig. 1.14), charged with q coulombs of electricity. As the shell is a conductor, the charge will distribute itself uniformly over it.

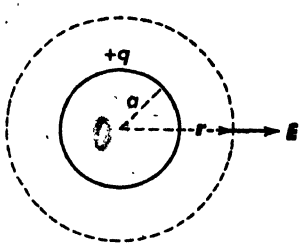


Fig. 1.14

In order to employ Gauss' theorem to calculate the field strength at a distance r from the center of the shell, let us draw a spherical Gaussian surface of radius r and concentric with the conducting shell. By *Gaussian surface* is meant one that is drawn for the purpose of applying Gauss' theorem. The surface is indicated in the diagram by the dotted lines.

Because of the spherical symmetry which exists in this case, the field vector emerges everywhere normal to this Gaussian surface. That is, the field and the normal component of the field are here synonymous. Also, the field has a constant value over the surface. By Gauss' theorem

$$\epsilon \int_s E_n \, ds = q \qquad [1\text{-}30]$$

or

$$\epsilon E \int_s ds = q \qquad [1\text{-}45]$$

or

$$\epsilon E(4\pi r^2) = q \qquad [1\text{-}46]$$

whence

$$E = \frac{q}{4\pi\epsilon r^2} \qquad [1\text{-}18]$$

which is identical with the expression already derived for the field at a distance r from a point charge (equation [1-18]). It is thus seen that the field strength at a point external to a charged conducting shell is the same as if the charge were concentrated at the center of the shell. From this important result, the general result may be deduced that *in calculating the field external to any spherical charge distribution, the charge may be considered as concentrated at the center.* This deduction forms the basis for Problem 1.10.

Field Inside a Charged Conductor

Let us consider first a charged conducting shell. It is proved in Problem 1.11 that the absence of charge within the shell results in the field in the interior of the shell being identically zero. Conversely, the presence of charge within the shell, if not compensated for, would give rise to an electric field.

Let us consider now the charge distribution on a *solid* conductor. The conductor can be pictured as a conducting shell enclosing a solid conducting core. We have seen above that the presence of net charge within the shell would give rise to a field in that region. The charges on a conductor, however, are free to move when subjected to an electric field, and any that were initially in the core would immediately be set in motion under the action of the internal field. The internal field is reduced to zero, and a static condition achieved, only when the charge in the interior has all migrated to the shell. For this reason the entire *free charge on a conductor locates itself on the surface*, regardless of the size or shape of the conductor.

It can be shown, although we shall not do so here, that the charge inside a spherical conductor is only zero if the electrical force between point charges varies inversely as the square of the distance between them. This theorem is due to Pierre Simon de Laplace (1749–1827), the son of a cottager in Normandy, who became an eminent applied mathematician and a Marquis of the Restoration. It provided the means, first employed by Cavendish and later by Clerk Maxwell (1831–1879), of verifying experimentally that the law of force is an inverse square one. If it is assumed that the force, instead of being inverse square, varies as $1/r^{2+x}$, Cavendish showed that $x \leq 0.02$, while Maxwell lowered the limit to $x \leq 0.000046$. Modern experiments have established that $x \leq 10^{-9}$.

Field of a Long Charged Wire

As a further illustration of the use of Gauss' theorem, we shall use it to obtain the strength of the field of a long charged wire. The desired result has already been derived, from first principles, in § 1.6.

The usefulness of Gauss' theorem depends to a large extent upon the shrewdness with which the Gaussian surfaces are drawn. As Gauss' theorem is concerned with the normal component of the field, the arithmetic is greatly simplified if the field at the various parts of a Gaussian surface has either a constant normal value or a zero normal component.

The Gaussian surface shown in Figure 1.15 satisfies this requirement. It is a cylinder, of length l and radius r, that is coaxial with the long wire. As the lines of force from the long wire are everywhere perpendicular to the axis of the wire (Fig. 1.6), none emerges through the ends, C, of the cylinder, while those passing through the sides, B, are normal to the surface. Everywhere on surface B (area $= 2\pi rl$) the normal component of the field has the value $E_n = E$. The charge enclosed by the Gaussian surface is λl, where λ is the charge per unit length. Gauss' theorem (equation [1-30]) takes, in this case, the form

$$\epsilon E(2\pi rl) = \lambda l \qquad [1\text{-}47]$$

whence

$$E = \frac{\lambda}{2\pi\epsilon r} \qquad [1\text{-}23]$$

in agreement with equation [1-23]. The simplicity of the above derivation, as compared to that in § 1.6, will be evident.

This analysis demonstrates that, for any cylindrical charge distribution, the field at an external point is identical with the field that would exist if the charge were concentrated along the axis of the cylinder.

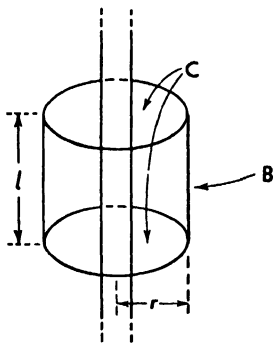


Fig. 1.15

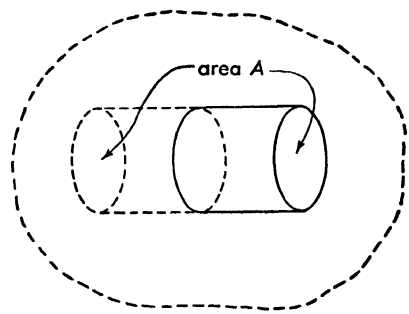


Fig. 1.16

Field of a Large Charged Sheet

This problem has been solved from first principles in § 1.7. It is handled much more easily by Gauss' theorem. A suitable Gaussian surface for this configuration is a right cylinder whose axis is perpendicular to the sheet and

whose ends lie above and below the sheet, respectively, as shown in Figure 1.16. Since the lines of force from the charged sheet are everywhere normal to the sheet (Fig. 1.7), they emerge normally through the ends of the Gaussian surface and not at all through the sides. $E_n = E$ has a constant value over the areas A. If the total surface charge density (both sides) is σ coul/m^2, Gauss' theorem,

$$\epsilon \int E_n \, ds = q \qquad [1\text{-}30]$$

takes, in this case, the form

$$\epsilon E 2A = \sigma A \qquad [1\text{-}48]$$

whence,

$$E = \frac{\sigma}{2\epsilon} \qquad [1\text{-}27]$$

in agreement with [1-27].

Field Close to a Charged Conductor

We wish to determine the field strength at a point close to the surface of a charged conductor. No matter how sharp the curvature of the surface of the conductor, provided the radius of curvature is large in comparison with the distance of the point from the surface, the surface can be considered an infinite plane. We can, therefore, assume that at points close to the conductor (in effect, *at* its surface) the electric field vector, $\boldsymbol{E}$, is normal to the surface. This is actually proved in the next chapter but, even here, it can be seen intuitively to be true.

In order to deduce the strength of the field at the surface of a conductor, let us use as a Gaussian surface the small pillbox shown in Figure 1.17. The top face of the box is just *outside* the surface of the conductor, and the bottom face is just *inside* the surface. The height of this pillbox is negligibly small. Hence, $\boldsymbol{E}$ is normal to the top face. Let us now apply Gauss' theorem to this surface.

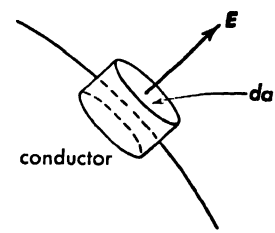


Fig. 1.17

Since the field inside a conductor is zero, $\epsilon E_n \, ds$ is zero everywhere on the bottom face of the pillbox and for the portions of the sides that lie within the conductor. $\epsilon E_n \, ds$ is also zero for the portions of the sides that lie outside the conductor, since $\boldsymbol{E}$ is parallel to them. The only non-zero $\epsilon E_n \, ds$-contributions are those associated with the top of the box. Thus,

$$\epsilon \int_s E_n \, ds = q \qquad [1\text{-}30]$$

becomes, in this case,

$$\epsilon E \, da = \sigma \, da \qquad [1\text{-}49]$$

where da is the area of the top of the box and σ is the superficial charge density on the conductor.

From equation [1-49] we conclude that the field very close to the surface of a conductor has the value

$$E = \frac{\sigma}{\epsilon} \qquad [1\text{-}50]$$

where ϵ is the dielectric constant of the medium surrounding the conductor and σ, as mentioned, is the charge per unit area on the conductor.

1.10 Electrical Stress on the Surface of a Charged Conductor

A charged conductor is always in a state of stress because of the forces of mutual repulsion between the individual charges.

Let us consider the electrical forces exerted upon an element of surface, ds, of an isolated charged conductor. The element under consideration is shown in Figure 1.18 as a thickened section. We shall assume the conductor to be located in free space.

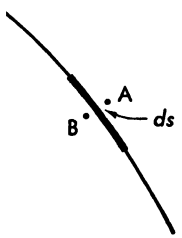


Fig. 1.18

At the point A, which is just outside the surface of the conductor, the field E consists of two components: E', caused by the charge $\sigma\, ds$ on the surface element, and E'', caused by the charge elsewhere on the conductor. Thus:

$$E = E' + E'' \qquad [1\text{-}51]$$

At the point B, which is very close to A but across the conductor surface from it, the field E', due to $\sigma\, ds$, is opposite in direction but equal in magnitude to that at point A. E'' is unchanged in value since A and B are infinitesimally close together. But B is located within the conductor, where the net field is zero, that is,

$$E'' - E' = 0 \qquad [1\text{-}52]$$

Adding equations [1-51] and [1-52] we find

$$E'' = \frac{E}{2} \qquad [1\text{-}53]$$

This is the field that acts upon the element of charge $\sigma\, ds$ on the surface of the conductor. The resulting force exerted upon the element is

$$dF = \frac{E}{2}\sigma\, ds \qquad [1\text{-}54]$$

or, per unit area,
$$\text{Stress} = \frac{dF}{ds} = \frac{E\sigma}{2} \qquad [1\text{-}55]$$

where E, as explicitly written in equation [1-51], is the field strength just outside the surface of the conductor. This same result is obtained if the conductor is assumed to be embedded in a medium of dielectric coefficient k_e. Under these circumstances, the internal values of E' and E'' are each larger than their external values by the factor k_e. Thus, equation [1-52] becomes $k_eE'' - k_eE' = 0$, which leads, as before, to equation [1-53].

We may now substitute in equation [1-55] the known value for the field close to a conductor ($E = \sigma/\epsilon$) to obtain the following equivalent forms of equation [1-55]:
$$\text{Stress} = \frac{\sigma^2}{2\epsilon} \qquad [1\text{-}56]$$

or
$$\text{Stress} = \frac{\epsilon E^2}{2} \qquad [1\text{-}57]$$

each of which expresses the electrical stress on the surface of a charged conductor.

1.11 Energy in an Electric Field

As the first step toward calculating the energy in an electric field, we shall calculate the pull on the plates of a parallel-plate capacitor.

In the case of an isolated conductor, the charges are distributed over its entire surface, with the result that the vector sum of all electric stresses on its surface is zero. The charges on the plates of a parallel-plate capacitor, however, are all on the inner surfaces and there give rise to an unbalanced electrical stress on each plate of amount $\epsilon E^2/2$, in accordance with equation [1-57]. The total electrical pull on a plate of area A is therefore:

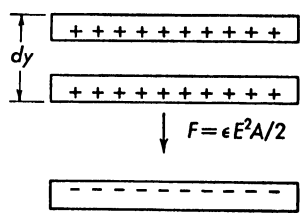


Fig. 1.19

$$\text{Force} = \frac{\epsilon E^2 A}{2} \qquad [1\text{-}58]$$

Let us now suppose that the upper plate of the capacitor in Figure 1.19 is raised a distance dy. The work done during this operation is
$$dW = \left[\frac{\epsilon E^2 A}{2}\right] dy \qquad [1\text{-}59]$$

The energy thus expended can be regarded as stored in the additional electric field thus created. As the volume of this additional electric field is $dv = A\,dy$, the work done in creating unit volume of field is:

$$\frac{dW}{dv} = \frac{\epsilon E^2}{2} \qquad [1\text{-}60]$$

This suggests that the

$$\textit{Volume density of electrical energy in region occupied by an electric field} = \frac{\epsilon E^2}{2}. \qquad [1\text{-}61]$$

It is easy to see how energy can be stored in an electric field if the field exists in a dielectric medium, because the rotation or creation of polar molecules represents the production of potential energy. Equation [1-61], however, implies that even empty space, when pervaded by an electric field, contains energy in the amount of $\epsilon_0 E^2/2$ joules/m^3. This suggests that space itself can be in a state of stress, a view that was mechanically acceptable as long as belief persisted in an all-pervading luminiferous ether. Whether or not it is valid to regard the medium surrounding a given charge configuration as the seat of all the electrical energy, it is without question true that any calculation based on such an assumption will not be incorrect on this account. In short, the mathematical idea expressed by equation [1-61] is correct, even though there may be doubt as to its physical interpretation.

Force on an Uncharged Conductor in an Electric Field

It is instructive to apply the above ideas to deduce the nature of the force experienced by an uncharged conductor in an electric field.

First, we recall that the space within a conductor is field-free. It can, therefore, contain no energy. Since the conductor is occupying space that would otherwise contain energy, its presence in the field results in a diminution in total energy. We next recall that all natural processes, if not interfered with, proceed inexorably from states of higher energy to states of lower energy. For this reason, the conductor will be urged toward the region in which the field is strongest since its presence there will result in removal from the field of the maximum possible amount of energy. Problem 1.18 deals with this subject.

Actually, the presence of a conductor distorts the field. Lines of force in the neighborhood of the conductor are deviated from their original paths so as to intercept the conductor at right angles (§ 2.8). Thus, the conductor not only removes energy from the field by virtue of its own volume but, in addition, weakens the field in the neighboring region. For this reason the calculation of

the magnitude of the force that is acting upon the conductor in the direction of increasing field is not as direct as the preceding paragraph might lead the reader to suppose.

1.12 Value of the Elementary Charge: The Oil Drop Experiment

The existence of the electron was unequivocally demonstrated in 1897 by J. J. Thomson (1856–1940), Professor of Experimental Physics in the Cavendish Laboratory, Cambridge. In this work cathode rays were deflected by electric and magnetic fields and shown to consist of streams of negatively charged particles with extraordinarily large magnitudes of *specific charge* (or charge per unit mass, usually written as $-e/m$). The accepted value for the specific charge of the electron is -1.759×10^{11} coulombs/kg.

Although Thomson and his colleagues also determined the magnitude, e, of the elementary charge, the first precision value for e was obtained by the American physicist, R. A. Millikan (1868–1953), who reported in 1917 the results of his celebrated *oil-drop experiment*, done at the University of Chicago.

In Figure 1.20, A and B are the plates of a parallel-plate capacitor. The

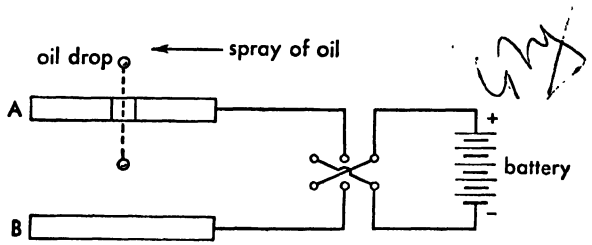


Fig. 1.20

upper plate A contains a number of small holes through which drops of oil, that have been sprayed into the region above the capacitor, may fall into the space between the capacitor plates. Many of these oil drops, which have been sprayed from an atomizer, are very small and have become charged in the atomizing process. Let us consider now a positively charged drop of mass m that has fallen to a point midway between the capacitor plates, at which time the switch is thrown to the left, thereby charging the top plate of the capacitor negatively and the bottom plate positively. This gives rise to an electric field E which exerts an upward force upon the drop, in opposition to the force of gravity. The electrical field is large enough for the electrical force on the drop to exceed the gravitational force. Thus, with the switch open, the drop falls under the action of gravity but, when the switch is closed, the drop rises. If the number of

electrons lost by the oil drop during the charging process is n, the net charge on the drop is $+ne$ coulombs and the electrical force exerted upon it is Ene newtons. The combined electric and gravitational forces acting on the drop with the switch open and closed are:

$$\text{Switch open} \qquad \text{Force down} = mg \tag{1-62}$$

$$\text{Switch closed} \qquad \text{Force up} = Ene - mg \tag{1-63}$$

These equations appear to neglect the buoyant force of the air. They are correct provided m is regarded as the "effective" mass of the drop—that is, the actual mass of the drop minus the mass of the air which it displaces.

An object moving in a viscous medium does not accelerate indefinitely but attains, instead, a certain *terminal velocity* which depends upon the nature of the object and of the medium. Fortunate it is that this is true; otherwise a stroll in the rain could be a fatal experience. The motion of objects in viscous fluids was studied by G. G. Stokes (1819–1903), a British physicist and mathematician, who showed that a sphere of radius r moving with velocity v through a medium whose coefficient of viscosity is η experiences a resisting force (*Stokes' law*)

$$F = 6\pi\eta rv \tag{1-64}$$

A sphere that is acted upon by an external force accelerates until such time as the resisting force, given by equation [1-64], equals the external force. The velocity possessed by the sphere at that time is the terminal velocity.

The terminal velocity of a small oil drop moving in air is low and is quickly attained. If we designate the terminal velocity during its rise as v_u, and during its fall as v_d, equations [1-62], [1-63], and [1-64] lead to the following relationships

$$6\pi\eta rv_u = Ene - mg \tag{1-65}$$

and

$$6\pi\eta rv_d = mg \tag{1-66}$$

whence after division and rearrangement,

$$ne = \frac{mg(v_u + v_d)}{Ev_d} \tag{1-67}$$

The upward and downward velocities, v_u and v_d, are observed experimentally by viewing the drop through a telescope and recording the time required for the drop to traverse a certain known distance. E is calculated from a knowledge of the capacitor spacing and the voltage applied across the plates (§ 2.4). The charge, ne, on the drop can therefore be computed if the effective mass, m, of the drop be known.

Stokes' law (equation [1-64]) is also used to obtain m for the drop. During its fall under the action of gravity the drop attains a terminal velocity given

by equation [1-66]. This, when written out to take into account the buoyant effect of the air, becomes

$$6\pi\eta r v_d = \tfrac{4}{3}\pi r^3 g(\rho - \sigma) \qquad [1\text{-}68]$$

where ρ is the density of the oil comprising the drop and σ is the density of the air.[10] Rearranging equation [1-68], one obtains

$$v_d = \frac{2gr^2(\rho - \sigma)}{9\eta} \qquad [1\text{-}69]$$

Thus, the radius of the drop can be calculated from the observed downward terminal velocity and from it the effective mass, $m = 4\pi r^3(\rho - \sigma)/3$. It is, of course, necessary to know to a high degree of precision the density, ρ, of the oil and the viscosity, η, of the air. Incidentally, Millikan discovered that it is necessary to modify Stokes' law when very small drops are employed, since for such drops the air can no longer be regarded as a homogeneous medium.

In practice, the rising and falling velocities of a given drop are observed for an extended period of time. The falling velocity v_d is always the same. Sooner or later the charge on the drop changes, as a result of encountering an ion created by cosmic radiation or x-rays (Millikan utilized an x-ray beam for this purpose), and the rising velocity v_u displays a new value. Let us suppose that the charge and upward velocity are $n'e$ and v_u', respectively, before the change in charge, and $n''e$ and v_u'' afterward. Then, from equation [1-67],

$$(n' - n'')e = \frac{mg(v_u' - v_u'')}{Ev_d} \qquad [1\text{-}70]$$

The advantage of using this expression to determine e rather than using expression [1-67] becomes evident when one reflects that, whereas the total charge ne on the drop may consist of a large (uncertain) number of elementary charges, the amount of the change in charge, $(n' - n'')e$, is generally only one or two elementary charges. As one observes the drop through several changes in charge, results such as those below are obtained.

Change in charge $(n' - n'')e$	*No. of elementary charges apparently involved* $(n' - n'')$	*Value of single elementary charge,* e
3.22×10^{-19} coulombs	2	1.61×10^{-19} coulombs
1.61×10^{-19}	1	1.61×10^{-19}
4.77×10^{-19}	3	1.59×10^{-19}
1.60×10^{-19}	1	1.60×10^{-19}
		Mean 1.602×10^{-19} coulombs

[10] In this section ρ and σ represent ordinary densities, and not charge densities.

The value of the elementary charge is easily recognized since it corresponds to the smallest charge change that is observed.

The value found by Millikan was $e = 1.59 \times 10^{-19}$ coulombs, which is to be compared to the value now accepted of $e = 1.60201 \pm 0.00003 \times 10^{-19}$ coulombs. The error in Millikan's work arose from an incorrect value for the viscosity of air. The reason for the discrepancy between Millikan's value and that obtained subsequently by other methods was for many years a mystery and even a cause of contention.

1.13 Problems *

E (1.1) The radius of the innermost (ground state) electronic orbit in the hydrogen atom is 0.53×10^{-10} m (0.53 Å). Calculate the velocity of the electron. [Ans.: 2.2×10^6 m/sec.]

E (1.2) Two pith balls of equal mass m and equal charge q are suspended from the same point by strings of equal length l. Show that each string makes an angle θ with the vertical given by $16\pi\epsilon_0 mgl^2 \sin^3\theta = q^2 \cos\theta$.

E (1.3) How many electrons flow per millisecond from the cathode to the plate of a diode if the current is 10 microamperes? [Ans.: 6.25×10^{10}.]

M (1.4) The volume charge density in a dielectric sphere of radius a meters is $\rho = 6 + 2r^2$ coul/m^3, where r is the distance from the center. What is the *average* value of the charge density?

[Ans.: $6 + 1.2a^2$ coul/m.^{3}]

E (1.5) A thin wire of length $2r$ is uniformly charged with Q coulombs. Show that the field strength at a distance r meters from the center of the wire is $Q/(4\sqrt{2}\pi\epsilon r^2)$ n/coul.

E (1.6) A circular ring of radius a meters has a uniform charge density of λ coul/m. Calculate the field strength along the axis of the ring at a distance d meters from the center. [Ans.: $\lambda ad/2\epsilon(a^2 + d^2)^{3/2}$.]

M (1.7) A conducting rod of length l meters is given a total charge of Q coulombs. Find the magnitude and direction of the force exerted upon a charge of +1 microcoulomb located at the point P, as shown in Figure 1.21.

[Ans.: $(Q \times 10^{-6})/(5.2\pi\epsilon l^2)$ newtons, making an angle of 22.5° with the horizontal.]

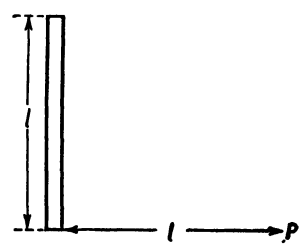


Fig. 1.21

* E = easy, M = medium hard, H = hard.

E (1.8) A charge of 5 microcoulombs is located at the origin of coordinates and one of -3 microcoulombs at the point (10,0). Calculate the field strength at the point (0,30). All coordinate distances are in meters. [Ans.: 26 newtons/coul]

E (1.9) For the charge configuration of Problem 1.8, at what point in the X-Y plane would a third charge of $+2$ coulombs be subject to no force? Is this a position of stable or unstable equilibrium? [Ans.: (44.4,0), unstable.]

E (1.10) Assume, as proved in § 1.9, that to the external observer the charge on a conducting spherical shell may be considered as located at the center of the sphere. Show that this same result holds for any spherical charge distribution, that is, one in which the charge density, $\rho = \rho(r)$, is a function of the radius only.

E (1.11) Prove, by the use of Gauss' theorem, that the electric field is zero inside a charged spherical conducting shell.

E (1.12) Calculate the field at a distance r meters from the surface of a charged conducting sphere of radius a meters, whose superficial charge density is σ coul/m^2. [Ans.: $E = a^2\sigma/\epsilon(a + r)^2$.]

M (1.13) In Figure 1.22 the shaded portion represents a hollow conducting sphere with zero total charge. Situated within the sphere and isolated from it is a body with total charge $+q$ coulombs. Show (a) that the total charge residing on the inner surface, A, of the hollow sphere is $-q$ and (b) that the total charge on the outer surface, B, is $+q$ coulombs.

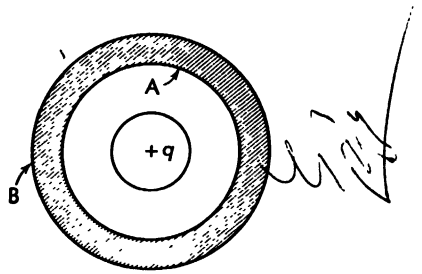


Fig. 1.22

M (1.14) The superficial charge density on a dielectric shell of radius a meters is given by $\sigma = k \sin \phi$, where ϕ is the angle between the radius vector and the Z-axis, and k is a constant. Calculate the total charge upon the sphere. [Ans.: $k\pi^2a^2$.]

M (1.15) In a solid dielectric sphere the volume charge density, ρ, is directly proportional to the distance from the center. If $\rho = \rho_0$ at the surface of the sphere, calculate the field strength at a distance r from the center of the sphere for the two cases $r > a$ and $r < a$, where a is the radius of the sphere. [Ans.: $\rho_0 a^3/4\epsilon r^2$, $\rho_0 r^2/4\epsilon a$.]

E (1.16) Assume that the electric field at the surface of a charged conducting shell is E. If a small hole is made in the wall of the shell, show that the electric field at the mouth of the hole is $E/2$.

M (1.17) An insulated soap bubble 0.1 m in radius is charged with 7×10^{-9} coulombs of electricity. Taking the atmospheric pressure as 10^5 newtons/m^2, find the increase in radius due to charging.

[Ans.: Approx. 6×10^{-11} m.]

E (1.18) What is the force on an uncharged conductor of volume v in a uniform electric field of strength E?

Chapter 2

Electric Potential

In this chapter we shall develop the notion of electric potential and show its relation to electric field. First, however, some useful vector terminology will be presented or recalled, as the case may be.

2.1 Some Elementary Vector Analysis

Unit Vectors

Vectors are expressed in Cartesian coordinates in terms of three *unit vectors* $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$, which are directed, respectively, along the X-, Y-, and Z-axes, as shown in Figure 2.1. The vector $\boldsymbol{F}$, with X-, Y-, and Z-components F_x, F_y and F_z, respectively, is written

$$\boldsymbol{F} = \boldsymbol{i}F_x + \boldsymbol{j}F_y + \boldsymbol{k}F_z \qquad [2\text{-}1]$$

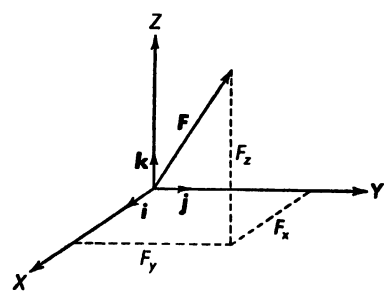


Fig. 2.1

Unit vectors are thus seen to serve the purpose of specifying *direction* only. Because they are *unit* vectors, any quantity multiplied by one of them is not changed in magnitude.

Multiplication of Vectors

Special rules apply for the multiplication of vectors; in fact, there are two forms of vector multiplication, one leading to a scalar and the other to a vector

product. These will now be illustrated by reference to two common examples in mechanics.

The *scalar* or "dot" product of two vectors $\boldsymbol{F}$ and $\boldsymbol{r}$, written as $\boldsymbol{F}\cdot\boldsymbol{r}$, is a scalar quantity of magnitude $Fr\cos\theta$, where θ is the angle between $\boldsymbol{F}$ and $\boldsymbol{r}$. It is useful, as suggested in Figure 2.2, in the calculation of *work*.

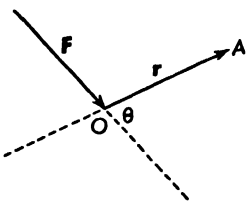


Fig. 2.2

The amount of work done when a force acts through a certain distance is defined as the product of the distance moved and the component of the force in the direction of motion. If the force $\boldsymbol{F}$ be exerted on, say, a lawnmower, moving it up a slight incline from position O (the origin) to position A, the work done in this operation is

$$\text{Work} = (F\cos\theta)(r) \qquad [2\text{-}2]$$

which can conveniently be written in terms of the scalar product, viz.,

$$\text{Work} = \boldsymbol{F}\cdot\boldsymbol{r} \qquad [2\text{-}3]$$

If $\boldsymbol{F}$ and $\boldsymbol{r}$ be written out in terms of unit vectors, equation [2-3] becomes

$$\text{Work} = (\boldsymbol{i}F_x + \boldsymbol{j}F_y + \boldsymbol{k}F_z)\cdot(\boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z) \qquad [2\text{-}4]$$

$$= \boldsymbol{i}\cdot\boldsymbol{i}F_x x + \boldsymbol{i}\cdot\boldsymbol{j}F_x y + \boldsymbol{i}\cdot\boldsymbol{k}F_x z + \boldsymbol{j}\cdot\boldsymbol{i}F_y x + \boldsymbol{j}\cdot\boldsymbol{j}F_y y + \boldsymbol{j}\cdot\boldsymbol{k}F_y z$$

$$+ \boldsymbol{k}\cdot\boldsymbol{i}F_z x + \boldsymbol{k}\cdot\boldsymbol{j}F_z y + \boldsymbol{k}\cdot\boldsymbol{k}F_z z \qquad [2\text{-}5]$$

Clearly, $\boldsymbol{i}\cdot\boldsymbol{i} = \boldsymbol{j}\cdot\boldsymbol{j} = \boldsymbol{k}\cdot\boldsymbol{k} = 1$, since the cosine of the angle between each unit vector and itself is unity, whereas $\boldsymbol{i}\cdot\boldsymbol{j} = \boldsymbol{i}\cdot\boldsymbol{k} = \boldsymbol{j}\cdot\boldsymbol{i} = \boldsymbol{j}\cdot\boldsymbol{k} = \boldsymbol{k}\cdot\boldsymbol{i} = \boldsymbol{k}\cdot\boldsymbol{j} = 0$, since the angle whose cosine is involved in these cases is 90°. Equation [2-5] may then be rewritten

$$\text{Work} = F_x x + F_y y + F_z z \qquad [2\text{-}6]$$

This multiplied-out example illustrates three important properties of a scalar product: (a) *the order of the factors in a scalar product is immaterial* ($\boldsymbol{F}\cdot\boldsymbol{r} = \boldsymbol{r}\cdot\boldsymbol{F}$), (b) *the scalar product of two perpendicular vectors is zero*, and (c) *the scalar product of two nonperpendicular vectors is a scalar.*

The *vector* or "cross" product of $\boldsymbol{r}$ and $\boldsymbol{F}$, written as $\boldsymbol{r}\times\boldsymbol{F}$, is a vector of magnitude $rF\sin\theta$ pointing in the direction a right-handed screw would advance in rotating (less than π radians) from $\boldsymbol{r}$ to $\boldsymbol{F}$. This direction is obviously

normal to both $\boldsymbol{r}$ and $\boldsymbol{F}$. *Torque* or the *moment of force* provides an example of such vector multiplication.

Consider a force $\boldsymbol{F}$ applied to the end of a rod which is pivoted at O, as in Figure 2.3. The torque M exerted on the rod is given by the product of the force and the perpendicular distance from O to the line of action of the force. Thus,

$$M = rF \sin \theta \qquad [2\text{-}7]$$

a quantity which has the same magnitude as the vector product

$$\boldsymbol{M} = \boldsymbol{r} \times \boldsymbol{F} \qquad [2\text{-}8]$$

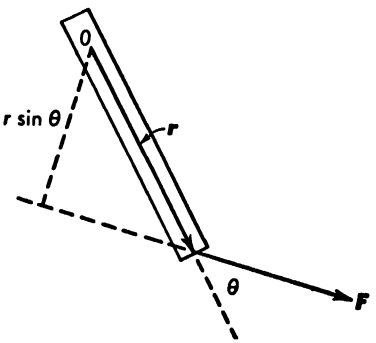


Fig. 2.3

Here $\boldsymbol{M}$ is directed out from the paper, coinciding in direction with the axis about which the rod rotates in response to the applied torque. The suitability of this direction is evident when one considers rotation under a constant torque $\boldsymbol{M}$. Although the direction of action of the force giving rise to the torque may change during the rotation, the axis of rotation remains constant and is therefore an appropriate direction to assign to $\boldsymbol{M}$. Also, with this convention for $\boldsymbol{M}$, two forces which produce the same rotating effects result in torques which are not only equal in magnitude but are, in addition, vectorially identical.

In terms of unit vectors equation [2-8] becomes

$$\boldsymbol{M} = (\boldsymbol{i}x + \boldsymbol{j}y + \boldsymbol{k}z) \times (\boldsymbol{i}F_x + \boldsymbol{j}F_y + \boldsymbol{k}F_z) \qquad [2\text{-}9]$$

$$= \boldsymbol{i} \times \boldsymbol{i}F_x x + \boldsymbol{i} \times \boldsymbol{j}F_y x + \boldsymbol{i} \times \boldsymbol{k}F_z x + \boldsymbol{j} \times \boldsymbol{i}F_x y + \boldsymbol{j} \times \boldsymbol{j}F_y y + \boldsymbol{j} \times \boldsymbol{k}F_z y$$

$$+ \boldsymbol{k} \times \boldsymbol{i}F_x z + \boldsymbol{k} \times \boldsymbol{j}F_y z + \boldsymbol{k} \times \boldsymbol{k}F_z z \qquad [2\text{-}10]$$

In this case, it is readily seen that $\boldsymbol{i} \times \boldsymbol{i} = \boldsymbol{j} \times \boldsymbol{j} = \boldsymbol{k} \times \boldsymbol{k} = 0$, whereas $\boldsymbol{i} \times \boldsymbol{j} = -\boldsymbol{j} \times \boldsymbol{i} = \boldsymbol{k}$, $\boldsymbol{j} \times \boldsymbol{k} = -\boldsymbol{k} \times \boldsymbol{j} = \boldsymbol{i}$ and $\boldsymbol{k} \times \boldsymbol{i} = -\boldsymbol{i} \times \boldsymbol{k} = \boldsymbol{j}$. Thus, equation [2-10] for the moment of a force may be rewritten

$$\boldsymbol{M} = \boldsymbol{i}(yF_z - zF_y) + \boldsymbol{j}(zF_x - xF_z) + \boldsymbol{k}(xF_y - yF_x) \qquad [2\text{-}11]$$

Three important properties of a vector product should be noted from this example: (a) *the order of the factors in a vector product is important* ($\boldsymbol{r} \times \boldsymbol{F} = -\boldsymbol{F} \times \boldsymbol{r}$), (b) *the vector product of two parallel vectors vanishes*, and (c) *the vector product of two nonparallel vectors is a third vector that is normal to each of the other two.*

Vector Differential Operators

The vector differential operator, ∇, known as *del*, is defined as

$$\nabla = i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \qquad [2\text{-}12]$$

It may operate directly upon a scalar quantity, V, viz.,

$$\nabla V = i\frac{\partial V}{\partial x} + j\frac{\partial V}{\partial y} + k\frac{\partial V}{\partial z} \qquad [2\text{-}13]$$

to give a vector quantity, known as the *gradient* of the scalar quantity. As the name gradient suggests, the three components of ∇V are the rates (say, per meter) at which V changes in each of the three coordinate directions.

Also, ∇ may operate upon a vector through either a dot or cross product. Thus:

$$\nabla \cdot \boldsymbol{E} = \left(i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}\right)\cdot\left(iE_x + jE_y + kE_z\right)$$

$$= \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \qquad [2\text{-}14]$$

which is known as the *divergence* of $\boldsymbol{E}$, or

$$\nabla \times \boldsymbol{E} = \left(i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}\right) \times (iE_x + jE_y + kE_z)$$

$$= i\left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right) + j\left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\right) + k\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right) \qquad [2\text{-}15]$$

which is known as the *curl* of $\boldsymbol{E}$.[1]

2.2 Gauss' Theorem in Vector Form

Gauss' theorem (§ 1.8) deals with the surface integral of the normal component of the field vector $\boldsymbol{E}$.

In order to express this theorem in vector form we can represent each element of the surface by a vector that is directed outward from the surface, as

[1] The terms "divergence" and "curl" arose in hydrodynamics. Let v equal the velocity at some point, P, in a moving fluid. The divergence of v is the rate per unit volume at which an infinitesimal portion of the fluid at P is changing volume. The curl of v is equal to twice the angular velocity of an infinitesimal portion of the fluid at P. The significance of the divergence and curl of the electrostatic field, $\boldsymbol{E}$, is commented upon in connection with equations [2-19] and [2-31].

shown in Figure 2.4. The quantity $\epsilon E_n\, ds$, which appears in Gauss' theorem, can then be written as

$$\epsilon \boldsymbol{E} \cdot d\boldsymbol{s}$$

Gauss' theorem for a continuous charge distribution (equation [1-44]) now takes the form

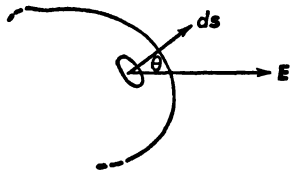


Fig. 2.4

$$\int_s \boldsymbol{E} \cdot d\boldsymbol{s} = \frac{1}{\epsilon} \int_v \rho \, dv \qquad [2\text{-}16]$$

Maxwell's First Equation

One of the basic theorems of vector calculus is the *divergence theorem*, relating the surface integral of the normal component of a vector to the volume integral of its divergence (see Appendix). This theorem is written as follows:

$$\int_s \boldsymbol{E} \cdot d\boldsymbol{s} = \int_v \nabla \cdot \boldsymbol{E} \, dv \qquad [2\text{-}17]$$

where it is understood that the surface in question bounds the volume under consideration.

By means of the divergence theorem it is often possible to transform a surface to a volume integral, or vice versa. Evidently, by it, equation [2-16] can be rewritten in the form

$$\int_v \nabla \cdot \boldsymbol{E} \, dv = \frac{1}{\epsilon} \int_v \rho \, dv \qquad [2\text{-}18]$$

This equality requires that the integrands be equal, that is,

$$\nabla \cdot \boldsymbol{E} = \frac{\rho}{\epsilon} \qquad [2\text{-}19]$$

This equation is one of *Maxwell's equations*. It is the first of the four fundamental equations which were used by Clerk Maxwell to develop the electromagnetic theory of light. Attention will be drawn to the other three equations as they are developed, and the four will be discussed as a group in Chapter 14.[8]

[8] We see from equation [2-19] that the divergence of the electric field is only zero at a point if an infinitesimal volume surrounding the point contains zero charge. A positive value for the divergence of $\boldsymbol{E}$ indicates that lines of force "diverge" from the infinitesimal volume under consideration, while a negative value for $\nabla \cdot \boldsymbol{E}$ indicates that lines of force "converge" upon it.

2.3 General Concept of Potential and Difference in Potential

If one raises an object of mass m kg from the surface of the earth to a point P, which is at a height of h m, he does mgh joules of work, and the object is said to possess mgh joules of *potential energy*. The energy is "potential" by virtue of the fact that it may subsequently be used to perform mgh units of work. If the object has *unit mass*, the work done (gh joules) is said to be the difference in *gravitational potential* between the surface of the earth and the point P. Although one always deals with *differences* in gravitational potential, by common consent the surface of the earth is taken to be the zero position. For this reason the statement that "the difference in gravitational potential between a point P and the surface of the earth is gh joules/kg" may be rewritten as "the gravitational potential at point P is gh joules/kg."

The analogy between *electric potential* and gravitational potential is close. In the electric case *the difference in potential between any two points is defined as the work per unit positive charge that is required to move a small test charge from the one point to the other.* This work may be positive or negative, depending upon whether the electric field in the region opposes or encourages the motion.

As in the case of gravity, it is convenient to agree upon a "zero" of potential. In the ideal case, *infinity is taken to be the zero of potential.* The *absolute potential at any* other *point is* then *the work per unit charge required to move the test charge from infinity to the point in question.* In the practical case, infinity is not an accessible reference point. What is readily accessible is the surface of the earth. Fortunately, the earth is a conductor and so provides a vast equipotential surface. Although its absolute potential (i.e. with respect to infinity) is not accurately known, *the earth* is the reference potential that is used in practical cases and, for simplicity, is *said to be zero.* The student will discover that these two usages, although inconsistent, are not confusing. Whenever the earth or *ground* appears as an element in a problem, it provides the zero of potential. If ground does not appear, and unless the zero of potential is explicitly defined in some other manner, then zero is taken to be at infinity.

The symbol for electric potential is V.

2.4 Relation Between Potential and Field

Let us consider the work done in slowly moving a test charge, $+q$, a vector distance $d\mathbf{r}$ in a region where the electric intensity is $\mathbf{E}$, as shown in Figure 2.5.

By the definition of $\mathbf{E}$, the force on the test charge is given by $\mathbf{E}q$. To over-

come this force an external force, $-Eq$, equal and opposite to the force exerted by the field, must be applied. The work done in moving the test charge the distance $d\boldsymbol{r}$ $(=\boldsymbol{i}\,dx + \boldsymbol{j}\,dy + \boldsymbol{k}\,dz)$ is then:

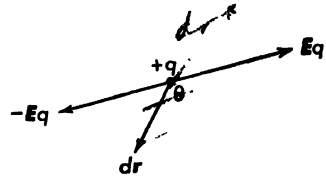


Fig. 2.5

$$dW = (-q\boldsymbol{E})\cdot d\boldsymbol{r} \qquad [2\text{-}20]$$

From this the potential difference dV (work per unit positive charge), between the tail and tip of $d\boldsymbol{r}$ is

$$dV = -\boldsymbol{E}\cdot d\boldsymbol{r} \qquad [2\text{-}21]$$

$$= -(E_x\,dx + E_y\,dy + E_z\,dz) \qquad [2\text{-}22]$$

The difference in potential between two points $A(x_1,y_1,z_1)$ and $B(x_2,y_2,z_2)$ is the integral of either [2-21] or [2-22] between the points in question. Thus:

$$V_B - V_A = \int_A^B dV = -\int_A^B \boldsymbol{E}\cdot d\boldsymbol{r} \qquad [2\text{-}23]$$

or

$$V_B - V_A = -\left[\int_{x_1}^{x_2} E_x\,dx + \int_{y_1}^{y_2} E_y\,dy + \int_{z_1}^{z_2} E_z\,dz\right] \qquad [2\text{-}24]$$

If $V_B - V_A$ is positive, point B is said to be at the higher potential.

The mks unit of potential, the joule/coulomb, is known as the *volt*. Thus, *two points differ in potential by 1 volt if energy in the amount of 1 joule per coulomb is expended in moving a test charge from the one point to the other.*

Potential Is a Function Only of Position

Because potential has the character of work, it is a scalar quantity. In the calculation of the potential difference between two points, as can be seen in equation [2-24], there is no specification that a particular route must be followed in moving the test charge from the one point to the other. From this we are led to suspect that the path of integration of $\boldsymbol{E}\cdot d\boldsymbol{r}$ (in equation [2-23]) is immaterial, and that the value of $V_B - V_A$ depends only upon the coordinates of the two end points.

We may easily convince ourselves of the irrelevance of the path of integration by employing a conservation-of-energy argument. In applying the argument, consideration need only be given to the work done in moving the test charge and the potential energy resulting therefrom. Kinetic energy does not enter the picture (since all motions are slow). Also, it is assumed that there is no dissipation of energy under these conditions. Any energy that is expended in moving the test charge along a certain path can be subsequently regained by

allowing the charge to reverse its direction and retrace its path. The argument now follows.

Let us suppose that the value of $-\int_A^B \boldsymbol{E}\cdot d\boldsymbol{r}$ *does* depend upon the path of integration and that (Fig. 2.6)

$$(V_B - V_A)_{\text{path } ACB} > (V_B - V_A)_{\text{path } ADC} \qquad [2\text{-}25]$$

Let us now move the test charge from A to B along the lower path and then allow it to return to A along the upper path. By so doing, we gain more energy during the return trip than was expended in moving the charge from A to B in the first place. Although the charge is back in its original position, and the rest of the system is in no way altered, we have gained energy—in contradiction to the law of conservation of energy. Thus, we conclude that the value of $V_B - V_A$ is independent of the path of integration of dV (i.e., $\boldsymbol{E}\cdot d\boldsymbol{r}$). This is equivalent to saying, in mathematical language, that dV is an *exact differential.*

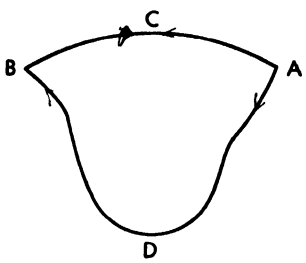


Fig. 2.6

Potential is a single-valued function of position. The integral of $\boldsymbol{E}\cdot d\boldsymbol{r}$ around any closed path is therefore zero, since it represents (equation [2-23]) the difference in potential between a given point and itself. This fundamental property of electrostatic fields is expressed as follows:

$$\oint \boldsymbol{E}\cdot d\boldsymbol{l} = 0 \qquad [2\text{-}26]$$

where $d\boldsymbol{l}$, corresponding to the $d\boldsymbol{r}$ used heretofore, is an infinitesimal vector that is tangent to the path in question. It is emphasized that equation [2-26] is true only of *static* fields. As we shall later see, the value of $\oint \boldsymbol{E}\cdot d\boldsymbol{l}$ can be significantly different from zero if the system contains sources of energy.

A field, such as the electrostatic field $\boldsymbol{E}$, whose line integral is independent of the path of integration, is known as a *conservative field.* Equation [2-26], which states that the line integral of $\boldsymbol{E}$ vanishes around any closed path, may be regarded as an alternative expression of the same property.

Electric Field Is the Gradient of the Potential

Since the potential is a function only of positional coordinates, its differential dV may be written in the following general form:

$$dV = \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz \qquad [2\text{-}27]$$

Comparing equations [2-22] and [2-27], one finds the potential and the components of the field to be related as follows:

$$E_x = -\frac{\partial V}{\partial x}, \quad Ey = -\frac{\partial V}{\partial y}, \quad E_z = -\frac{\partial V}{\partial z} \qquad [2\text{-}28]$$

Relationships analogous to these also exist for other coordinate systems. Problem 2.4 is concerned with the spherical polar equivalents.

Finally, if we take the gradient of the potential,

$$\nabla V = \boldsymbol{i}\frac{\partial V}{\partial x} + \boldsymbol{j}\frac{\partial V}{\partial y} + \boldsymbol{k}\frac{\partial V}{\partial z}$$

we find, by substitution of equation [2-28] in the above, that

$$\nabla V = -\boldsymbol{i}E_x - \boldsymbol{j}E_y - \boldsymbol{k}E_z$$

or

$$\boldsymbol{E} = -\nabla V \qquad [2\text{-}29]$$

We have so far expressed the electric field in the mks system, in units of newtons/coulomb. In keeping with the relationship between $\boldsymbol{E}$ and V that we have established in equation [2-29], that the field is the gradient of the potential, the field may also be expressed in volts/meter. The student should satisfy himself that the newton/coulomb and the volt/meter are equivalent. The volt/meter will be used hereafter. For example, in the particular case of the parallel-plate capacitor (§ 1.7), the field is constant and its value is given by

$$E = \frac{V}{d} \text{ volts/m} \qquad [2\text{-}30]$$

where V is the potential difference (in volts) between the plates and d is the separation (in meters) of the plates.

The result expressed in equation [2-29], that *the electrostatic field is the (negative) gradient of the potential*, is of paramount importance. Incidentally, it typifies an additional property of a conservative field: that it can be derived from a scalar potential.

Presumably the reader has already demonstrated to his own satisfaction, as required by Problem 2.3, that $\nabla \times \nabla V = 0$. If so, he may now conclude, on the basis of equation [2-29], that the electrostatic field is governed by the following relation:

$$\nabla \times \boldsymbol{E} = 0 \qquad [2\text{-}31]$$

or, in words, that the curl of the electrostatic field is zero.[3] This criterion is probably the most useful shibboleth for identifying a conservative field.

[3] Equation [2-31], when taken in conjunction with equation [2-26], provides a physical interpretation of the curl of $\boldsymbol{E}$: when $\nabla \times \boldsymbol{E}$ is zero, the work done in moving a charge around any closed path is zero.

Electric Field Is Normal to Equipotential Surfaces

Equation [2-29] permits us to deduce that the electric field is perpendicular to any surface V = constant. First, we recall (equations [2-21] and [2-29]) that the difference in potential between the two ends of the vector $d\boldsymbol{r}$ is given by

$$dV = -\boldsymbol{E}\cdot d\boldsymbol{r} = \nabla V\cdot d\boldsymbol{r} \qquad [2\text{-}32]$$

Next, let us suppose that $d\boldsymbol{r}$ lies entirely in a surface for which the potential has a constant value. If so, dV is zero and equation [2-32] becomes

$$\nabla V\cdot d\boldsymbol{r} = 0 \qquad [2\text{-}33]$$

The gradient of the potential is not zero (except in the special case when the entire region is at the same potential). ∇V and $d\boldsymbol{r}$ must, therefore, be normal to one another. Since $d\boldsymbol{r}$ lies in an equipotential surface, the gradient of V must be normal to this surface. Thus we conclude that the electric field ($= -\nabla V$) is everywhere normal to the equipotential surfaces.

Surfaces of Conductors Are Equipotential Surfaces

When two points are at different potentials, a field, $\boldsymbol{E} = -\nabla V$, exists in the region between them. If the two points are located on the surface of a conductor, charges are free to move under the action of the field and do so in such a direction as to destroy the difference in potential. Thus, unless there is provision for *maintaining* a difference in potential between two points on a conductor, by use of a battery, for example, any difference that initially exists will swiftly disappear. In short, for the static cases that we are concerned with here, *conducting surfaces are equipotential surfaces.*

Poisson's Equation and Laplace's Equation

We may now combine the two equations

$$\nabla\cdot\boldsymbol{E} = \frac{\rho}{\epsilon} \qquad [2\text{-}19]$$

and

$$\boldsymbol{E} = -\nabla V \qquad [2\text{-}29]$$

to obtain

$$\nabla\cdot\nabla V = -\frac{\rho}{\epsilon}$$

or

$$\nabla^2 V = -\frac{\rho}{\epsilon} \qquad [2\text{-}34]$$

Equation [2-34] is *Poisson's equation*. It is readily seen, by writing out the product $\nabla \cdot \nabla$, that another form of Poisson's equation is

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -\frac{\rho}{\epsilon} \qquad [2\text{-}35]$$

In regions where there is no charge, $\rho = 0$, and Poisson's equation reduces to

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0 \qquad [2\text{-}36]$$

This is *Laplace's equation*, one of the most important relationships in mathematical physics. The operator ∇^2 is known as the *Laplacian*. The use of Laplace's equation in the solution of certain electrostatic problems is described in § 2.12.

2.5 Potential Arising from a Point Charge and from a System of Point Charges

Let us now, by reference to Figure 2.7, calculate the potential at a distance r m from a point charge of $+q$ coulombs.

As we have seen in the preceding section, the difference in potential between the two ends of the infinitesimal vector $d\boldsymbol{r}$ is

$$dV = -\boldsymbol{E} \cdot d\boldsymbol{r} \qquad [2\text{-}21]$$

The field strength at a vector distance $\boldsymbol{r}$ from the point charge q we have already written as

$$\boldsymbol{E} = \frac{q}{4\pi\epsilon r^3}\boldsymbol{r} \qquad [1\text{-}14]$$

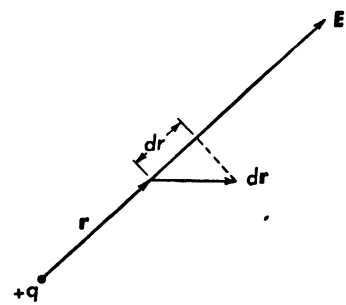


Fig. 2.7

whence

$$dV = \frac{-q}{4\pi\epsilon r^3}\boldsymbol{r} \cdot d\boldsymbol{r} \qquad [2\text{-}37]$$

It will be seen from Figure 2.7 that $\boldsymbol{r} \cdot d\boldsymbol{r}$ is a scalar of magnitude $r\,dr$, where dr is the projection of $d\boldsymbol{r}$ in the direction of $\boldsymbol{r}$. We may then rewrite [2-37] as

$$dV = \frac{-q}{4\pi\epsilon r^2}\,dr \qquad [2\text{-}38]$$

The absolute potential at a distance r from the point charge is now found by integrating [2-38] between the limits $r = \infty$ and $r = r$:

$$V(r) = \frac{-q}{4\pi\epsilon}\int_{\infty}^{r} \frac{dr}{r^2} = \frac{q}{4\pi\epsilon r}\Bigg]_{\infty}^{r} = \frac{q}{4\pi\epsilon r} \qquad [2\text{-}39]$$

If a number of point charges inhabit the region under consideration, as in Figure 2.8, the potential arising from each of the point charges takes the form of equation [2-39]. Hence, because of the scalar nature of potential, the *total* potential, $V(P)$, at the point P, is the algebraic sum of the potentials arising from each of the individual charges. Thus,

$$V(P) = \frac{1}{4\pi\epsilon}\left[\frac{q_1}{r_1} + \frac{q_2}{r_2} + \cdots + \frac{q_n}{r_n}\right] = \frac{1}{4\pi\epsilon}\sum_1^n \frac{q_i}{r_i} \qquad [2\text{-}40]$$

If the charge distribution is a continuous one, the total potential is evidently given by the integral

$$V(P) = \frac{1}{4\pi\epsilon}\int_v \frac{\rho dv}{r} \qquad [2\text{-}41]$$

where the integration extends over the entire volume occupied by the charge, and r is the distance from the volume element to the point P.

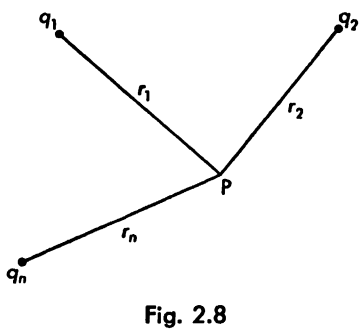


Fig. 2.8

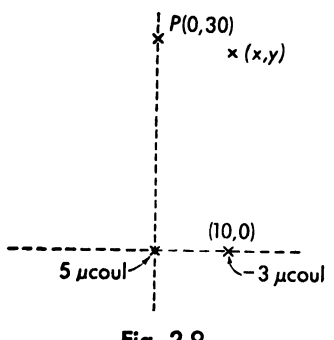


Fig. 2.9

Because of the simplicity of thus calculating the total potential, often the simplest way to compute the field of a number of point charges or of a continuous charge distribution is to compute the potential first and obtain the field components therefrom by the use of the relations given in equation [2-28], Problem 2.4, etc. This is demonstrated in the following illustrative example.

Illustrative Example: Let us solve Problem 1.8 by first calculating the potential. The charge configuration is shown in Figure 2.9.

First, we calculate the potential, V, at some general point (x,y). Thus, from equation [2-40],

$$V = \frac{5 \times 10^{-6}}{4\pi\epsilon\sqrt{x^2 + y^2}} - \frac{3 \times 10^{-6}}{4\pi\epsilon\sqrt{(x-10)^2 + y^2}} \qquad [2\text{-}42]$$

Then

$$E_x = \frac{-\partial V}{\partial x} = \frac{10^{-6}}{4\pi\epsilon}\left[\frac{5x}{[x^2 + y^2]^{3/2}} - \frac{3(x-10)}{[(x-10)^2 + y^2]^{3/2}}\right] \qquad [2\text{-}43]$$

and $$E_y = \frac{-\partial V}{\partial y} = \frac{10^{-6}}{4\pi\epsilon}\left[\frac{5y}{[x^2+y^2]^{3/2}} - \frac{3y}{[(x-10)^2+y^2]^{3/2}}\right]$$

E_x and E_y are now evaluated at the point (0,30) to obtain

$$E_x = \frac{10^{-6}}{4\pi\epsilon}\left[0 + \frac{30}{1000^{3/2}}\right] = \frac{7.5 \times 10^{-11}}{\epsilon} \qquad [2\text{-}44]$$

and $$E_y = \frac{10^{-6}}{4\pi\epsilon}\left[\frac{150}{900^{3/2}} - \frac{90}{1000^{3/2}}\right] = \frac{2.2 \times 10^{-10}}{\epsilon}$$

Thus, $$\boldsymbol{E} = i E_x + j E_y \text{ volts/meter} \qquad [2\text{-}45]$$

and makes an angle with the X-axis of $\theta = \tan^{-1}(22/7.5)$ radians.

2.6 Potential Arising from Spherically and Cylindrically Symmetric Charge Distributions

We have already seen in § 1.9 and Problem 1.10 that in calculating the field strength at a point external to any spherically symmetric charge distribution we may assume that the total charge is concentrated at the center of the sphere. It follows therefore that the potential at any external point takes the same form as for a point charge (equation [2-39]), namely, $V = q/4\pi\epsilon r$, where r is the distance from the center of the charge distribution to the point in question. This result may be derived directly, say, for a conducting sphere, by finding the potential arising from an element of surface and integrating over the surface (Problem 2.13).

The potential at the *surface* of a charged conducting sphere or shell is of particular interest. This has the value $V = q/4\pi\epsilon R$ (where R is the radius of the spherical conductor) and is said to be the *potential of the conductor*. The state of charge of a conductor is usually described by stating its potential.

Inside a conductor, as we have seen in § 1.9, the electric field is zero. Consequently, no work is done in moving a test charge from one internal point to another. It follows that the interior of a conductor is an equipotential region in which the value of the potential is identical with its value at the surface.

Potential Arising from a Line Charge and a Cylindrically Symmetric Charge

We make use here of the expression derived in § 1.6 for the field of a long charged wire:

$$\boldsymbol{E} = \frac{\lambda}{2\pi\epsilon r^2}\boldsymbol{r}$$

where r, normal to the wire, is the vector distance from the axis of the wire to the point at which the field is evaluated, and λ is the charge per unit length.

In keeping with our convention that the zero of potential is at infinity, the potential at the point in question would be obtained by evaluating the integral

$$V = -\int_{\infty}^{r} \boldsymbol{E}\cdot d\boldsymbol{r} = \frac{-\lambda}{2\pi\epsilon}\int_{\infty}^{r}\frac{dr}{r} = \frac{\lambda}{2\pi\epsilon}\ln r\Big]_{r}^{\infty} \qquad [2\text{-}46]$$

This unfortunately has an infinite value, requiring that the conventional zero of potential be abandoned for this particular charge configuration. Let us agree that the zero of potential is located at some distance r_0 from the wire, in which case the potential at a distance r is given by

$$V = \frac{\lambda}{2\pi\epsilon}\ln r\Big]_{r}^{r_0} = \frac{\lambda}{2\pi\epsilon}\ln\left(\frac{r_0}{r}\right) \qquad [2\text{-}47]$$

The reader should not be distressed by this rather arbitrary forsaking of the conventional zero of potential. What is of basic importance is the difference in potential between two points. The adoption of a zero of potential simply provides an agreed-upon reference point from which such differences may be measured.

We have noted in § 1.9 that the charge on a long cylindrical conductor may be considered, by an outside observer, to be concentrated along the axis. The potential of the surface of a conducting cylinder of radius R is, therefore,

$$V = \frac{\lambda}{2\pi\epsilon}\ln\left(\frac{r_0}{R}\right) \qquad [2\text{-}48]$$

2.7 Distribution of Charge on an Irregularly Shaped Conductor

Knowing that the electric field strength is given by the gradient of the potential, we may draw qualitative conclusions concerning the distribution of electric charge on the surface of an irregularly shaped conductor. Let us consider the needle-shaped conductor shown in Figure 2.10.

In the first place, since the surface of the conductor is an equipotential surface, the other equipotential surfaces that are *close* to the conductor, such as $V = V_1$, have shapes that closely resemble the surface of the conductor. On the other hand, at a *distance* from the conductor, the conductor appears as a point charge, and the equipotential surfaces are spherical in shape, as for $V = V_3$. Between these two extreme cases the equipotential surfaces possess transitional shapes, such as $V = V_2$.

Evidently, particularly in the region close to the conductor, the space rate of change of potential is larger along the line AX than it is along the line BY. Clearly, then, the field strength at A is larger than at B. Further, since the magnitude of the field is related to the surface charge density, as follows,

$$E = \frac{\sigma}{\epsilon} \qquad [1\text{-}50]$$

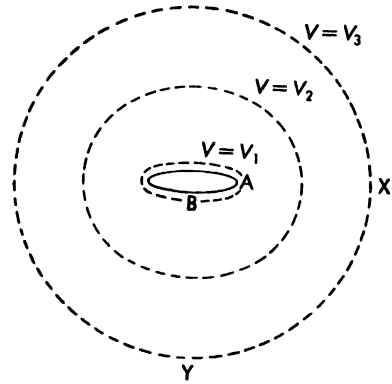


Fig. 2.10

we may conclude that the charge density at A is greater than at B. In general, the superficial charge density is largest on those portions of the surface where the curvature is greatest.

At sharp points on a conductor the charge density is exceptionally high, and the resulting field in the nearby air may be large enough to ionize molecules of the air. If the conductor is positively charged, the positive ions that are created are repelled, while the electrons that have been released in the ionization process are attracted to the conductor, where they neutralize a portion of its charge. If the conductor is negatively charged, it is the positive ions that are attracted to it. In both cases the net effect is loss of charge by the conductor, and the phenomenon is known as *corona discharge*. In the process light is emitted by some of the gas molecules that have been excited by collision with electrons.

The presence of other conductors also affects the distribution of charge on a given conductor.

2.8 Equipotential Surfaces and Lines of Force

The conclusion drawn in § 2.4, that the electric field is everywhere normal to an equipotential surface, is of the greatest value in determining graphically the direction of the lines of force in the region of a given charge configuration. This will be illustrated with reference to Figure 2.11, where charges $q_1 = +10(4\pi\epsilon)$ microcoulombs and $q_2 = -10(4\pi\epsilon)$ microcoulombs are located at the points (0,0) and (2,0), respectively.

The first step in drawing the equipotential lines for the entire configuration is to draw those for the individual charges involved. Each of these is a set of concentric circles surrounding the charge in question, the potential at each point being given by $V_i = q_i/4\pi\epsilon r_i$. In the charge configuration of Figure 2.11 there are two such sets of equipotential lines. These are shown in the diagram

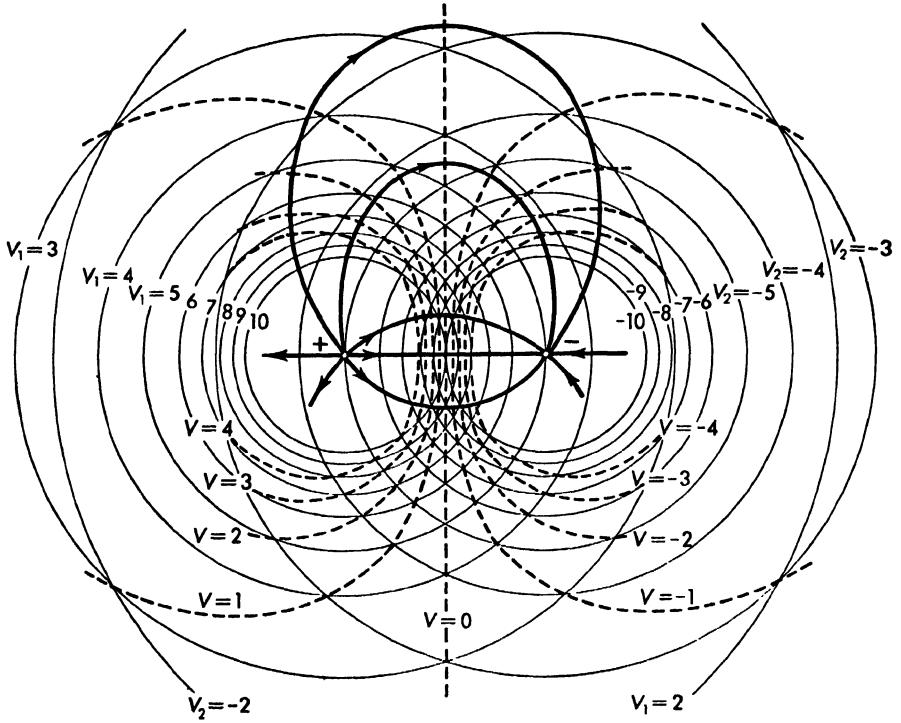


Fig. 2.11

as thin solid lines and are labeled V_1 and V_2. All values of potential shown in the diagram are in microvolts.

Next the total potential, $V = \sum_1^n q_i/4\pi\epsilon r_i$, is computed by adding algebraically the separate potentials. In Figure 2.11, for example, the total potential is easily evaluated at the points of intersection of the equipotential circles belonging to the two individual charges. The total potential is evaluated at a sufficiently large number of points, so that, when points of the same potential are joined together, a representative picture is obtained of the equipotential lines for the entire system. These equipotential lines are shown in Figure 2.11 as broken lines (– – – –).

Finally, the lines of force are drawn so as to be everywhere orthogonal to the equipotential lines. A few such lines are shown on the diagram as heavy solid lines. If the diagram had been completed, the resulting pattern of lines of force would have been similar to that on the right in Figure 1.4. At each point the direction of the field is given by the tangent to the line of force passing through the point. In addition, the spacing of the equipotential surfaces is a measure of the strength of the field. In keeping with the definition of potential, the field vectors point in the direction of decreasing potential.

Experimental Plotting of Equipotential Lines

The question of electrostatic lenses for the focusing of charged particles is discussed below in § 2.10. In designing these lenses and, for that matter, in designing any equipment in which charged particles move under the action of electric fields, it is important to have knowledge of the directions of the fields involved. This may be got by plotting experimentally the equipotential lines and obtaining the field directions therefrom in the manner described above. Apparatus which is commonly used for this purpose is shown schematically in Figure 2.12.

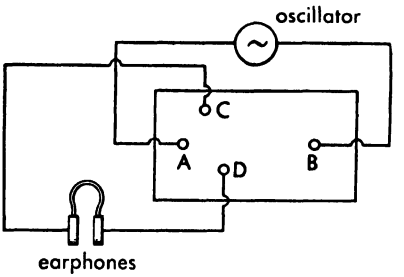


Fig. 2.12

Two electrodes, A and B, connected to a source of audio-frequency power, dip into a tray of slightly conducting water and establish the electric field characteristic of two, equal, unlike charges. Two point probes, C and D, also dipping into the solution, are connected through earphones. When C and D are at the same potential, zero current passes through the earphones.[4] Thus, an equipotential line may be plotted by leaving one of the probes in a fixed position and moving the other to positions where there is a sound null in the earphones.

In order to plot the equipotential lines of other charge configurations, the electrodes A and B are replaced by those of the desired geometry.

2.9 Acceleration of Charged Particles

The difference in potential between two points is the work done in moving a unit positive charge from one point to the other. This work, ΔV joules per coulomb, is positive if the motion is from a position of lower potential to a position of higher potential. The work is then done by an external agent. Conversely, if the particle possessing unit positive charge is released at the position of higher potential and allowed to move under the action of the field to the position of lower potential, work in the amount of ΔV joules is done upon it by the system, and the particle consequently gains ΔV joules of kinetic energy. For a particle possessing a charge of $+q$ coulombs the gain in kinetic energy is $q\ \Delta V$ joules. Negatively charged particles, of course, gain kinetic energy in moving to regions of higher potential.

[4] This method assumes that the potential distribution for a fixed system of charges can be found by studying a system in which charges are in motion. This assumption is justified in § 5.5.

The Electron-Volt

The unit of energy most commonly used in describing the energy of atoms and subatomic particles is the *electron-volt*, abbreviated as eV. *One eV is the kinetic energy gained by one electron when it is accelerated through a potential difference of 1 volt.* The relationship between the electron-volt and the joule is found by expressing the electronic charge in coulombs, as follows:

$$1eV = (1.6 \times 10^{-19} \text{ coulombs}) \times (1 \text{ volt}) = 1.6 \times 10^{-19} \text{ joules} \quad [2\text{-}49]$$

Although the electron-volt is defined in terms of an acceleration involving an electron, it is simply a convenient unit of energy in terms of which the energy of any particle, electron or otherwise, may be expressed (cf. Problem 2.15).

Let us now calculate the potential difference that is required to give an electron a velocity of 10^6 m/sec.

$$\begin{aligned} \tfrac{1}{2}mv^2 &= \tfrac{1}{2}(9 \times 10^{-31})(10^6)^2 \text{ joules} \\ &= 4.5 \times 10^{-19} \text{ joules} \\ &= 2.8eV \end{aligned}$$

Thus, a modest accelerating voltage of less than 3 volts is sufficient. The plate voltage in a vacuum tube is often ~300 volts. Under these conditions the electrons striking it possess a velocity of $\sim 10^7$ m/sec.

Relativistic Effects

If we make a calculation similar to the above for an electron that has fallen through a potential of 500,000 volts, we find that the electron velocity corresponding to $5 \times 10^5 eV$ is 4.2×10^8 m/sec, a value which exceeds the velocity of light ($c = 3 \times 10^8$ m/sec). The theory of relativity, however, decrees that *the velocity of a particle can never exceed the velocity of light* and that the usual expression for the kinetic energy ($\frac{1}{2}mv^2$) is not adequate if the particle velocity approaches this limiting value. For such relativistic velocities the correct expression is

$$\text{K.E.} = m_0c^2\left[\frac{1}{\sqrt{1-\beta^2}} - 1\right] \quad [2\text{-}50]$$

where m_0 is the mass of the particle at rest (the *rest mass*) and $\beta = v/c$ is the ratio of the particle velocity to the velocity of light. Equation [2-50] can be shown (Problem 2.19) to reduce to the classical value, $\frac{1}{2}m_0v^2$, for nonrelativistic values of velocity.

The reference above to rest mass implies that the mass of a particle in motion differs from its value at rest, as is indeed the case. In relativistic mechanics

the mass of a particle is given by

$$m = \frac{m_0}{\sqrt{1 - \beta^2}} \qquad [2\text{-}51]$$

which approaches infinity as the particle velocity approaches that of light. It is thus seen that the work done in accelerating a particle that is already traveling at relativistic velocities manifests itself largely as an increase in the particle mass. In the case of electrons, an acceleration through 10,000 volts results in a 4 percent increase in effective mass.

Van de Graaff Generator

In 1913 it was demonstrated by Ernest Rutherford (1871-1937), the eminent New Zealand-born British physicist, that the positive charges in the atom are located in a diminutive central region (the nucleus). The fact that they are confined to this small volume, in spite of the electric forces of repulsion between them, indicates that powerful attractive forces are also operative. These *nuclear forces* act between both like and unlike nucleons, and have an extremely short range. One of the most effective means of investigating the nature of these nuclear forces is to strike a nucleus with an energetic particle and observe its reaction. If the bombarding particle is positively charged, it must overcome a *Coulomb barrier* (see Problem 2.20) before it can deliver its blow: it must, therefore, possess a certain minimum initial kinetic energy. As this *threshold energy* usually represents an acceleration through a voltage that is much larger than is provided by the conventional electronic power supplies, the desire to study nuclear reactions has led to the invention of several ingenious devices for the acceleration of charged particles. Perhaps the most widely used of these is the Van de Graaff generator, shown schematically in Figure 2.13.

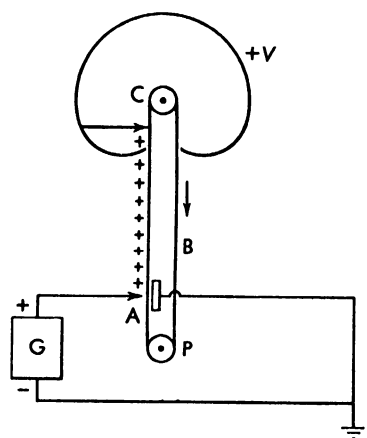


Fig. 2.13

An endless belt, B, made of flexible, insulating material, is driven rapidly by the pulley P in the direction indicated. Electric charge from the generator G is sprayed from the sharp point A onto the upward-moving belt. At the sharp point C this charge is "combed" from the belt and immediately makes its way to the outer surface of the large, hollow metal sphere. Regardless of the density of charge on the outer surface of the sphere, the field inside remains zero and

each cargo of charge, as it arrives, can be removed from the belt. In this way the charge on the conductor is continually augmented until the potential reaches an equilibrium value that is determined principally by the rate at which the charge leaks from the outer surface. The limiting voltage for an air-surrounded installation is ~2.5 MV (megavolts), a figure which can be more than doubled if the accelerator is surrounded by a gas with high ionization potential and at a pressure of several atmospheres.

Charged particles are accelerated along an evacuated tube between the charged sphere and ground. The beam of accelerated charged particles from such a generator is characterized by its high intensity (several ma) and by the homogeneity of its energy. Although the principle on which this accelerator is based is easily understood, the construction of the instrument presents many engineering difficulties. Other types of accelerator will be described in Chapters 8 and 9.

2.10 Electrostatic Lenses

Electrostatic lenses are widely used in the electronics industry and elsewhere for the focusing of beams of charged particles, usually electrons. The focusing action is secured by an arrangement of electrodes that are at different potentials. A simple converging lens will now be described.

The left side of Figure 2.14 shows some of the equipotential surfaces for a

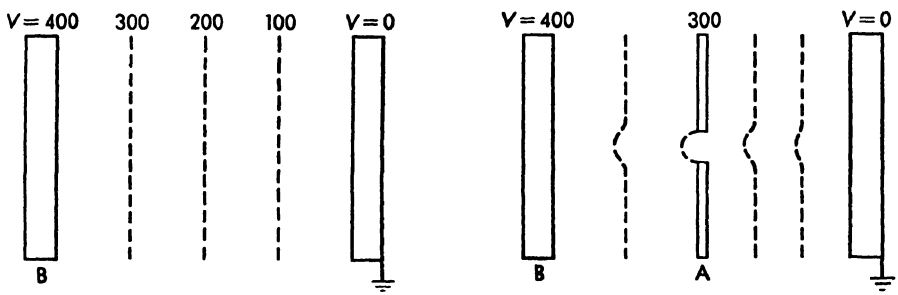


Fig. 2.14

parallel-plate capacitor, one of whose plates is grounded while the other is at a potential of +400 volts. If the edge effects are disregarded, these equipotential surfaces are equally spaced and parallel to the capacitor plates. Let us now, as in the right side of Figure 2.14, place a conducting plate A, whose potential is +300 volts, in the position normally occupied by the 200-volt equipotential surface. Plate A contains a circular aperture. Here the 300-volt surface, which elsewhere coincides with plate A, bulges out toward the

position it occupied before plate A was introduced. Similarly, other nearby equipotential surfaces also bulge to the left, towards their unperturbed positions, the degree of bulging being greatest in the aperture itself. A magnified picture of the shape of the equipotential surfaces and the corresponding field lines in the region of the aperture is shown in Figure 2.15.

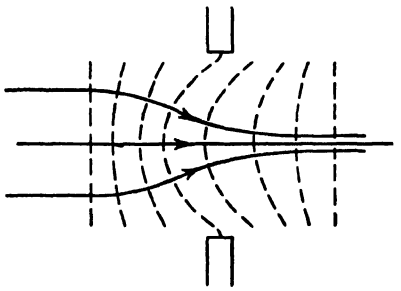


Fig. 2.15

The field is such that a beam of positively charged particles traveling from left to right experiences a focusing force, both those above and below the central line being pushed toward it.[5] This electrode arrangement is thus the electrostatic equivalent of a convex lens. The charged particles to be focused have access to the lens through a small hole in plate B. This type of lens, with potentials reversed, is suitable for the focusing of the electron beam in a cathode ray or television tube. The part of the tube that contains the filament from which the electrons are emitted, and the electrodes for accelerating and focusing the electron beam, is known as the *electron gun.*

An electrostatic focusing system is far more flexible than its optical analogue in that its focal point and magnification may be altered by simply changing the electrode potentials.

The Electron Microscope

The ordinary optical microscope can be used to observe small objects provided their dimensions are no smaller than the wavelength of the light by which they are being viewed. Otherwise, the light simply passes over the object without being significantly affected by it and cannot, therefore, transmit information concerning it. As the wavelength of visible light is 4000Å – 7000Å, the microscope is not useful in the study of objects smaller than $\sim 5 \times 10^{-5}$ cm. Shorter electromagnetic waves exist and could, in theory, be used to reduce this limit (e.g., x-rays have wavelengths of $\sim$1 Å). No method, however, is available for the construction of lenses that will focus such short waves.

On the other hand, in the *electron microscope*, very small objects can be viewed by means of electrons, a fact undreamt of prior to 1924. In that year

---

[5] The reader is reminded that the charged particle trajectories do not coincide with the field lines shown in Figure 2.15. The direction of motion of a charged particle at any instant is determined not only by the electric field but also by the momentum that the particle already possesses. In any case, however, as stated, the electric field shown in Figure 2.15 will have a focusing effect upon a left-to-right traveling group of positively charged particles.

it was speculated by Louis de Broglie (1892–), a young French physicist, that a particle in motion might display a wave nature, its wavelength λ (known as the *de Broglie wavelength*) being given by

$$\lambda = \frac{h}{mv} \qquad [2\text{-}52]$$

where m and v are the mass and velocity, respectively, of the particle, and h is a fundamental constant, known as *Planck's constant*, having the value $h = 6.62 \times 10^{-34}$ joule-sec. The correctness of this speculation was confirmed in 1927 by C. J. Davisson (1881–) at the Bell Telephone Laboratories (United States of America) and by G. P. Thomson (1892–) at the University of Aberdeen, each of whom independently observed, with electrons, certain diffraction effects which could be attributed to waves of length given by equation [2-52].

The electron microscope not only depends for its operation upon the fact that electrons possess a wave nature but also presupposes that methods are available by which they can be focused. Electrostatic lenses which operate on the principle described above are entirely suitable for this latter purpose. Magnetic lenses of a type to be described in § 8.17 are also widely used.

An electron that has fallen through a potential difference of 50,000 volts has a velocity of $\sim 10^8$ m/sec and a de Broglie wavelength of ~ 0.07 Å. Such an electron is, therefore, well suited for the transmission of information concerning objects that are too small to be observed by the usual optical means. With the electron microscope, successful observations have been made of objects that are only a few Ångstrom units in size. In this instrument the electrons pass through the object under study and subsequently strike a fluorescent screen or a photographic plate where an enlarged image of the object is recorded.

The wave nature of the electron provides a natural explanation for the "permitted" orbits in an atom. This question is raised in Problem 2.22.

2.11 Electrical Images

Many special problems in electrostatics can be greatly simplified by utilizing the method of electrical images, invented by Lord Kelvin (1824–1907), British mathematician and physicist. In this method, complicated distributions of charge are replaced, for purposes of calculation, either by a single point charge or by a much-simplified charge distribution. The method will be illustrated with reference to a couple of simple, but interesting, cases.

Charge Induced on a Grounded Conductor by a Point Charge

Let us suppose that two equal and opposite point charges are located as shown in Figure 2.16.

At any point P along the central plane the potential vanishes since

$$V = \frac{1}{4\pi\epsilon}\left[\frac{q}{r} - \frac{q}{r}\right] = 0 \quad [2\text{-}53]$$

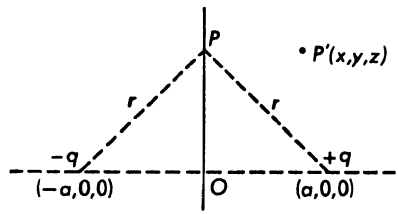


Fig. 2.16

One may, therefore, place a grounded conducting sheet ($V = 0$) in the central plane without in any way disturbing the field, because the potential of such a conductor coincides at every point with the potential existing before its insertion. The lines of force from $+q$ to the grounded conductor follow precisely the same path that they did before the introduction of the conductor.

Conversely, the field between the point charge $+q$ and a grounded plane conductor can be found by replacing the conductor with a point charge $-q$. This *image charge* is located as far behind the surface of the conductor as the real charge was in front, and the line joining it to the real charge intersects the plane of the conductor at right angles.

With the help of this image charge one can readily compute the density of the charge induced on the conducting plane by the point charge. To do so, let us calculate the potential at any point $P'(x,y,z)$ arising from the actual charge and its image:

$$V(P') = \frac{q}{4\pi\epsilon}\left[\frac{1}{\sqrt{(x-a)^2 + y^2 + z^2}} - \frac{1}{\sqrt{(x+a)^2 + y^2 + z^2}}\right] \quad [2\text{-}54]$$

From this the field components at the point P' can be obtained by differentiation. In particular, at the point $P(0,y,z)$ these have the value

$$E_x = -\left(\frac{\partial V}{\partial x}\right)_{x=0}, \quad E_y = E_z = 0 \quad [2\text{-}55]$$

The point P, however, lies on the surface of the conductor, in which case the field can be written in terms of the charge density on the conductor, viz.,

$$E = E_x = \frac{\sigma}{\epsilon} \quad [1\text{-}50]$$

Equating [2-55] and [1-50], one obtains:

$$\sigma = -\epsilon\left(\frac{\partial V}{\partial x}\right) \quad [2\text{-}56]$$

or
$$\sigma = \frac{-q}{4\pi}\left[\frac{2a}{(a^2 + y^2 + z^2)^{3/2}}\right] \qquad [2\text{-}57]$$

or
$$\sigma = \frac{-aq}{2\pi(a^2 + y^2 + z^2)^{3/2}} \qquad [2\text{-}58]$$

As the quantity $(a^2 + y^2 + z^2)^{1/2}$ is the distance from the point charge to the point on the conducting plane, it is seen that the superficial charge density on the plane varies inversely as the cube of the distance from the point charge. Problem 2.24 is an application of this result.

General Method of Electrical Images

The problem treated above, that of the charge density induced on a conducting plane by a point charge, provides the simplest example of the method of images. Let us review the general principle underlying the method.

In a system of point charges the potential at any point is given by equation [2-40]

$$V = \frac{1}{4\pi\epsilon}\sum_{1}^{n}\frac{q_i}{r_i} \qquad [2\text{-}40]$$

The surface of zero potential (that is, the grounded conductor) is the locus of points for which

$$V = 0 = \frac{1}{4\pi\epsilon}\sum_{1}^{n}\frac{q_i}{r_i} \qquad [2\text{-}59]$$

We may rewrite equation [2-59] as the sum of contributions from two groups of charges:

$$V = 0 = \frac{1}{4\pi\epsilon}\sum_{1}^{j}\frac{q_i}{r_i} + \frac{1}{4\pi\epsilon}\sum_{j+1}^{n}\frac{q_i}{r_i} \qquad [2\text{-}60]$$

Thus, if $q_1, q_2, \cdots, q_j$ constitute the system of real charges, and the grounded conductor possesses the shape required by equation [2-59], the group of charges $q_{j+1}, \cdots, q_n$ constitutes the electrical image. In the calculation of electrical effects in the region outside, or on the surface, of the grounded conductor, the conductor may be replaced by the system of charges constituting the image. As expressed in equation [2-60], this is permissible because the image charge and the system of real charges combine to produce zero potential over the surface occupied by the grounded conductor.

If the conductor under discussion is at a potential $V \neq 0$, the system of charges and its image must combine to produce the actual potential. Equation [2-60] then becomes:

$$V = \frac{1}{4\pi\epsilon}\sum_{1}^{j}\frac{q_i}{r_i} + \frac{1}{4\pi\epsilon}\sum_{j+1}^{n}\frac{q_i}{r_i} \qquad [2\text{-}61]$$

Charge Induced on a Conducting Sphere by a Point Charge

Next to an infinite plane, the surface whose shape can be most easily expressed in the form of equation [2-59] is that of a sphere. In Figure 2.17 a point charge q is brought near to a grounded conducting sphere. Let us now calculate the distribution of the charge that is induced upon the sphere by the external point charge.

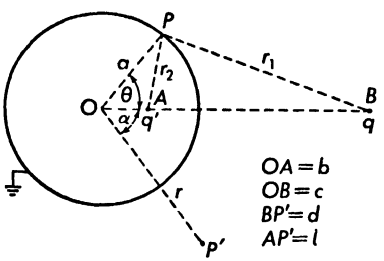


Fig. 2.17

Let us suppose that an image charge q' placed at the point A (Figure 2.17), when combined with $+q$, leads to zero potential at any point P on the surface of the sphere. This assumption presupposes that the $V = 0$ surface for a pair of opposite (but not equal) charges can be spherical in shape. If this hypothesis is untrue, it will not be possible to obtain a solution for the position and magnitude of q'.

The potential at P, due to the real charge q and its image charge q', is given, by definition, as

$$V(P) = 0 = \frac{1}{4\pi\epsilon}\left[\frac{q}{r_1} + \frac{q'}{r_2}\right] \qquad [2\text{-}62]$$

Substituting

$$r_1^2 = a^2 + c^2 - 2ac\cos\theta \qquad [2\text{-}63]$$

and

$$r_2^2 = a^2 + b^2 - 2ab\cos\theta$$

we obtain

$$\frac{q'}{\sqrt{a^2 + b^2 - 2ab\cos\theta}} = \frac{-q}{\sqrt{a^2 + c^2 - 2ac\cos\theta}} \qquad [2\text{-}64]$$

If we now choose $b = a^2/c$, equation [2-64] becomes, after both sides are multiplied by a/c,

$$\frac{q'}{\sqrt{a^2 + c^2 - 2ac\cos\theta}} = \frac{-q\left(\frac{a}{c}\right)}{\sqrt{a^2 + c^2 - 2ac\cos\theta}} \qquad [2\text{-}65]$$

Thus, the image charge in this case is a point charge of magnitude $q' = -q(a/c)$, placed a distance a^2/c from the center of the sphere along the line connecting the center of the sphere to the real point charge q.

As before, in order to calculate the density of charge induced on the sphere due to the presence of q, we can evaluate the potential at some external point P' that is located a distance r from the center of the sphere. Thus,

$$V(P') = \frac{q}{4\pi\epsilon d} - \frac{q\left(\frac{a}{c}\right)}{4\pi\epsilon l} \qquad [2\text{-}66]$$

where d and l are the distances BP' and AP', respectively. Substituting the values of d and l (corresponding to equation [2-63]), we obtain

$$V(P') = \frac{1}{4\pi\epsilon}\left[\frac{q}{\sqrt{c^2 + r^2 - 2cr\cos\alpha}} - \frac{q\left(\frac{a}{c}\right)}{\sqrt{\frac{a^4}{c^2} + r^2 - 2\left(\frac{a^2 r}{c}\right)\cos\alpha}}\right] \qquad [2\text{-}67]$$

The field at the surface of the sphere, always given by $E = \sigma/\epsilon$ (equation [1-50]), is now obtained by differentiating equation [2-67] with respect to r and evaluating the derivative at the surface of the sphere:

$$E = -\left(\frac{\partial V}{\partial r}\right)_{r=a} \qquad [2\text{-}68]$$

From this it may readily be shown that the superficial induced-charge density on the sphere is:

$$\sigma = \epsilon E = \frac{-q}{4\pi a}\,\frac{c^2 - a^2}{(a^2 + c^2 - 2ac\cos\alpha)^{3/2}} \qquad [2\text{-}69]$$

In the two examples which have been given, both the real and the image charges have been single point charges. A somewhat more complicated case, in which the image charge consists of a group of charges, is considered in Problem 2.27. The problem is considerably more complicated if the real charge is not a point but is, say, a charged conducting sphere. When this sphere is brought near a grounded conducting plane, the charge induced on the plane exerts an attraction on the charges located on the sphere, with the result that the original uniform charge distribution is altered. This problem and others in which the method of images provides a valuable method of attack in spite of the complexity of the situation are discussed in advanced texts on electricity.[6]

---

[6] See, for example, Sir James Jeans: *The Mathematical Theory of Electricity and Magnetism*, Cambridge University Press, 1925.

2.12 Use of Laplace's Equation in the Solution of Electrostatic Problems [7]

We have briefly referred in § 2.4 to Laplace's equation,

$$\nabla^2 V = 0 \qquad [2\text{-}36]$$

which describes the potential distribution in a charge-free region. The solving of any electrostatic problem involving such a region is equivalent to finding a solution of Laplace's equation which satisfies the conditions at the boundaries of the space under consideration. Usually these boundaries are infinity, on the one hand, and the surfaces of the various conductors that are present, on the other. At infinity the potential must be zero, while at the surfaces of the conductors it must have the values that are associated with them. It can be shown, although it will not be done here, that any solution to Laplace's equation which satisfies the boundary conditions is unique. The use of Laplace's equation is demonstrated below with reference to three special problems. Problems 2.28–2.32 are intended to involve its use.

Potential Outside a Charged Conducting Sphere

Let us solve Laplace's equation for the region outside a charged metal sphere of radius a and potential V_a. For a problem such as this, which possesses spherical symmetry, it is convenient to express the Laplacian operator, ∇^2, in terms of spherical polar coordinates (see Appendix). Laplace's equation then becomes

$$\nabla^2 V = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial V}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 V}{\partial \phi^2} = 0 \qquad [2\text{-}70]$$

Because of the spherical symmetry, V is a function of r only, and equation [2-70] can be rewritten

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2 \frac{dV}{dr}\right) = 0 \qquad [2\text{-}71]$$

This second-order differential equation is to be solved in such a way that $V = 0$ at infinity and $V = V_a$ at the surface of the sphere.

Evidently, the quantity within parentheses in equation [2-71] must be constant, viz.,

$$r^2 \frac{dV}{dr} = C_1 \qquad [2\text{-}72]$$

This, when integrated, yields

$$V = -\frac{C_1}{r} + C_2 \qquad [2\text{-}73]$$

as the general expression for the potential.

[7] This is the first of a number of "small-type" topics that are discussed in the text. These topics, designated by the symbol ■, are more advanced than the rest of the material in the text, but *none of them is prerequisite for any of the normal-type sections.* Better students will find that a perusal of these small-type sections is rewarding.

Since $V = 0$ when $r = \infty$ it follows that $C_2 = 0$. Further, requiring that $V = V_a$ when $r = a$, we have

$$V_a = -\frac{C_1}{a}, \quad \text{or} \quad C_1 = -aV_a \qquad [2\text{-}74]$$

The potential in the region surrounding the charged metal sphere is, therefore,

$$V = \frac{a}{r} V_a \qquad [2\text{-}75]$$

When the known value of V_a is substituted in the above expression, it takes a form with which we are familiar:

$$V = \frac{a}{r}\left[\frac{q}{4\pi\epsilon a}\right] = \frac{q}{4\pi\epsilon r} \qquad [2\text{-}76]$$

■ *Potential Between Coaxial Conducting Cylinders*

This is an instance where infinity is not one of the boundaries. Instead, the space under consideration is bounded by an inner cylinder of radius a and an outer cylinder of radius b. The arrangement is similar to that of the cylindrical capacitator shown in Figure 3.12. The potential of the inner cylinder is V_a, while the outer cylinder is grounded ($V = 0$).

The geometry of this problem suggests that the Laplacian operator should be expressed in terms of cylindrical coordinates (see Appendix). Laplace's equation then becomes

$$\nabla^2 V = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 V}{\partial \theta^2} + \frac{\partial^2 V}{\partial z^2} = 0 \qquad [2\text{-}77]$$

It is assumed that the cylinders are long and that end effects may be disregarded. If so, the potential is independent of both θ and z, and equation [2-77] takes the form

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dV}{dr}\right) = 0 \qquad [2\text{-}78]$$

As before, the quantity within parentheses must be constant—equal, say, to C_1. Integrating

$$r\frac{dV}{dr} = C_1 \qquad [2\text{-}79]$$

we obtain

$$V = C_1 \ln r + C_2 \qquad [2\text{-}80]$$

as the general expression for the potential. The constants C_1 and C_2 are evaluated by applying the boundary conditions, viz.,

$$\text{when} \quad r = b, V = 0, \qquad \text{whence} \quad 0 = C_1 \ln b + C_2 \qquad [2\text{-}81]$$

$$\text{when} \quad r = a, V = V_a, \qquad \text{whence} \quad V_a = C_1 \ln a + C_2 \qquad [2\text{-}82]$$

These two equations when solved yield

$$C_1 = \frac{V_a}{\ln\left(\frac{a}{b}\right)} \quad \text{and} \quad C_2 = \frac{-V_a \ln b}{\ln\left(\frac{a}{b}\right)} \qquad [2\text{-}83]$$

Equation [2-80], giving the potential in the space between the coaxial conducting cylinders, then takes the form

$$V = V_a \frac{\ln\left(\frac{r}{b}\right)}{\ln\left(\frac{a}{b}\right)} \qquad [2\text{-}84]$$

Earnshaw's Theorem

A point charge located in an electric field cannot exist in a state of stable equilibrium under the action of the electric field alone. This rule, known as Earnshaw's theorem, is easily deduced from Laplace's equation.

At any point in the field where the potential is V, a point charge q has potential energy in the amount $W = qV$. A position of stable equilibrium for the point charge is one in which W has a minimum value. Since V is a solution of Laplace's equation, $W = qV$ is also a solution, inasmuch as q is a constant. This is readily proved by substituting qV in Laplace's equation. Thus,

$$\nabla^2 W = \frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2} + \frac{\partial^2 W}{\partial z^2} = 0 \qquad [2\text{-}85]$$

A necessary condition for W to be a minimum is that its second derivative with respect to each of the position coordinates be positive. This is clearly impossible since in equation [2-85] their total is zero.

Before the modern picture of the atom emerged (c1913) Earnshaw's theorem provided the basis for concluding either (a) that atoms did not consist of a group of point charges *at rest* or (b) that the inverse-square law of electric force did not hold at distances comparable to atomic dimensions. The acceptance of the fact that electrons are *in motion* within the atom explained how stable equilibrium could be achieved.

2.13 Problems

E (2.1) A unit vector corresponding to any coordinate η is perpendicular to the surface η = constant and points in the direction in which η is increasing. The student should convince himself that the Cartesian unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ satisfy this description and should then draw at some point in space the unit vectors $\hat{\boldsymbol{r}}$, $\hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{\phi}}$, corresponding to spherical polar coordinates,[8] and $\hat{\boldsymbol{r}}$, $\hat{\boldsymbol{\theta}}$, $\boldsymbol{k}$, corresponding to cylindrical coordinates.[9]

[8] The spherical polar coordinates of a point P are its distance $r = OP$ from the origin, the angle ϕ between OP and the Z-axis, and the angle θ in the X-Y plane between the plane OPZ and the X-axis.

[9] The cylindrical coordinates of a point P are (r, θ, z), where r and θ have the same values as for spherical polar coordinates, and z is the projection of r on the Z-axis.

E (2.2) Show that the two vectors $A = ix + jy + kz$ and

$$B = i\alpha + j\beta - k\left(\frac{x\alpha}{z} + \frac{y\beta}{z}\right)$$

are normal to one another.

E (2.3) Show that $\nabla \times \nabla V = 0$.

M (2.4) By analogy to equation [2-28], show that the field components in the spherical polar system of coordinates are:

$$E_r = -\frac{\partial V}{\partial r}, \quad E_\theta = -\frac{1}{r}\frac{\partial V}{\partial \theta}, \quad E_\phi = -\frac{1}{r \sin\theta}\frac{\partial V}{\partial \phi}$$

E (2.5) The potential of a certain charge configuration is given by $V = 2x + 3xy + y^2$ volts. What acceleration does an electron experience in the X-direction at the point (5,2)?

[Ans.: 1.4×10^{12}m/sec.^{2}]

E (2.6) What is the potential of a uniformly charged spherical shell of radius a meters at a point (a) outside, (b) on the surface, and (c) in the interior of the shell?

E (2.7) What energy is expended in placing in vacuo a +8 millicoulomb charge at the point (0,0,0) and a +10 millicoulomb charge at the point (0,5,0)? The coordinate distances are in meters.

[Ans.: $4 \times 10^{-6}/(\pi\epsilon_0)$ joules.]

M (2.8) Show that the *total* energy of an electron following a circular orbit in a one-electron atom is $-Ze^2/8\pi\epsilon_0 r$ joules, where Ze is the nuclear charge and r is the radius of the orbit.

E (2.9) Calculate the potential at a distance r from an infinite charged sheet.

E (2.10) Assume two isolated conducting spheres of radii a meters and b meters, respectively. One of the spheres is given a charge q and the two spheres are then connected by a long fine wire. How does the charge distribute itself between the two conductors? Note that this result is in qualitative agreement with the conclusion arrived at in § 2.7 concerning the distribution of charge on an irregularly shaped conductor.

H (2.11) A thin dielectric rod of length l meters lies along the X-axis with one end at the origin and the other at the point $(l,0)$. It is charged uniformly with a total charge of Q coulombs. Calculate the potential and field at the point $(x,0)$ when $x > l$.

$$[\text{Ans.: } V = \frac{Q}{4\pi\epsilon l}\ln\frac{x}{x-l}, \; E = Q/4\pi\epsilon x(x-l).]$$

H (2.12) If the charge density on the rod of Problem 2.11 is $\lambda = kx$ coulombs/m, derive expressions for the potential and field at a point $(h,0)$ beyond the end of the wire.

H (2.13) Derive the expression for the potential at a point outside a charged spherical conductor by finding the potential arising from an element of the surface and integrating over the surface of the sphere.

E (2.14) A solid metal sphere of 0.1 m radius is given a charge of 1 microcoulomb. A metal shell of 0.25 m radius is given a charge of .0025 millicoulomb. When the two conductors are connected by a wire, in which direction does the current flow?

E (2.15) What energy in electron volts is gained by a proton during acceleration through 300 volts?

E (2.16) In a two-electrode radio tube electrons are emitted from a plane cathode and pass to a flat metal plate .02 m away. A 200-V battery is connected externally between cathode and plate with the latter at the higher potential. How long does it take the electrons to travel from cathode to plate? [Ans.: 0.005 μsec.]

M (2.17) If the current in Problem 2.16 is 6 ma, what is the force exerted on the plate by the arriving electrons? [Ans.: 3×10^{-7} newtons.]

M (2.18) In a Geiger counter a thin wire extends along the axis of a cylindrical metal tube. This wire and the cylinder are connected externally through a 2000-V battery, with the cylinder at the higher potential. If the radius of the wire is 0.1 mm and the inner radius of cylinder is 2 cm, at what distance from the surface of the wire is the potential 1000 V above that of the wire?

M (2.19) Show that equation [2-50] for the relativistic kinetic energy of a particle reduces to the usual value for the kinetic energy if $v/c \ll 1$.

M (2.20) The radius of an atomic nucleus is given to a good approximation by $r = r_0 A^{1/3}$, where $r_0 = 1.4 \times 10^{-13}$ cm and A is the mass number of the atom in question. What energy must an α-particle (a doubly-charged atom of He^4) possess in order to strike the nucleus of Te^{125} ($Z = 52$)? This value of energy is the *Coulomb barrier* which the α-particle must overcome before reaching the nucleus of Te^{125}. Actually, α-particles with somewhat less energy can reach the nucleus by "tunneling" through the final portion of the Coulomb barrier.

H (2.21) The following is quoted (with some changes in symbols) from a paper published by Rutherford in 1911. "In order to form some idea of the forces required to deflect an α-particle through a large angle, consider an atom containing a positive charge Ze at its center, and surrounded by a distribution of negative electricity, $-Ze$, *uniformly* distributed within a sphere of radius R. The electric force E and the potential V at a distance r from the center of an atom, for a point inside the atom, are given by

$$E = \frac{Ze}{4\pi\epsilon_0}\left[\frac{1}{r^2} - \frac{r}{R^3}\right]$$

and

$$V = \frac{Ze}{4\pi\epsilon_0}\left[\frac{1}{r} - \frac{3}{2R} + \frac{r^2}{2R^3}\right]$$

Prove the correctness of these expressions. When predictions based

on this model were found to agree with experiment, the idea of a *nuclear* atom was established.

M (2.22) In the ground state of the hydrogen atom the electron's angular momentum (see footnote to § 1.2) is $h/2\pi$. Calculate the de Broglie wavelength of the electron and compare it with the length of its orbit.

E (2.23) A point charge of 30 microcoulombs is located 3 cm from a grounded conducting sheet. With what force is it attracted toward the sheet?

[Ans.: $10^{-6}/(16\pi\epsilon)$ newtons.]

M (2.24) In Problem 2.23 show that the charge induced on the portion of the plane that is contained by the circumference of a circle of radius 4 cm, whose center is directly opposite the point charge, is numerically equal to 12 microcoulombs.

M (2.25) A grounded conducting sphere of radius 4 cm is coated with a plastic layer, as shown in Figure 2.18. The plastic layer is given a uniform charge of Q coulombs after which a point charge of q coulombs is brought to a point 8 cm from the center of the sphere. What is the value of the field at the point P midway between q and the surface of the sphere?

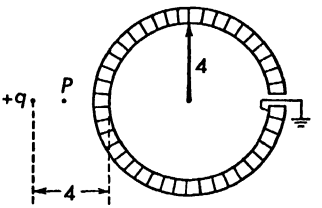


Fig. 2.18

M (2.26) A point charge q is placed inside a grounded metal spherical shell (radius a) at a distance r from its center. Prove that this charge is subject to a force of magnitude

$$\frac{q^2 ar}{4\pi\epsilon_0(a^2 - r^2)^2}$$

M (2.27) A point charge $+q$ is located near the intersection of two grounded conducting planes that intersect at right angles, as shown in Figure 2.19. What is the electrical image in this case?

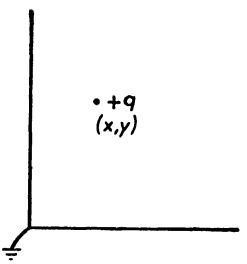


Fig. 2.19

The remaining five problems are based on the small-type material presented in § 2.12.

M (2.28) Show that the potential cannot have a maximum or minimum value at any point in space that is not occupied by an electric charge.

M (2.29) If V is a solution of Laplace's equation, show that $A\cdot\nabla V$ is also a solution, where A is a constant vector.

E (2.30) Use Laplace's equation to obtain the potential in the space between the plates of a parallel-plate capacitor. Assume the plates are normal to the X-axis, are located at $x = 0$ and $x = a$, and are at potentials $V = 0$ and $V = V_a$ respectively.

E (2.31) Use Laplace's equation to obtain the potential in the space between two concentric metal spheres, of radii a and b, respectively. The inner sphere is at the potential V_a and the outer sphere is grounded.

M (2.32) By use of the vector identity $\nabla \times \nabla \times \boldsymbol{E} = \nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E}$, show that $\nabla^2 \boldsymbol{E} = 0$. $\boldsymbol{E}$ is the electrostatic field in a charge-free region.

Chapter 3

Capacitance

3.1 Concept and Definition of Capacitance

In a given environment the potential of a given conductor is directly dependent upon the total charge which it possesses, i.e.,

$$V \propto Q \qquad [3\text{-}1]$$

or

$$V = \frac{Q}{C} \qquad [3\text{-}2]$$

where C, the constant of proportionality between Q and V, is known as the *capacitance of the conductor*. It is seen, by rewriting equation [3-2] in the form

$$C = \frac{Q}{V} \qquad [3\text{-}3]$$

that the capacitance of a conductor is numerically equal to the quantity of charge which is required to raise the potential of the conductor from zero to unity. Evidently, *a conductor has unit capacitance when the addition of unit charge raises its potential by one unit*. This definition holds for any system of units—mks, electrostatic, or other.

The term "capacitance" suggests an ability to store or to contain; in this case the material that is stored is charge. As equation [3-3] predicts, a conductor that has a large capacitance will, for the same value of potential, contain a greater charge than a conductor of small capacitance. Alternatively, if two conductors contain the same charge, the potential of the one with the

larger capacitance is the lower, indicating that less work is required to bring to this conductor a given additional quantity of charge.

The somewhat grandiose term "capacitance," which we have used, is to be preferred to the more familiar term "capacity." The latter, when used, for example, in connection with a pail, refers to the amount of material that the pail can hold—a quantity that is essentially constant. By analogy, one might falsely conclude that electric "capacity" refers to the amount of charge that a conductor can hold. As we have seen, the amount of charge held by a conductor is not a constant but varies with the potential. The use of the word "capacitance" is, therefore, not a sign of pedantry but of the desire unequivocally to designate the ratio Q/V.

3.2 Units of Capacitance

The unit of capacitance in the mks system is the *farad*, named in honor of the English physicist, Michael Faraday, whose work in electrolysis and electromagnetic induction will be discussed later. *A conductor has a capacitance of 1 farad if the addition to it of 1 coulomb of charge changes its potential by 1 volt.* One farad is an exceptionally large capacitance: the entire earth, considered as a spherical conductor of radius 4000 miles, has a calculated capacitance of only 0.0007 farad. Capacitances that are actually encountered are expressed in *microfarads*, or even in *micromicrofarads*. A micromicrofarad is sometimes referred to as a *picofarad:* [1] thus 1 pf = 1 $\mu\mu$f, where $\mu = 10^{-6}$.

3.3 Energy of a Charged Conductor

The operation of assembling upon a conductor a group of charges that mutually repel one another requires work and therefore results in the produc-

[1] A list of prefixes representing the most commonly used powers of 10 has been agreed upon by the International Union of Pure and Applied Physics and the International Electrochemical Commission. These prefixes, many of which are coming into common use, are shown below.

Prefix	*Value*	*Prefix*	*Value*
Tara (T)	10^{12}	Deci (d)	10^{-1}
Giga (G)	10^{9}	Centi (c)	10^{-2}
Mega (M)	10^{6}	Milli (m)	10^{-3}
Kilo (k)	10^{3}	Micro (μ)	10^{-6}
Hecto (H)	10^{2}	Nano (n)	10^{-9}
Deka (D)	10^{1}	Pico (p)	10^{-12}

tion of potential energy. It is customary to say that this potential energy is possessed by the charged conductor itself although, as suggested in § 1.11, it may be more correct to picture the energy as stored in the field surrounding the conductor.

The amount of this stored energy can be computed by considering the charging operation as a piecemeal affair in which the final charge is built up of small increments. Let us suppose that at some stage during this operation the conductor possesses a total charge Q'. Its potential is then

$$V' = \frac{Q'}{C} \qquad [3\text{-}4]$$

Since equation [3-4] represents the work involved in bringing unit charge to the surface of the Q'-charged conductor, that required to deposit an increment of charge dQ' is

$$dW = V'\,dQ' = \frac{Q'\,dQ'}{C} \qquad [3\text{-}5]$$

The total work to attain the final charge Q is

$$\text{Work} = \frac{1}{C}\int_0^Q Q'\,dQ' = \frac{Q^2}{2C} \qquad [3\text{-}6]$$

or

$$\text{Work} = \frac{CV^2}{2} \qquad [3\text{-}7]$$

or

$$\text{Work} = \frac{QV}{2} \qquad [3\text{-}8]$$

The equivalent expressions [3-6], [3-7], and [3-8] thus represent the energy associated with a charged conductor. If the conductor is one of the plates of a capacitor (see below) expressions [3-6] to [3-8] still hold and in this case give the energy of the charged capacitor.

3.4 Principle of the Capacitor

The potential of a charged conductor is affected by the approach of another charged conductor. In particular, its potential is reduced if the approaching conductor is oppositely charged, in accordance with:

$$V = \frac{1}{4\pi\epsilon}\sum_1^n \frac{q_i}{r_i} \qquad [2\text{-}40]$$

As the presence of the second, oppositely charged conductor does not alter the charge, Q, on the first conductor, the decrease in potential which it occasions must be interpreted as an increase in the capacitance ($C = Q/V$) of the first conductor. Further, it is seen from equation [2-40] that an increase in the dielectric constant of the surrounding medium also results in a lowering of the conductor's potential and, consequently, leads to an increase in its capacitance. These two devices, bringing near an oppositely charged conductor and increasing the dielectric constant of the surrounding medium, are in common use. They make it possible to place upon a conductor, for the same expenditure of energy ($= QV/2$), a greater charge than would be possible if the conductor were isolated in free space.

Any device for increasing the capacitance of a conductor is known as a capacitor. Capacitors normally consist of two *plates* placed close together and insulated from one another by air or some dielectric material.

In many applications one of the plates is grounded, as in Figure 3.1. When

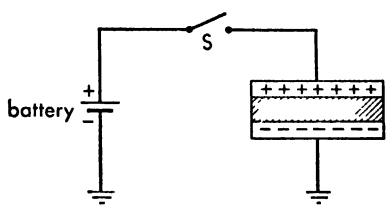


Fig. 3.1

a charge is given to the upper or "hot" plate, by closing the switch S, an *equal and opposite* charge is attracted, from ground, to the lower plate. When the lower plate is connected to a point that is at other than ground potential, its charge is again equal and opposite to that on the top plate, although the amount of the charge on each is different, in conformity with the basic capacitor equation $Q = CV$.

The various shapes of plates that are commonly employed and the relative merits of each are discussed in § 3.7 *et seq.* When capacitors are shown in diagrams they are represented as parallel-plate capacitors unless there is some particular reason for indicating their actual geometry.

A dielectric material between the capacitor plates may serve the additional purpose of allowing a much higher potential difference to exist between the plates, without discharging, than would be the case with an air dielectric. With this use in mind, reference is often made to the *dielectric strength* of a material, defined as the field to which the dielectric material must be subjected before a disruptive discharge occurs. Values of the dielectric coefficient, $k_e = \epsilon/\epsilon_0$, and the dielectric strength for some common dielectrics are shown in Table 1.

TABLE 1
Dielectric Coefficients and Dielectric Strengths of Some Common Materials

	Dielectric coefficient, ($k_e = \epsilon/\epsilon_0$)	*Dielectric strength* (volts/mil [a])
Air	1.000585	75
Water	80	—
Mica	7–8	—
Hard rubber	2.8	470
Paraffin	2.25	250
Quartz, fused	3.85	—
Pyrex glass	4–6	—
Sulphur	4.0	—
Transformer oil	2.2	260–300
Silicones	2.8	250–300
Paper (high quality)	2.23	1230
Bakelite (general purpose)	6.0	300
Lucite	3.3	—
Plexiglass	3.4	—
Polystyrene	2.5	500
Polyethylene	3.3	415
Steatite ceramics	5.5–7.5	200 350
Titanium dioxide ceramics	15–500	—
Barium titanate	500–6000	30–50

[a] A *mil* is 0.001 in.

3.5 Capacitors in Series and Parallel

By grouping individual capacitors in series, in parallel, or in some series-parallel combination, it is possible to secure an effective capacitance which differs in value from that of the individual capacitors and/or can withstand a higher applied voltage.

Series Connection

The series arrangement is shown in Figure 3.2. Here the positive charge $+Q$ on plate A of the first capacitor attracts an equal and opposite charge $-Q$ to plate B. This leaves a charge $+Q$ on plate C of the second capacitor which attracts a charge $-Q$ to the opposite plate D, etc. In short, the magnitude of *the charge Q on each capacitor plate is the same*, regardless of the capacitances of the individual capacitors. The potential differences across the separate capacitors

are, respectively:

$$V_1 = \frac{Q}{C_1}, \quad V_2 = \frac{Q}{C_2}, \quad \cdots, \quad V_n = \frac{Q}{C_n} \qquad [3\text{-}9]$$

Also, we may regard the capacitor combination as a single capacitor of capacitance C, with potential difference V between the two plates A and F.

Thus: $$V = \frac{Q}{C} \qquad [3\text{-}10]$$

Evidently, $$V = V_1 + V_2 + \cdots + V_n \qquad [3\text{-}11]$$

whence $$\frac{Q}{C} = \frac{Q}{C_1} + \frac{Q}{C_2} + \cdots + \frac{Q}{C_n} \qquad [3\text{-}12]$$

and $$\frac{1}{C} = \sum_1^n \frac{1}{C_i} \qquad [3\text{-}13]$$

or, in words, *the reciprocal of the effective capacitance of a number of capacitors connected in series is the sum of the reciprocals of their separate capacitances.*

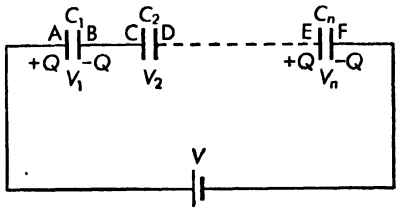


Fig. 3.2

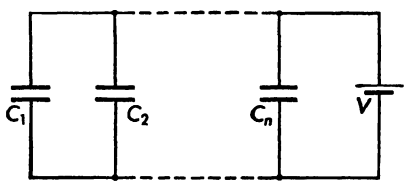


Fig. 3.3

Parallel Connection

The parallel arrangement is shown in Figure 3.3. Here the potential across the individual capacitors is a constant, equal to the battery voltage.

$$V = \frac{Q_1}{C_1} = \frac{Q_2}{C_2} = \cdots = \frac{Q_n}{C_n} \qquad [3\text{-}14]$$

If the arrangement is regarded as one large capacitor of capacitance C, the total charge Q is given by

$$Q = CV \qquad [3\text{-}15]$$

where $$Q = Q_1 + Q_2 + \cdots + Q_n \qquad [3\text{-}16]$$

whence $$CV = C_1V + C_2V + \cdots + C_nV \qquad [3\text{-}17]$$

and $$C = \sum_1^n C_i \qquad [3\text{-}18]$$

or, in words, *the effective capacitance of a number of capacitors connected in parallel is the sum of their separate capacitances.*

It is likely that the reader has at some time been introduced to the rules (§ 5.6) for combining resistors in series or in parallel. If so, he will note that the rules for combining resistors in parallel and capacitors in series are the same. Also, resistors in series are treated in the same manner as capacitors in parallel.

Problems 3.5 and 3.6 provide practice in applying the rules for combining capacitors.

3.6 Uses of Capacitors

Before discussing various types of capacitors it might be helpful in developing intuition concerning their action to describe briefly some of the practical uses to which they are put. Let us consider three important applications.

Smoothing Fluctuating Voltages

Presumably all readers are familiar with the fact that AC voltages may be rectified, or made unidirectional. One means of accomplishing this is to include a "metallic" *rectifier* (e.g., selenium or copper oxide) in one of the leads from the AC generator (Fig. 3.4a). The rectifier offers low resistance to current

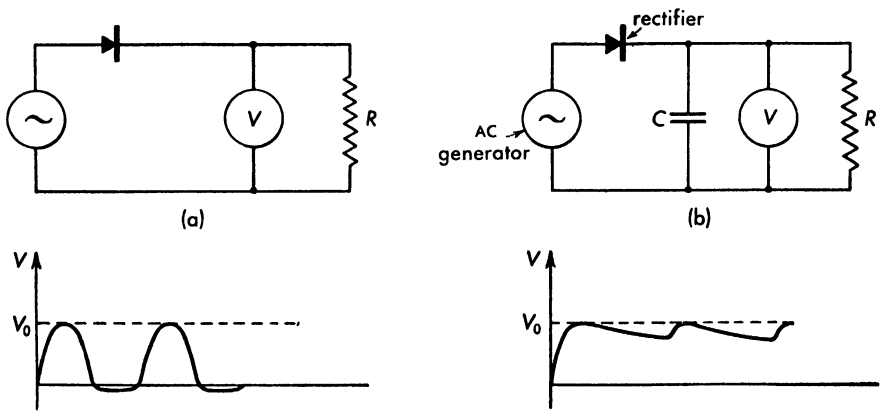


Fig. 3.4

flowing in one direction but high resistance to current flowing in the opposite direction. The voltage across the load resistor R is, for this reason, a pulsating one. For half the time, during the periods when the voltage across the rectifier is such as to send a current in the high-resistance direction, most of the voltage

drop appears across the rectifier and the voltage, V, across the load is negligibly small. During the other half-cycles the voltage drop across the rectifier is small and the voltage across the load is approximately equal to the generator voltage. The current during these half-cycles is large. The load voltage V varies with time as shown in the graph in Figure 3.4a; although it is DC in the sense that it does not change polarity (at least, not to any significant extent), it is anything but steady.

If, however, a capacitor is inserted in the circuit, as shown in Figure 3.4b, it becomes charged to the potential V_0 during the conducting half-cycle of the rectifier and discharges through the load resistor R during the "off" half-cycle. If R and C are large, this discharge takes place slowly, with the result that V does not appreciably drop before the capacitor is recharged, during the next "on" half-cycle, to the initial value V_0. In this way, by virtue of their ability to store charge, capacitors can materially reduce fluctuations in pulsating voltages: they are integral components of most rectifying arrangements.

Transmission of AC Signals

It is frequently necessary in a circuit to transmit an AC signal between two points that are maintained at different DC potentials and cannot, therefore, be directly connected. A capacitor provides the required DC isolation and at the same time allows the AC signal to be transmitted, as a consideration of Figure 3.5 will disclose.

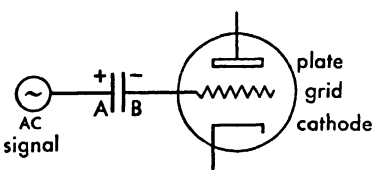


Fig. 3.5

Let us suppose that the signal from an AC generator is to be transmitted to the grid of a tube, there to be amplified. When the generator voltage is positive, the positive charge on plate A of the capacitor attracts an equal negative charge to plate B, leaving the grid deficient in electrons. The charge on the grid thus follows that on plate A. Although there is no actual transmission of charge through the capacitor, there is, in effect, a transmission of signal.

Selection of Signal for Wireless Reception

A radio or television set must contain the means of selecting one signal from the several that are available. This takes the form of a *tuned circuit*, shown in Figure 3.6, across which the current from the antenna can develop a voltage. A capacitor of capacitance C and an inductor of inductance L in parallel with it constitute such a tuned circuit. As we shall see in Chapter 13, this circuit offers

maximum impedance to the flow of AC current of a particular frequency, namely, $f_r = 1/2\pi\sqrt{LC}$. For signal frequencies higher or lower than this value, because of the low impedance offered by the LC-combination, the signal voltages at point A are negligible in comparison with those of frequency f_r. The tuned circuit thus selects the signal frequency f_r for amplification by the receiver. The frequency to which the receiver is receptive can be altered by varying the value of C. The variable capacitor is appropriately known as the "tuning" capacitor.

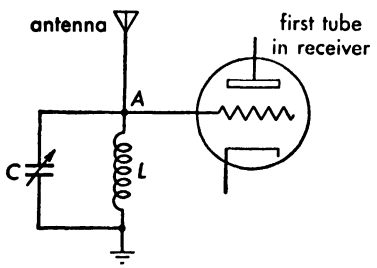


Fig. 3.6

3.7 The Parallel-Plate Capacitor

The parallel-plate capacitor (Fig. 3.7), is the most commonly used of the various capacitor types.

It follows from what has been said in § 1.7 that the field intensity between these capacitor plates has the constant value

$$E = \frac{\sigma}{\epsilon} \qquad [1\text{-}29]$$

The potential of plate P is, therefore,

$$V = Ed = \frac{\sigma d}{\epsilon} \qquad [3\text{-}19]$$

or

$$V = \frac{d}{\epsilon}\left(\frac{Q}{A}\right) \qquad [3\text{-}20]$$

where A m^2 is the plate area and Q is the total plate charge.

Comparing equation [3-20] with the general relationship $V = Q/C$, one finds that the capacitance of the plate P, in the presence of the grounded plate S,

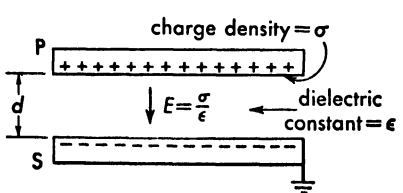


Fig. 3.7

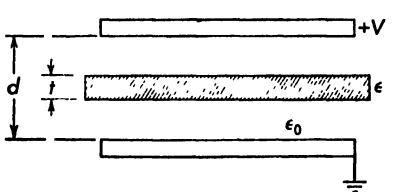


Fig. 3.8

is

$$C = \frac{\epsilon A}{d} \qquad [3\text{-}21]$$

or

$$C = \frac{k_e \epsilon_0 A}{d} \qquad [3\text{-}22]$$

This is said to be the capacitance of the two-plate combination.

The calculation of the capacitance of any capacitor follows the above plan. Before discussing further the particular case of parallel-plate capacitors, the steps in the calculation will be summarized. They are:

a. Calculate the potential difference between the capacitor plates, as in equation [3-19];

b. Write this potential difference V in terms of the net charge Q on one of the capacitor plates, as in equation [3-20];

c. Obtain the value of the capacitance from $C = Q/V$.

One concludes from equation [3-22] that *the capacitance of the parallel-plate capacitor varies* (1) *directly as the area of the plates*, (2) *directly as the dielectric coefficient of the medium between the plates, and* (3) *inversely as the separation of the plates.*

It can be seen from Table 1 that a wide variety of dielectric coefficients is available for the construction of capacitors. Of the time-honored dielectrics, mica ($7 < k_e < 8$) has the highest value of k_e, a fact which accounts for its widespread use in commercial capacitors. The recently developed titanium-dioxide ceramics ($15 < k_e < 500$) make it possible to achieve high values of capacitance in capacitors that are physically very small. This is even more true of the miniature capacitors that employ barium titanate as a dielectric ($500 < k_e < 6000$). The special characteristic of barium titanate that is responsible for its large values of dielectric coefficient is discussed in § 4.13.

Parallel-Plate Capacitor with Compound Dielectric

The capacitor shown in Figure 3.8, in which a portion of the interplate region is occupied by a dielectric slab of thickness t, the rest of the region being air-filled, provides an example to which the method, outlined above, for computing capacitance may be instructively applied. The calculation of the capacitance of this arrangement is the subject of Problem 3.10. The capacitance will be found to have the value:

$$C = \frac{\epsilon_0 A}{\left[\quad t\right]} \qquad [3\text{-}23]$$

Elimination of Edge Effects by Use of Guard Ring

We have so far assumed that the field between the plates of a parallel-plate capacitor is everywhere constant and drops sharply to zero at the edges of the plates. In practice this is not true, as the field begins to weaken near the edges of the capacitor plates and then actually bulges out beyond their physical boundary, as shown on the left of Figure 3.9. Because of these edge effects it is

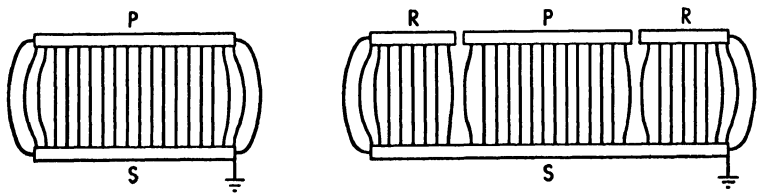


Fig. 3.9

difficult to calculate the capacitance of a parallel-plate capacitor from its physical dimensions. The effective area A in the expression for the capacitance, $C = \epsilon A/d$ (equation [3-21]), is not known. This difficulty was overcome by Lord Kelvin by introducing a *guard ring* for the upper plate, as shown on the right side of Figure 3.9. The lower plate is enlarged and the upper plate, here assumed to be circular, is surrounded by the guard ring R, which is in the same plane as P and almost touches it. As the guard ring is at the same potential as P, the combination $P + R$ may be regarded as constituting one plate of a parallel-plate capacitor of which S is the other plate. The edge effects now occur at the outer edge of the guard ring, while the field between P and S is practically uniform throughout. The capacitance of plate P is now accurately given by the theoretical value $C = \epsilon A/d$, where A is the area of P plus half the area of the gap between it and the guard ring. The guard ring thus makes it possible to construct capacitors whose *absolute* values are known with precision, and which, consequently, find important use as laboratory standards.

Kelvin employed a capacitor of the type described above to make absolute determinations of potential differences. In this work the central portion P of the upper capacitor plate was not fixed in position, but was attached to one arm of a balance. In this way the pull between the two capacitor plates could be weighed, and the difference in potential between them calculated therefrom. It can readily be shown that the total force of attraction between plates P and S is given by

$$F = \frac{\epsilon_0 V^2 A}{2d^2} \qquad [3\text{-}24]$$

where A is the effective area of plate P, and V is the potential difference between plates P and S.

Some Commonly Employed Parallel-Plate Capacitors

Two embodiments of the parallel-plate capacitor are shown in Figure 3.10. That on the left is a variable air capacitor (that is, a capacitor with air dielectric) of the type used as a frequency-tuning element in radio receivers. Alternate plates are connected together, one group being fixed in position while the other group can be rotated with respect to them in order to vary the capacitance. In Problem 3.11 it is shown that n such plates, each of area A and separated from adjacent plates by a distance d, provide a maximum capacitance of $(n - 1)\epsilon_0 A/d$ farads. This decreases as the movable plates are rotated out of coincidence with the stationary ones.

The construction of the common paper capacitor is shown on the right in Figure 3.10. A long, thin, parallel-plate capacitor, consisting of two sheets of

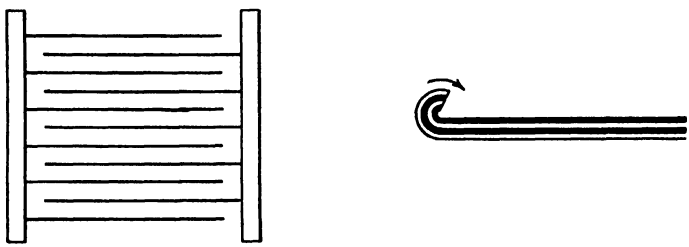


Fig. 3.10

metal foil separated by and covered with a thin layer of impregnated paper, is wound into a roll. This not only leads to a rugged, compact structure but approximately doubles the capacitance, as each side of the metal foil now serves as a capacitor plate. After winding, the unit is not literally a parallel-plate capacitor.

3.8 Spherical Capacitor

The spherical capacitor consists of one spherical metal shell concentric with and surrounding a smaller insulated one, as shown in Figure 3.11. As a rule, the outer spherical plate is grounded, and contact is made to the inner plate through a small hole in the outer one.

As before, the capacitance is computed by first calculating the potential difference between the two plates. The field at point P is caused entirely by the charge Q on the inner sphere and has the value

$$\boldsymbol{E} = \frac{Q}{4\pi\epsilon r^3}\boldsymbol{r} \qquad [3\text{-}25]$$

The potential difference between the two spheres is then:

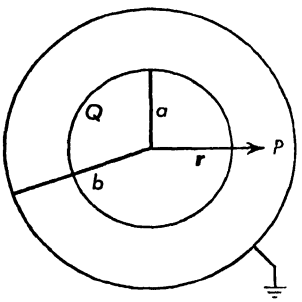


Fig. 3.11

$$V = -\int_b^a \boldsymbol{E}\cdot d\boldsymbol{r} = -\int_b^a \frac{Q}{4\pi\epsilon r^2}\,dr \qquad [3\text{-}26]$$

$$= \frac{Q}{4\pi\epsilon r}\Big]_b^a = \frac{Q}{4\pi\epsilon}\,\frac{(b-a)}{ab} \qquad [3\text{-}27]$$

from which

$$C = \frac{Q}{V} = \frac{4\pi\epsilon\, ab}{b-a} \qquad [3\text{-}28]$$

This type of capacitor is not of much practical use, although the fact that its capacitance can be accurately computed from its geometry has led to its occasional use as a standard.

3.9 Cylindrical Capacitor

The plates of the cylindrical capacitor, shown in Figure 3.12, are coaxial cylinders, the outer one necessarily being hollow. Let us suppose that the charge per unit length on the inner cylinder is λ. The field at the point P is that of the inner cylinder alone and is given by the expression for the field of a long charged wire:

$$\boldsymbol{E} = \frac{\lambda}{2\pi\epsilon r^2}\,\boldsymbol{r} \qquad [1\text{-}23]$$

The potential difference between the outer and inner cylinder is

$$V = -\int_b^a \boldsymbol{E}\cdot d\boldsymbol{r} = \frac{\lambda}{2\pi\epsilon}\int_a^b \frac{dr}{r} = \frac{\lambda}{2\pi\epsilon}\ln\frac{b}{a} \qquad [3\text{-}29]$$

The capacitance per unit length is then

$$C = \frac{Q}{V} = \frac{\lambda}{V} = \frac{2\pi\epsilon}{\ln\left(\frac{b}{a}\right)} \qquad [3\text{-}30]$$

This type of capacitor, in contrast with the spherical type, is of great practical interest. For example, coaxial cables consist of a cylindrical metal shield, a coaxial central conductor, and an interposed dielectric. These are widely used in the transmission of high-frequency signals, and in their use account must be taken of the

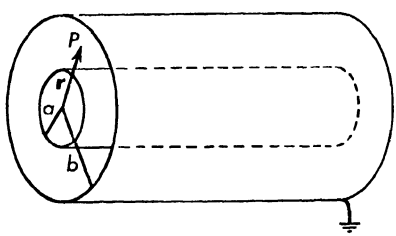


Fig. 3.12

capacitance which they introduce. Also, a submarine cable, consisting of an insulated single conductor, forms an effective cylindrical capacitor in which the salt water plays the role of the outer cylinder.

3.10 Comparison of Dielectric Constants

The capacitance of any capacitor is directly proportional to the dielectric constant of the material between the capacitor plates. Hence, the dielectric constants of two materials can be readily compared by introducing the materials, in turn, into a given capacitor and determining the resulting capacitances.

For a given material, the change in dielectric constant as a function of pressure, temperature, or some other variable can be measured with high precision by employing the material-filled capacitor as the capacitative element in a tuned circuit (Fig. 3.6). If the circuit is sharply resonant, a small change in the capacitance of the capacitor results in a significant change in the resonant frequency of the circuit. By this means, for example, even the small changes in the dielectric constants of gases which occur when the temperature is altered have been accurately studied. The manner in which the dielectric constant of a polar material changes with temperature is discussed in detail in §§ 4.11 and 4.12.

The variation of the dielectric constant of a material as the frequency of the applied field is changed is a subject of considerable interest. When a DC voltage is applied to a capacitor, the polar molecules in the dielectric orient themselves under the action of the electric field. When the applied voltage is an alternating one, the polar molecules again attempt to line up with the field and are, in fact, equally successful if the frequency of the AC voltage is low. As the polarity of the voltage changes, the polar molecules obligingly change their direction. When the frequency of the applied field is high, however, the polar molecules may not have time to orient themselves to the same extent before the polarity changes. For this reason, *in a material that possesses permanent polar molecules, the dielectric constant decreases with increasing frequency.* If, on the other hand, *the polar molecules in the dielectric are induced ones*, resulting from a displacement of the planetary electron systems (§§ 1.5 and 4.1), *there is no observed decrease with increasing frequency*, because this displacement is practically instantaneous.

In most materials, both permanent and induced polar molecules contribute to the polarization. For such materials the behavior of the dielectric coefficient as a function of the frequency of the applied field is shown in Figure 3.13. For ex-

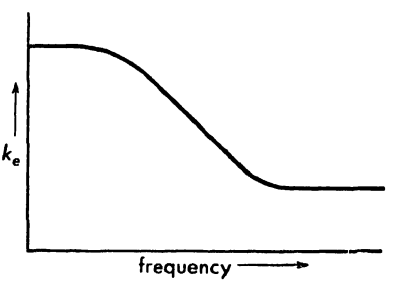


Fig. 3.13

ample, the dielectric coefficient of water falls from its low-frequency value of 80 to less than 2 at optical frequencies (10^{14}–10^{15}/sec). The observations of k_e at optical frequencies are, of course, not made within a capacitor but by optical methods (§ 4.9).

3.11 Electrostatic Voltmeters

Voltmeter is the name given to an instrument that measures the difference in potential between two points. Most voltmeters require for their operation a steady passage of charge. This, as we shall have occasion to observe in Chapter 6, is usually a disadvantage, and pains are taken to construct voltmeters which operate with as small a flow of charge as possible. Two types of voltmeter exist, however, which require an initial flow of charge only, and that a relatively small one. These are the *vacuum-tube voltmeter*, discussed in § 6.8, and the *electrostatic voltmeter*, discussed below.

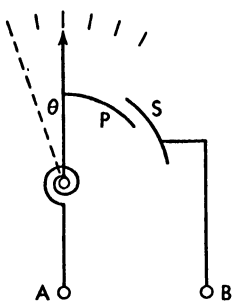


Fig. 3.14

One form of the electrostatic voltmeter is shown schematically in Figure 3.14. Plates P and S constitute a capacitor whose capacitance increases as the needle moves to the right. The motion of the needle is opposed by a spring which also serves the purpose of providing electrical contact between terminal A and plate P. When A and B are attached to points that are at different potentials, the attraction between plates P and S causes the needle to move to the right until the torque arising from the electrical attraction is equal to the restoring torque provided by the spring.

Let us first consider the case where the charges ($\pm Q$) on plates P and S remain constant. This would happen if the external connections to A and B were made only briefly, and were then broken before any motion of the needle had occurred.

As we have seen in § 3.3, the energy of the capacitor formed by plates P and S is given by

$$U = \frac{Q^2}{2C} \qquad [3\text{-}6]$$

Plate P moves so as to reduce this energy. The amount of the electrical torque is given by

$$T = -\frac{\partial U}{\partial \theta} = \frac{Q^2}{2C^2}\frac{dC}{d\theta} \qquad [3\text{-}31]$$

which becomes, upon substitution of $V = Q/C$,

$$T = \frac{V^2}{2}\frac{dC}{d\theta} \qquad [3\text{-}32]$$

In the equilibrium position, this electrostatic torque is equal to the restoring torque, $\tau\theta$, of the spring (τ is the torsional constant of the spring). That is,

$$\frac{V^2}{2}\frac{dC}{d\theta} = \tau\theta$$

or

$$\theta = \frac{V^2}{2\tau}\frac{dC}{d\theta} \qquad [3\text{-}33]$$

Thus, if $dC/d\theta$ were constant, the deflection of the needle would be proportional to the square of the difference in potential between A and B. In practice, the value of $dC/d\theta$ is somewhat dependent upon θ, and the instrument must therefore be calibrated by the manufacturer.

It is important to note that the electrostatic voltmeter (since its deflection depends upon the *square* of the applied voltage) is equally useful for AC and DC voltages. In the case of AC voltages, the remark made above, that only an initial flow of charge occurs, is not applicable.

■ *Source of the Energy Expended in Deflecting the Electrostatic Voltmeter*

In the case described above, where the charge on the capacitor plates remains constant, the work to deflect the needle against the restoring torque of the spring is done at the expense of the electrostatic energy of the system. When the instrument is in actual use, however, the external connections to A and B are maintained until the needle takes up its final position. Thus, it is the potential difference between the plates, rather than the charge upon them, that remains constant. Under these conditions, although the electrostatic torque is still given by equation [3-32], the energy expended in deflecting the needle and, indeed, additional energy as well, is supplied by the source of potential that is attached to terminals A and B. We shall now analyze this case.

In general, any change in potential across the electrostatic voltmeter is given by

$$dV = d\left(\frac{Q}{C}\right) = -\frac{Q}{C^2}dC + \frac{dQ}{C} \qquad [3\text{-}34]$$

In the situation under study the potential remains constant, and therefore,

$$\frac{dQ}{C} = \frac{Q}{C^2}dC \qquad [3\text{-}35]$$

Equation [3-34] suggests that the operation may be visualized as taking place in two stages. In the first stage, represented by the first term on the right of equation [3-34], the charge remains constant, while the capacitance increases (plate P moves to the right). This results in a lowering of the potential. In the second stage, represented by the

second term on the right of equation [3-34], plate P remains motionless while charge is added to it to raise the potential to its original value.

The electrostatic energy of the system is altered during these two stages by the following amount:

$$dU = d\left(\frac{Q^2}{2C}\right) = -\frac{Q^2}{2C^2}dC + \frac{Q}{C}dQ \qquad [3\text{-}36]$$

where the terms on the right represent the energy changes during the first and second stages, respectively. It will be noted that the first term describes the case discussed earlier, in which the charge on the capacitor remains constant. As mentioned in that discussion, the electrical energy of the voltmeter system decreases during this operation, as mechanical work is done. The amount of mechanical work done is, therefore:

$$dW = \frac{Q^2}{2C^2}dC \qquad [3\text{-}37]$$

When equation [3-35] is substituted into [3-36], the total change in the electrostatic energy of the constant-potential system is seen to be

$$dU = -\frac{Q^2}{2C^2}dC + \frac{Q^2}{C^2}dC = \frac{Q^2}{2C^2}dC \qquad [3\text{-}38]$$

Evidently, when a constant potential is applied to an electrostatic voltmeter, the electrostatic energy increases by an amount equal to the mechanical work done. The source of potential, therefore, not only supplies the energy needed to deflect the needle but, in addition, contributes an equal amount of electrostatic energy to the voltmeter system.

Since the mechanical work is given by the product of the torque and the angular displacement, we conclude that

$$dW = T\,d\theta = dU \qquad [3\text{-}39]$$

whence

$$T = \frac{\partial U}{\partial \theta} \qquad [3\text{-}40]$$

This expression for the torque, it will be noted, differs in sign from equation [3-31]. In the case of constant charge, $T = -\partial U/\partial\theta$, whereas in the case of constant potential, $T = \partial U/\partial\theta$. In each case the torque has the same value, viz.,

$$T = \frac{V^2}{2}\frac{dC}{d\theta} \qquad [3\text{-}32]$$

This is to be expected since the torque arises from the attraction between the charges on plates P and S and depends only upon the relative positions of the plates and the charges they contain. It is a matter of indifference as to how the existing state of affairs was reached.

■ *Force on a Dielectric in an Electric Field*

The introduction of a dielectric, as in Figure 3.8, increases the capacitance of a capacitor and, consequently, affects its electrostatic energy. If the dielectric slab is

partially withdrawn, it will experience, in keeping with the considerations elaborated above, a restoring force. This matter is further developed in Problem 3.21.

Thus, both conductors (§ 1.11) and dielectrics experience forces which urge them from the weaker to the stronger parts of an electric field.

3.12 Problems

E (3.1) Show that when two charged conductors are connected together the total charge divides between them in the ratio of their capacitances.

E (3.2) Show that, on the basis of energy considerations (§ 3.3), the charge on a conducting sphere will always be found on the surface.

E (3.3) Using equation [1-61] for the energy density of space pervaded by an electric field, derive the expression for the energy stored in the space between the plates of a parallel-plate capacitor. Compare this value with that for the energy of a charged capacitor, i.e., equations [3-6], [3-7], and [3-8].

E (3.4) A metal sphere of 20 pf capacitance is charged to a potential of 100 V. Find its energy. It is then joined by a long thin wire to an uncharged metal sphere of similar size. Find the energy of the system. Account for the loss in energy.

E (3.5) If several 1 μf capacitors are available, each capable of withstanding 200 V, how would you construct (a) a unit possessing a capacitance of 0.2 μf and (b) a unit possessing a capacitance of 0.6 μf, each capable of withstanding 1000 V?

E (3.6) Figure 3.15 shows a form of variable capacitor that is in common use. Various values of capacitance are obtained between the terminals A and B by inserting metal plugs in some of the spaces 1–12. For example, plugs in spaces 1 and 8, the other spaces remaining empty, give a capacitance of 0.05 μf. In which spaces should plugs be inserted to obtain a capacitance of (a) 1.0 μf and (b) 0.15 μf?

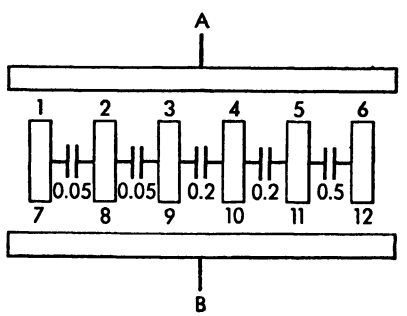
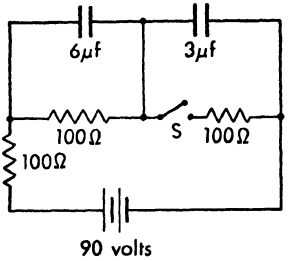


Fig. 3.15

E (3.7) Calculate the charge on each capacitor and the potential difference across it in the circuit shown in Figure 3.16, on the left, for the two cases (a) switch S closed and (b) switch S open.



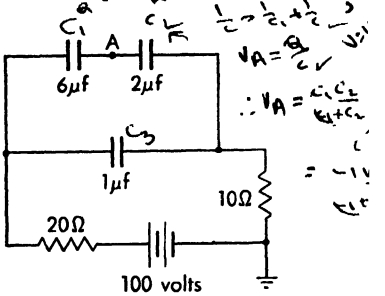


Fig. 3.16

E (3.8) Calculate the potential of point A in the circuit shown in Figure 3.16 on the right. [Ans.: 75 volts.]

E (3.9) One plate of a parallel-plate capacitor is grounded and the other is charged with Q coulombs of electricity. If the plates each have an area of A m^2 and the plate separation is d meters, what is the potential difference between them? How should the separation be changed in order to increase their potential? This method was actually used by Volta (1745–1827), Italian pioneer in the study of electricity, to detect the small difference in potential provided by the first voltaic cell.

M (3.10) Show that the capacitance of the capacitor with compound dielectric, pictured in Figure 3.8, is given by equation [3-23].

E (3.11) Show that a pile of n plates, each of area A m^2 and separated from one another by a dielectric of thickness d meters and dielectric constant ϵ, constitutes a capacitor of capacitance $(n - 1)\epsilon A/d$ farads. Alternate plates are connected together.

M (3.12) The separation of the plates in a parallel-plate air capacitor is 0.75 cm. The plates are connected to a battery of voltage V. What is the effect on the attraction between the plates when a slab of dielectric, $k_e = 5$ and thickness 0.5 cm, is inserted between them, as in Figure 3.8? [Ans.: 4.5 times greater.]

M (3.13) Suppose, in Problem 3.12, that the battery had been disconnected before inserting the dielectric slab. What would then have been the effect on the attraction between the plates?

M (3.14) What is the capacitance of the spherical capacitor of Figure 3.11 if the inner plate is grounded and the outer sphere is given the charge Q? What is the disadvantage of this arrangement?

M (3.15) In the arrangement suggested in the preceding problem, what charge is induced upon the inner sphere?

E (3.16) Show that expression [3-28] for the capacitance of a spherical capacitor approaches the parallel-plate value as a approaches b.

M (3.17) A capacitor is formed from two long coaxial cylinders of radii a and b, respectively (Fig. 3.12). The space between $r = a$ and $r = c$, where $a < c < b$, is filled with material of dielectric coefficient k_e, the remaining space between the cylinders being air-filled. If the outer cylinder is grounded, show that the capacitance per unit length of the arrangement is:

$$C = 2\pi\epsilon_0 \left[\frac{1}{k_e} \ln \frac{c}{a} + \ln \frac{b}{c} \right]^{-1}$$

H (3.18) Show that the capacitance of two long and oppositely-charged wires, each of radius a, is

$$C = \frac{\pi\epsilon}{\ln\left(\dfrac{d-a}{a}\right)}$$

where d ($> 2a$) is the distance between axes of the wires. This result assumes that the charge on each wire is uniformly distributed. As suggested in the concluding paragraph of § 2.11, this is not a correct assumption. The error introduced by it, however, is small if the wires are far apart.

M (3.19) Shown in Figure 3.17 is a schematic diagram of the electron gun and vertical-deflecting system of a cathode ray oscilloscope. Calculate the vertical component of the electron's velocity as the electron emerges from the parallel-plate capacitor. Also calculate the deflection d.

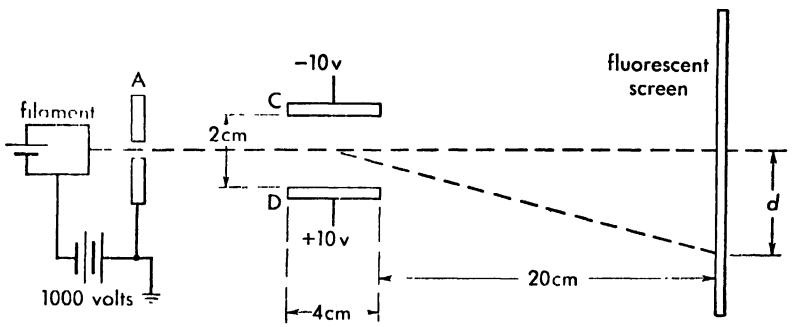


Fig. 3.17

H (3.20) How would the deflection d be altered if the potential of plate C were -20 V and plate D were grounded?

M (3.21) (Based on small-type material in § 3.11.) The plates of the parallel-plate capacitor shown in Figure 3.8 have a length l and a width w. If the dielectric slab (also of width w) is withdrawn in the direction of its length until it extends only a distance x within the plates, what force, if any, is exerted upon it? The potential between the plates is maintained constant.

Chapter 4

Dielectric Theory and Behavior

Substances can be classified from the viewpoint of their electrical behavior into one of three groups: *conductors*, *semiconductors*, and *insulators* (or *dielectrics*). The reasons for the differences in behavior of these three classes of material are well understood and will be discussed in Chapter 6. For the purpose of this chapter, in which we are concerned with insulators or dielectrics, it is sufficient to agree that an insulator is a material that will not permit the *passage* of electric charge, although local microscopic *displacements* of charge may take place within it.

4.1 Electric Dipoles: Permanent and Induced

We have already noted in §§ 1.5 and 3.10 that permanent polar molecules exist in certain dielectrics and that, in addition, *all* atoms or molecules can be temporarily polarized under the action of an electric field.

An example of a permanent polar molecule is provided by HCl. Here each of the atoms is one planetary electron short of a complete shell.[1] By a mutual sharing of two electrons, as shown in Figure 4.1, each atom achieves a reasonable facsimile of a closed shell configuration.

[1] The reader will recall, probably from studies in chemistry, that the planetary electrons in an atom are arranged in shells. Two electrons complete the first shell, eight comprise the second, eight more complete an additional subshell, and so on, each closed shell configuration representing a conspicuously stable arrangement. Since the first shell constitutes the lowest energy state, it is filled in first, followed by the second and third. In hydrogen ($Z = 1$) the first shell is filled but for one electron, while in chlorine ($Z = 17$) an additional electron is needed to complete a subshell.

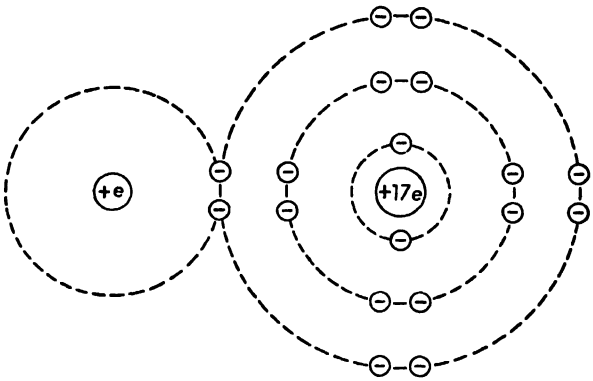


Fig. 4.1

This sharing of electrons, known as a *covalent bond*, produces an unsymmetric charge distribution in which there is a net positive charge at the hydrogen end of the molecule and a net negative charge at the chlorine end. To the outside observer, such a molecule is equivalent to the *dipole* shown in Figure 4.2.

A dipole, by definition, consists of *two equal and opposite charges*, $+q$ and $-q$, which are *separated by a* vector *distance* $\boldsymbol{d}$. The product

$$\boldsymbol{p} = q\boldsymbol{d} \qquad [4\text{-}1]$$

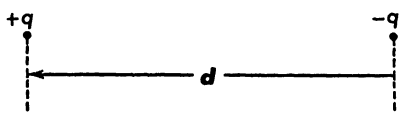


Fig. 4.2

known as the *dipole moment*, is a vector quantity pointing from the negative to the positive charge. All permanent polar molecules possess electric dipole moments by virtue of the fact that the "center of gravity" of the negative charge in the molecule is displaced from the center of gravity of the positive charge.

The production of *induced dipole moments* is illustrated schematically in Figure 4.3. In a field-free environment, as shown on the left, the charge dis-

Fig. 4.3

tribution in an atom of helium is entirely symmetric, and the atom consequently possesses zero dipole moment. In an electric field, however, as shown on the right, the center of gravity of the planetary electron system is displaced from

A to B, a distance d from the center of gravity (the nucleus) of the positive charge system. Effectively, there is now a charge of $+2e$ at A and a charge of $-2e$ at B, the entire system possessing a dipole moment

$$\boldsymbol{p} = 2e\boldsymbol{d} \qquad [4\text{-}2]$$

The interest in dipoles stems from the fact that they give rise to electric fields. Evidently, if two equal and opposite charges are located at the same point, the field of the one is exactly canceled by that of the other and there is zero external field. When, however, one of the charges is displaced with respect to the other, as in a dipole, the fields of the two charges no longer completely neutralize one another. The magnitude of the resulting field is derived in the following section.

4.2 Potential and Field Arising from an Electric Dipole

Let us calculate, with the aid of Figure 4.4, the potential and the field of an electric dipole of moment $\boldsymbol{p} = q\boldsymbol{d}$. First, the potential will be evaluated at the point P, whose polar coordinates are (r,θ) with respect to the mid-point and axis of the dipole. It is assumed that r is large compared to the length d of the dipole.

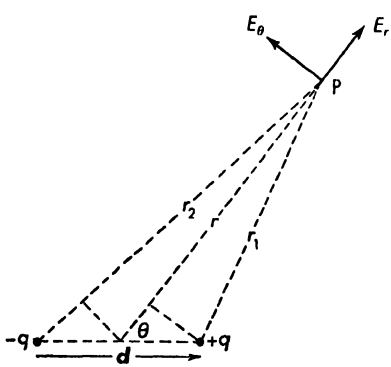


Fig. 4.4

The potential at P, arising from the two point charges q and $-q$, is given by equation [2-40]. Here,

$$V(P) = \frac{1}{4\pi\epsilon}\left[\frac{q}{r_1} - \frac{q}{r_2}\right] \qquad [4\text{-}3]$$

or

$$V(P) = \frac{q}{4\pi\epsilon}\left[\frac{r_2 - r_1}{r_1 r_2}\right] \qquad [4\text{-}4]$$

Because $r \gg d$, we can make two useful approximations. First, we may write

$$r_1 \doteq r - \frac{d}{2}\cos\theta$$

and

$$r_2 \doteq r + \frac{d}{2}\cos\theta \qquad [4\text{-}5]$$

in which case equation [4-4] takes the form

$$V(P) \doteq \frac{qd\cos\theta}{4\pi\epsilon\left[r^2 - \frac{d^2}{4}\cos^2\theta\right]} \qquad [4\text{-}6]$$

Second, we can neglect terms in d^2 in comparison to those in r^2. Equation [4-6] then becomes

$$V(P) \doteq \frac{qd\cos\theta}{4\pi\epsilon r^2} = \frac{p\cos\theta}{4\pi\epsilon r^2} = \frac{\boldsymbol{p}\cdot\boldsymbol{r}}{4\pi\epsilon r^3} \qquad [4\text{-}7]$$

where $\boldsymbol{p}$ $(= q\boldsymbol{d})$, and $\boldsymbol{r}$ extends from the center of the dipole to the point P.[8]

The radial and tangential components of the field at P are now obtained from the potential by differentiation (see Problem 2.4):

$$E_r = -\frac{\partial V}{\partial r} \doteq \frac{2p\cos\theta}{4\pi\epsilon r^3} \qquad [4\text{-}8]$$

and

$$E_\theta = -\frac{1}{r}\frac{\partial V}{\partial \theta} \doteq \frac{p\sin\theta}{4\pi\epsilon r^3} \qquad [4\text{-}9]$$

Evidently $E_\theta = 0$ along the axis of the dipole ($\theta = 0$), while $E_r = 0$ along the right bisector of the dipole ($\theta = \pi/2$).[9]

4.3 Quadrupole and Higher Moments

It is not uncommon to have a charge configuration for which the dipole moment is zero, yet which gives rise to a potential and a field. Such a one is shown in Figure 4.5. The total dipole moment is the vector sum of two separate moments, each of magnitude qd, which are antiparallel to one another. It is therefore zero.

Let us calculate the potential of this charge configuration along its axis, that is, at the point P shown in Figure 4.5.

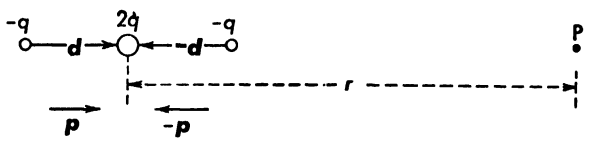


Fig. 4.5

[8] The reader may readily demonstrate to himself that $\nabla(1/r) = -\boldsymbol{r}/r^3$, where $\boldsymbol{r}$ is the radius vector $i\boldsymbol{x} + j\boldsymbol{y} + k\boldsymbol{z}$. In the light of this relationship, equation [4-7] may be written in the alternative form

$$V(P) = -\frac{1}{4\pi\epsilon}\boldsymbol{p}\cdot\nabla\left(\frac{1}{r}\right)$$

[9] Equations [4-7], [4-8], and [4-9] are usually written as exact equations with the understanding that $(r/d) \longrightarrow \infty$.

$$V(P) = \frac{2q}{4\pi\epsilon r} - \frac{q}{4\pi\epsilon(r - d)} - \frac{q}{4\pi\epsilon(r + d)} \qquad [4\text{-}10]$$

$$= \frac{-2qd^2}{4\pi\epsilon(r^3 - rd^2)} \qquad [4\text{-}11]$$

Assuming, as before, that $r \gg d$, we may rewrite [4-11] as:

$$V(P) \doteq \frac{-qd^2}{2\pi\epsilon r^3} \qquad [4\text{-}12][4]$$

Here the potential varies as r^{-3}, in contrast with the r^{-2} variation found for a dipole (equation [4-7]) and the r^{-1} variation characteristic of a point charge or *monopole* (equation [2-39]). Any charge configuration that gives rise to a potential that varies as r^{-3} is said to be an electric *quadrupole*. Likewise, if $V \propto r^{-4}$, the charge configuration is an *octupole*, and so forth.

If, in discussing the dipole in the previous section, we had written the *exact* expression for the potential at $P(r,\theta)$, we would have had an infinite series of terms, the first (given by equation [4-7]) containing r^{-2}, the second containing r^{-3}, and so on. That is, the two charges shown in Figure 4.4 constitute not only a dipole, but also a quadrupole, an octupole, etc. Similarly, the charge configuration in Figure 4.5 gives rise to an infinite number of terms in the expression for the potential at P. The first of these is given by equation [4-12].

In these infinite series, for large values of r, the subsequent terms are always negligible in comparison to the first one. They are ignored, therefore, as has been done above, and the charge configuration derives its name from the dependence upon r of the leading term in the expression for the potential.

4.4 Electric Polarization and Electric Susceptibility

We have already referred, in § 1.5, to a dielectric in which polar molecules have been induced, or in which permanent polar molecules have become aligned under the action of an external electric field, as being *polarized*. The extent to which these effects take place is described by the term *electric polarization*, $\boldsymbol{P}$, defined as the *electric moment per unit volume*. Thus,

$$\boldsymbol{P} = \frac{d\boldsymbol{p}}{dv} \qquad [4\text{-}13]$$

Let us consider, as an example of polarization, the dielectric that is between the plates of a charged parallel-plate capacitor (Fig. 4.6). Throughout

[4] It is emphasized that this is not a general result but holds only for points located on the axis. At non-axial points, however, V still varies as r^{-3}.

the body of the dielectric the charges on adjacent ends of polar molecules neutralize one another. At both the upper and lower boundaries of the dielectric, however, ends of polar molecules are exposed. Let the surface density of these exposed charges be σ_P. Consider specifically a tube of the dielectric, of cross-sectional area dA, extending from one capacitor plate to the other. Because the charges at the two ends are $-\sigma_P\, dA$ and $+\sigma_P\, dA$, respectively, the dipole moment of this tube of dielectric is

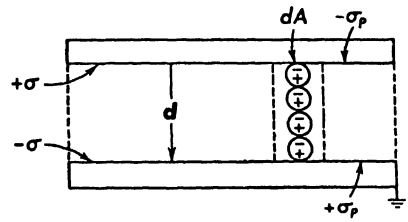


Fig. 4.6

$$d\boldsymbol{p} = (\sigma_P\, dA)\boldsymbol{d} \qquad [4\text{-}14]$$

Consequently, in keeping with the definition of polarization,

$$\boldsymbol{P} = \frac{d\boldsymbol{p}}{dv} = \frac{d\boldsymbol{p}}{\text{(volume of tube of dielectric)}} = \frac{(\sigma_P\, dA)\boldsymbol{d}}{(dA)d} \qquad [4\text{-}15]$$

or

$$P = \sigma_P \qquad [4\text{-}16]$$

This presupposes, of course, that polarization throughout the dielectric is uniform, in which case the only unneutralized charges are those located at the surfaces of the dielectric. The *polarization* is then *equal in magnitude to the charge per unit area on the surface of the dielectric.*

For isotropic dielectrics (those whose electrical properties are identical in all directions), the polarization that occurs under the action of an electric field has the same direction as the field. Also, the degree of polarization is a function of the strength of the field. This dependence is expressed as

$$\boldsymbol{P} = \chi_e \epsilon_0 \boldsymbol{E} \qquad [4\text{-}17]$$

where χ_e, the *electric susceptibility* of the material, is a dimensionless quantity. For most isotropic dielectrics χ_e is constant except for very strong fields.

For anisotropic dielectrics, $\boldsymbol{P}$ and $\boldsymbol{E}$ are normally not parallel. Although their relationship is still expressed by equation [4-17], the susceptibility is no longer a scalar, as for isotropic materials, but is instead a *tensor.*

Electrets

The electret is an interesting example of an isotropic dielectric material whose susceptibility is not constant. These materials, which include a group of organic waxes, can display permanent macroscopic polarization. When such a wax (e.g., carnauba wax) is subjected, in the molten state, to a strong electric field (say, by pouring it into the region between the plates of a charged

parallel-plate capacitor), the permanent polar molecules of which it is composed align themselves with the field. When the wax is allowed to solidify *in situ*, it is found to be permanently polarized.[5]

Electrets may be used as substitutes for batteries in certain applications requiring a source of potential but drawing no current. They were utilized in this way by the Japanese during the last war. For further details of the behavior of electrets, which can be much more complicated than the foregoing remarks would suggest, the reader is referred to a review article by F. Gutman (*Reviews of Modern Physics*, **20**, 457, 1948).

4.5 Electric Displacement

Before the introduction of the dielectric into the space between the plates of the capacitor of Figure 4.6, the field strength was (from [1-29]):

$$E = \sigma/\epsilon_0$$

where $\pm\sigma$ are the surface densities of "free" charges on the plates. The introduction of the dielectric causes the field to decrease to the value

$$E = \frac{1}{\epsilon_0}(\sigma - \sigma_P) \qquad [4\text{-}18]$$

reflecting the fact that σ_P coulombs/m^2 of the free charges on the plates are now neutralized by the polarization charges on the surface of the dielectric. Rewriting equation [4-18], we have

$$\sigma = \epsilon_0 E + \sigma_P = \epsilon_0 E + P \qquad [4\text{-}19]$$

The left-hand side of this equation, and consequently the right-hand side as well, depends only on the density of free charges on the capacitor plates. The right-hand side is defined as the *electric displacement*, D. Thus

$$D = \epsilon_0 E + P$$

or

$$\boldsymbol{D} = \epsilon_0 \boldsymbol{E} + \boldsymbol{P} \qquad [4\text{-}20]$$

Since the value of D depends only upon the density of free charges, it is not altered by the introduction of the dielectric. For this reason, it is a particularly useful quantity. It is directly related to E in a manner now to be shown.

An alternative expression to equation [4-18] for the intensity between the plates of the dielectric-filled capacitor is given by equation [1-29], namely,

---

[5] Another group of materials (the ferroelectrics), which can retain electric polarization when the polarizing field is removed, is discussed in § 4.13.

$$E = \frac{\sigma}{\epsilon} \qquad [1\text{-}29]$$

Recognizing from equations [4-19] and [4-20] that $D = \sigma$, we see that

$$D = \epsilon E$$

or

$$\boldsymbol{D} = \epsilon \boldsymbol{E} \qquad [4\text{-}21]$$

Thus, the factor of proportionality relating electric displacement and electric field strength is simply the dielectric constant of the medium.

Evidently, it also follows from equations [4-17], [4-20], and [4-21] that

$$\epsilon \boldsymbol{E} = \epsilon_0 \boldsymbol{E} + \boldsymbol{P} \qquad [4\text{-}22]$$

$$= (\epsilon_0 + \chi_e \epsilon_0)\boldsymbol{E}$$

whence

$$\epsilon = \epsilon_0(1 + \chi_e) \qquad [4\text{-}23]$$

relating the susceptibility of the medium, the dielectric constant of the medium, and the permittivity (or "dielectric constant") of free space. Equation [4-23] may also be written in terms of the dielectric coefficient, k_e:

$$k_e = \frac{\epsilon}{\epsilon_0} = 1 + \chi_e \qquad [4\text{-}24]$$

It is emphasized that the preceding relationships only apply to isotropic dielectrics. In such materials $\boldsymbol{D}$, $\boldsymbol{E}$, and $\boldsymbol{P}$ all have the same direction, and the values of ϵ and χ_e are usually independent of the strength of the field. For *anisotropic dielectrics*, such as quartz, where the electrical properties depend upon direction, $\boldsymbol{D}$, $\boldsymbol{E}$, and $\boldsymbol{P}$ are rarely parallel. Anisotropic dielectrics will not be discussed further.

Gauss' Theorem and Maxwell's First Law in Terms of Electric Displacement

The fact that $\boldsymbol{D} = \epsilon \boldsymbol{E}$ enables us to rewrite Gauss' theorem in terms of electric displacement. Equation [2-16] becomes

$$\int_s \boldsymbol{D} \cdot d\boldsymbol{s} = \int_v \rho \, dv \qquad [4\text{-}25]$$

where ρ is the volume density of free charge.[6] Likewise, the differential form of Maxwell's first equation (equation [2-19]) may be rewritten:

$$\nabla \cdot \boldsymbol{D} = \rho \qquad [4\text{-}26]$$

It should also be noted, in passing, that equation [1-61], giving the energy content of space that is pervaded by an electric field, can now be succinctly

[6] This follows since the value of D is independent of the presence of polarization charges (see preceding sub-section).

expressed as

$$\text{Energy density} = \frac{DE}{2} \qquad [4\text{-}27]$$

Flux and Flux Density

The left-hand side of equation [4-25] provides a useful interpretation of $\boldsymbol{D}$. Shown in Figure 4.7 are the two vectors $d\boldsymbol{s}$ and $\boldsymbol{D}$, the latter representing something which "passes through" the surface. The product $\boldsymbol{D}\cdot d\boldsymbol{s}$ is the product of the normal component of $\boldsymbol{D}$ and the area of the surface element. The integral of this expression over a closed surface is defined as the *electric flux*, Φ_e, emerging from the surface. That is,

$$\Phi_e = \int_s \boldsymbol{D}\cdot d\boldsymbol{s} = \int_v \rho \, dv \qquad [4\text{-}28]$$

In the light of this definition, $\boldsymbol{D}$ is seen to represent a *flux density*. It can be seen from equation [4-28] that the mks unit of Φ_e is the coulomb, and that of $\boldsymbol{D}$ is the coulomb/m^2. Gauss' theorem can now be stated in the following words: the *total electric flux emerging from any closed surface is equal to the net positive charge enclosed within the surface.* If the net charge enclosed within the surface is negative, Φ_e is also negative, indicating that the flux *enters* the surface.

In terms of lines of force, one can picture $\boldsymbol{D}$ as the density of lines of force, a quantity which is independent of the nature of the dielectric medium. Then $d\Phi_e = \boldsymbol{D}\cdot d\boldsymbol{s}$ is the product of the number of lines of force per unit area and the component of $d\boldsymbol{s}$ in the plane normal to the direction of the lines of force.

The use of the term "electric flux" is not fashionable nowadays, and likewise the interpretation of $\boldsymbol{D}$ as a flux density is frequently not presented. Notwithstanding, many readers will discover that the mental picture provided by the notion of flux is a helpful one, particularly in view of the fact that magnetic flux density, to be studied later, is a highly useful quantity. The interpretation of $\boldsymbol{D}$ which gives rise to the name "displacement" is described in the following section.

4.6 Maxwell's Displacement Current

Let us now, in the light of the concepts that we have been developing, consider the current that flows in the circuit of Figure 4.8 immediately after the switch is closed.

Evidently, a current flows through the conducting section of the circuit, resulting in an accumulation of positive charge on one plate of the capacitor

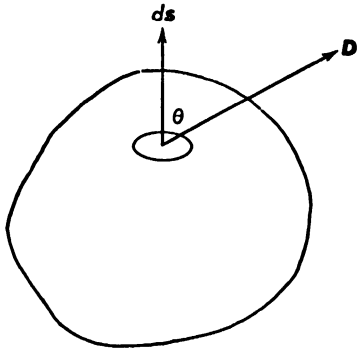


Fig. 4.7

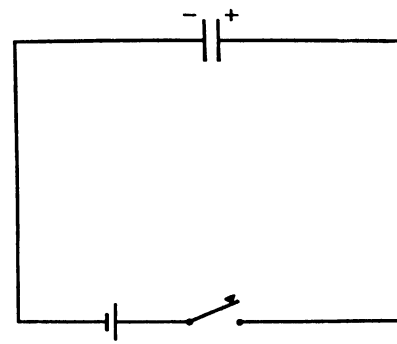


Fig. 4.8

and an equal accumulation of negative charge on the other plate. This is the usual conduction current with which we are familiar. Maxwell showed that it was necessary to assume that a current of the same value also flowed in the dielectric gap.

For a given charge density σ on the capacitor plates, the electric field in the dielectric gap is given by:

$$E = \frac{\sigma}{\epsilon} \qquad [1\text{-}29]$$

whence

$$\sigma = \epsilon E = D \qquad [4\text{-}29]$$

During the charging of the capacitor, the charge density on the plates is changing at a rate $d\sigma/dt$, and the value of the displacement in the dielectric gap changes accordingly:

$$\frac{d\sigma}{dt} = \frac{dD}{dt} \qquad [4\text{-}30]$$

or

$$\frac{d\sigma}{dt} = \epsilon_0 \frac{dE}{dt} + \frac{dP}{dt} \qquad [4\text{-}31]$$

The right-hand terms in equations [4-30] and [4-31] are descriptive of the dielectric gap and have (because the left-hand term has) the dimensions of current per unit cross-sectional area.[7] Let us examine, in equation [4-31], the two components of this current:

dP/dt is the rate of change of polarization. This is associated with the actual motion of charge in the dielectric, corresponding as it does to the rotation of permanent dipoles, or the formation of induced ones. In short, it arises from a "displacement" of

[7] The electric current (§ 5.1) through any cross-sectional area is the net charge that is transferred through it per second.

charge in the dielectric and possesses a current character that is easily recognized.

$\epsilon_0 \, dE/dt$ represents the current that is associated with a change in the electric field strength, even when there is a vacuum between the capacitor plates. It does not correspond to a movement of charge and is therefore difficult to interpret physically. However, regardless of the difficulty of interpretation, something occurring in the dielectric space is equivalent to a current of this magnitude. At least, as we shall see later, there are magnetic effects associated with $\epsilon_0 dE/dt$ that are identical with those which would arise from the passage of actual current of this amount.

The sum of these two current components, namely, dD/dt (equation [4-30]), was called by Maxwell the *displacement current*, the name obviously stemming from the physical effect associated with the dP/dt component. It was in this context that D, the *electric displacement*, was introduced. The *total* current in the gap, $A \, dD/dt$ (where A is the area of the capacitor plates), is thus equal to $A \, d\sigma/dt$, or dq/dt, the current flowing elsewhere in the circuit.

This recognition that the electric current is constant in all parts of a circuit, even in the space between the plates of a capacitor, was one of Maxwell's major contributions to physics. Without it, he could not have developed his electromagnetic theory of light.

4.7 Boundary Conditions at the Interface Between Two Dielectric Materials

In Figure 4.9, the line AB lies in the boundary plane between two isotropic dielectric media that have dielectric constants ϵ_1 and ϵ_2, respectively. The intensities and displacements in the two media are given, respectively, by $\boldsymbol{E}_1$, $\boldsymbol{D}_1$ and $\boldsymbol{E}_2$, $\boldsymbol{D}_2$.

In order to study the behavior of $\boldsymbol{D}$ at the boundary between the two media, let us consider the pillbox shown on the left of Figure 4.9. This pillbox, of

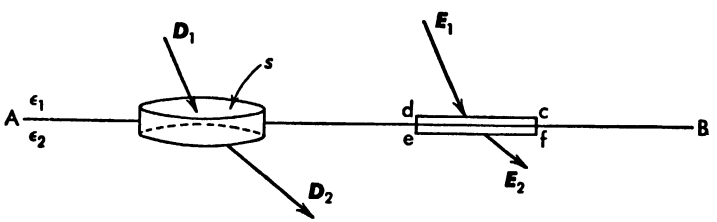


Fig. 4.9

negligible height, encloses a portion of the interface of area s. Applying Gauss' theorem to the region enclosed by the pillbox, we have:

$$\int_s \boldsymbol{D} \cdot d\boldsymbol{s} = \int_v \rho \, dv = 0 \qquad [4\text{-}25]$$

if the free charge within the box is zero (we shall assume the boundary to be charge-free). Also, because we can neglect the area of the sides of the pillbox, equation [4-25] reduces to

$$-D_{1_n} s + D_{2_n} s = 0 \qquad [4\text{-}32]$$

or

$$D_{1_n} = D_{2_n} \qquad [4\text{-}33]$$

where D_{1_n} and D_{2_n} represent, respectively, the normal components of the displacement in the two media. Therefore, we conclude that *the normal component of the displacement has the same value on the two sides of the boundary.*

The behavior of $\boldsymbol{E}$ at the boundary can be deduced by considering the rectangular path *cdef* shown on the right side of Figure 4.9. The length of this rectangle is l and its height is negligible. As we have seen, the work done in carrying a unit charge around any closed path in a region pervaded by an electrostatic field is zero, a fact expressed symbolically in the form:

$$\oint \boldsymbol{E} \cdot d\boldsymbol{l} = 0 \qquad [2\text{-}26]$$

This becomes, when applied to the rectangular path *cdef*,

$$-E_{1_t} l + E_{2_t} l = 0 \qquad [4\text{-}34]$$

or

$$E_{1_t} = E_{2_t} \qquad [4\text{-}35]$$

where E_{1_t} and E_{2_t} represent, respectively, the tangential components of the intensity in the two media. Therefore, the *tangential component of the intensity has the same value on both sides of the boundary.*

D and E Within Cavities in a Polarized Dielectric

The principles expressed in equations [4-33] and [4-35] may be used to arrive at useful operational definitions of $\boldsymbol{D}$ and $\boldsymbol{E}$.

Let us suppose (Fig. 4.10) that two cavities are cut in an extensive polarized dielectric medium. Cavity *1* is thin and disk-shaped with its axis parallel to the direction of $\boldsymbol{D}$ in the dielectric. Within such a cavity the effect of the thin edges is negligible: we need consider only the effect of the flat faces. We recall (equation [4-33]) that the normal component of $\boldsymbol{D}$ is continuous across a dielectric boundary, and we further note that, at the flat faces of the disk, the

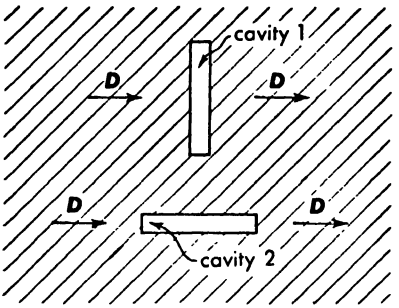


Fig. 4.10

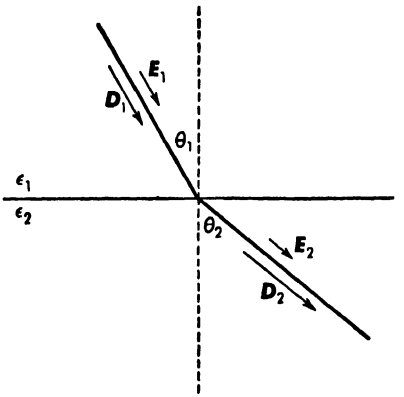


Fig. 4.11

normal component of $\boldsymbol{D}$ is $\boldsymbol{D}$ itself. Thus $\boldsymbol{D}$ within the cavity is identical with $\boldsymbol{D}$ within the dielectric.

Cavity *2* is a long needle-like cavity whose axis is parallel to the direction of $\boldsymbol{D}$ in the dielectric. At the center of such a cavity the effect of the ends is negligible. Also, it is of little interest that the normal component of $\boldsymbol{D}$ is continuous along the sides of the cavity, because there the normal component of $\boldsymbol{D}$ is zero. Instead, we invoke equation [4-35], which states that the tangential component of $\boldsymbol{E}$ is continuous across a dielectric boundary. Since, along the edges of the cavity, the tangential component of $\boldsymbol{E}$ is $\boldsymbol{E}$ itself, it follows that $\boldsymbol{E}$ within the air-filled cavity is identical with $\boldsymbol{E}$ within the dielectric.

Thus, in principle at least, the values of $\boldsymbol{D}$ and $\boldsymbol{E}$ within a dielectric can be ascertained by creating cavities of types *1* and *2*. When test charges are introduced into these cavities the fields in the two cases, say, $\boldsymbol{E}_0$ and $\boldsymbol{E}$, respectively, can be measured. $\boldsymbol{E}_0$, through the relationship $\boldsymbol{D} = \epsilon_0 \boldsymbol{E}_0$, gives the value of $\boldsymbol{D}$ within the dielectric, whereas $\boldsymbol{E}$ gives the corresponding value of the electric field.

The value of the field within a spherical cavity in a polarized dielectric is discussed in § 4.10.

4.8 Refraction of Lines of Force in Passing from One Dielectric Medium to Another

Lines of force undergo refraction in passing from one dielectric medium to a second of different dielectric constant. Let us consider this effect by reference to Figure 4.11. In the first medium, of dielectric constant ϵ_1, the displacement and intensity have the directions shown, making an angle θ_1 with

the normal to the boundary. Likewise, in the second medium, both vectors make an angle θ_2 with the normal to the boundary. We have already derived, in equations [4-33] and [4-35], the rules governing the behavior of $\boldsymbol{D}$ and $\boldsymbol{E}$ at the boundary. These rules can be rewritten in the terms of the angles θ_1 and θ_2. Equation [4-33] becomes

$$D_1 \cos \theta_1 = D_2 \cos \theta_2 \qquad [4\text{-}36]$$

Equation [4-35] becomes

$$E_1 \sin \theta_1 = E_2 \sin \theta_2 \qquad [4\text{-}37]$$

Dividing equation [4-37] by [4-36] and rearranging, we obtain

$$\frac{\tan \theta_1}{\tan \theta_2} = \frac{E_2 D_1}{E_1 D_2} \qquad [4\text{-}38]$$

which becomes, upon substitution of $D = \epsilon E$,

$$\frac{\tan \theta_1}{\tan \theta_2} = \frac{\epsilon_1}{\epsilon_2} \qquad [4\text{-}39]$$

This describes the direction and degree of the refraction of lines of force as they pass from one dielectric medium to another.

One sees from equation [4-39] that lines of force are bent away from the normal as they pass from empty space into a dielectric. This is the reverse of the behavior of light as it passes from empty space into a transparent medium. Figure 4.12 shows the distortion of lines of force that results when a glass disk

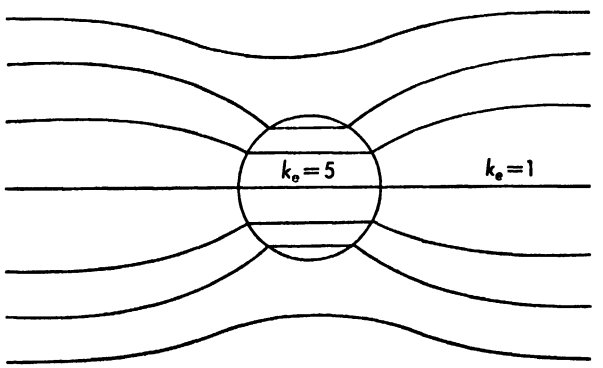


Fig. 4.12

($k_e = 5$) is placed in an air-filled region in which the field was previously uniform.

4.9 Relation Between the Dielectric Constant and the Index of Refraction

Maxwell's study of electromagnetic effects led him to predict the existence of electromagnetic waves. Associated with these waves are time-varying electric and magnetic fields whose directions are mutually perpendicular. In unbounded diclectric media the waves are transverse. The velocity of propagation of such a wave is dependent upon the electric and magnetic properties of the medium through which it is traveling. For an unbounded medium the velocity is given (see §§ 14.2 and 14.3) by

$$v = \frac{1}{\sqrt{\mu\epsilon}} \qquad [4\text{-}40]$$

where μ, the *permeability* of the medium, is the magnetic analogue of the dielectric constant ϵ. Permeability is further discussed in Chapter 8 *et seq.*

Ordinary light consists of the particular group of electromagnetic waves whose wavelengths lie in the approximate range 4000–7000 Å. If we are dealing with *nonmagnetic media*, for which the permeability is μ_0, the velocity of the wave is

$$v = \frac{1}{\sqrt{\mu_0\epsilon}} \qquad [4\text{-}41]$$

The change in its velocity as it passes from one medium to another is responsible for the fact that light is refracted. Indeed, the *index of refraction*, n, of a material is inversely proportional to the velocity with which light is propagated through it:

$$\frac{n_2}{n_1} = \frac{\text{velocity of light in medium 1}}{\text{velocity of light in medium 2}} \qquad [4\text{-}42]$$

Combining equations [4-41] and [4-42], we find

$$\frac{n_2}{n_1} = \frac{v_1}{v_2} = \sqrt{\frac{\mu_0\epsilon_2}{\mu_0\epsilon_1}} = \sqrt{\frac{\epsilon_2}{\epsilon_1}} = \sqrt{\frac{k_{e_2}}{k_{e_1}}} \qquad [4\text{-}43]$$

Further, if medium 1 be taken as free space ($n_1 = 1 = k_{e_1}$), we find for any nonmagnetic material that

$$n = \sqrt{k_e} \qquad [4\text{-}44]$$

This prediction of Maxwell's electromagnetic theory, that the index of refraction of a medium is equal to the square root of its dielectric coefficient, originally served as a basis for criticizing the theory. The reason for the

criticism will be understood by studying the data presented below. The values of k_e that are given are DC values

	n	$\sqrt{k_e}$
Air	1.000294	1.000295
H_2	1.000134	1.000132
Benzene	1.50	1.51
Carbon disulphide	1.64	1.62
Glass	1.6	2.3
Water	1.3	8.9

Although the predictions of equation [4-44] are strikingly borne out for the first four materials, the two last-named substances represent exceptions to the rule. The cause of these discrepancies is now known. Both glass and water contain permanent polar molecules. As a result, the value of k_e decreases with increasing frequency in the manner described in § 3.10. The relevant values of k_e to be used in the comparison with n are, of course, the values at optical frequencies. When these values for glass and water are used, rather than the DC values listed above, the agreement between n and $\sqrt{k_e}$ is found to be satisfactory.

On the other hand, the four first-named materials, which obey the prediction even for DC values of k_e, contain *no* permanent polar molecules. The polarization displayed by these substances is entirely attributable to the displacement, under the action of the applied field, of the electronic clouds of their constituent atoms. Since this displacement occurs with great rapidity, k_e displays no frequency dependence.

When the frequency is comparable to the orbital frequency of the electrons in the material, absorption and re-emission of waves can take place. Under the circumstances, the index of refraction can display an appreciable frequency dependence, as witness, for example, the dispersion of visible light which results from its passage through a glass prism. The subject of dispersion will not be discussed further.

4.10 Field Acting upon a Molecule Within a Dielectric: The "Inner Field"

We now remind ourselves that a dielectric is not continuous but consists, instead, of molecules and atoms. In preparation for studying the polarization

of individual molecules, we shall calculate the so-called *inner field*, E_i, which is the field that acts upon an individual molecule within the dielectric. We shall consider the dielectric, as in Figure 4.13, to occupy the space between plates of a parallel-plate capacitor.

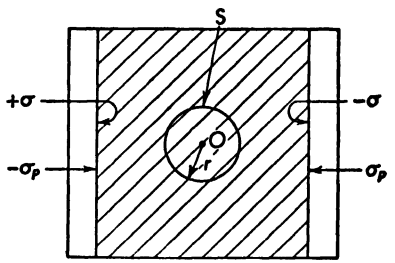


Fig. 4.13

For the purpose of this calculation, we shall regard the dielectric as consisting of two parts: (a) a sphere of radius r (centered at O, where the molecule under consideration is located) within which the discontinuous nature of the dielectric must be taken into account, and (b) the rest of the dielectric slab, which is sufficiently distant from O that it may be regarded as a continuous medium. We shall calculate the inner field at the point O.

The field at O consists of four components:

$$E_i = E_1 + E_2 + E_3 + E_4 \qquad [4\text{-}45]$$

E_1 is the field that would exist at O if the dielectric were absent. Its value is determined solely by the charge density on the plates and is known from equation [1-29] to be:

$$E_1 = \sigma/\epsilon_0 \qquad [4\text{-}46]$$

E_2 is the field component that arises from the polarization charges on the surfaces of the dielectric that are in contact with the capacitor plates. Because the density of these charges on the surface next to the positive capacitor plate is $-\sigma_P$, and $+\sigma_P$ on the surface next to the negative plate, E_2 has the value

$$E_2 = -\frac{\sigma_P}{\epsilon_0} = -\frac{P}{\epsilon_0} \qquad [4\text{-}47]^8$$

E_3 is the field at O arising from the polarization charges on the surface S which would be formed if the spherical section of the dielectric were removed. It will be computed below.

E_4 is the field at O arising from the polar molecules within the spherical section.

[8] The field is taken as positive if it points toward the right in Figure 4.13.

Calculation of E_3

Figure 4.14 shows an enlarged view of the region containing the dielectric sphere. In order to clarify the method by which E_3 is computed, it is assumed

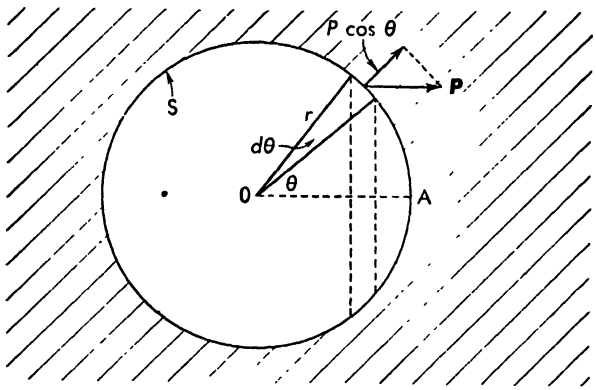


Fig. 4.14

that the spherical section is, for the time being, removed. The polarization in the rest of the dielectric is, however, the same as it would be in the presence of the sphere. The value of E_3 is now found by summing the contributions to the field at O of all elements of polarization charge on the surface S.

The superficial charge density, σ_S, on the spherical surface S is given by the component of the polarization normal to S, that is,

$$\sigma_S = -P \cos \theta \qquad [4\text{-}48]$$

Thus, the charge on the surface element dS that lies between θ and $\theta + d\theta$ is

$$\sigma_S \, dS = (-P \cos \theta) 2\pi r \sin \theta \, r d\theta \qquad [4\text{-}49]$$

Because of the symmetry of the arrangement, all components of the field that are not normal to the capacitor plates cancel one another. Therefore, the polarization charges on dS give rise to a field at O, in the direction OA, of value

$$dE_3 = \frac{(P \cos \theta)(\cos \theta) 2\pi r^2 \sin \theta \, d\theta}{4\pi \epsilon_0 r^2} \qquad [4\text{-}50]$$

The value of E_3 is then given by

$$E_3 = \int_0^{\pi} dE_3 \qquad [4\text{-}51]$$

which has the value

$$E_3 = \frac{P}{3\epsilon_0} \qquad [4\text{-}52]$$

Concerning the Value of E_4

The value of E_4, the field at O of the other dipoles within the spherical section, will not be calculated here. Indeed, it cannot be calculated exactly for a general case.

It can be shown, however,[9] that, for dielectrics that are cubic crystals, E_4 is identically zero. Also, for gases and dilute solutions (where short-range forces do not come into play), E_4 may be neglected. Thus, although E_4 is not always zero, there are many interesting dielectric materials for which it may be neglected. We shall limit ourselves to such cases.

Combining equations [4-45], [4-46], [4-47], and [4-52], we find the inner field to be

$$E_i = \frac{\sigma}{\epsilon_0} - \frac{P}{\epsilon_0} + \frac{P}{3\epsilon_0} \qquad [4\text{-}53]$$

or

$$E_i = \frac{D - P}{\epsilon_0} + \frac{P}{3\epsilon_0} \qquad [4\text{-}54]$$

or

$$E_i = E + \frac{P}{3\epsilon_0} \qquad [4\text{-}55]$$

We shall find this result to be useful in the discussions that follow.

The Clausius-Mossotti Relation

Equation [4-55] gives the electric field E_i that acts upon a single molecule of the dielectric. Under the action of this field the molecule displays an electric moment p' in the direction of the field which is dependent upon the strength of the field, viz.,

$$p' = \alpha E_i \qquad [4\text{-}56]$$

The molecular constant α is known as the *molecular polarizability*. From the definition of polarization (electric moment/unit volume) we see that

$$P = np' = n\alpha E_i = n\alpha(E + P/3\epsilon_0) \qquad [4\text{-}57]$$

where n is the number of molecules per unit volume of the dielectric.

If we now eliminate P between equations [4-22] and [4-57], we obtain

$$\frac{n\alpha}{3} = \epsilon_0 \frac{(k_e - 1)}{(k_e + 2)} \qquad [4\text{-}58]$$

[9] See, for example, H. Fröhlich, *Theory of Dielectrics*, Oxford University Press, 1949.

which indicates that the molecular polarizability of a dielectric of known density can be determined by measuring its dielectric coefficient.[10]

Equation [4-58] is known as the *Clausius-Mossotti relation.* Reference will be made to it again in § 4.12. It is applicable, as mentioned above, to those dielectrics for which $E_4 = 0$.

4.11 Effect of Temperature on Molecular Polarizability

The electric moment of a molecule may consist of two parts: (a) its intrinsic or permanent moment, $\boldsymbol{p}_p$, and (b) a moment, $\boldsymbol{p}_i$, that is induced under the action of an applied field.[11] We shall consider now the average value of the former in the direction of the applied field.

The permanent electric dipoles that characterize the molecules of polar materials tend to orient themselves in the direction of an applied field. These orientations, however, are being continually disturbed as a result of thermal agitation. Hence, a state of equilibrium is established in which the aligned-with-the-field orientation is preferred but is by no means predominant. The potential energy associated with any given orientation is given by $-\boldsymbol{E}_i \cdot \boldsymbol{p}_p$ (Problem 4.6).

The Boltzmann Probability Factor

Whenever thermal equilibrium exists, the probability that a state with energy W will be occupied has been shown by the Austrian physicist, Ludwig Boltzmann (1844–1906), to have the following dependence upon energy:

$$\text{Probability} \propto e^{-W/kT} \qquad [4\text{-}59]$$

In this expression T is the absolute temperature and k, known as Boltzmann's constant, is the ratio of the gas constant to Avogadro's number ($k = R/N$). This result was derived by Boltzmann by statistical methods in connection with his important studies in the kinetic theory of gases.

A simple example will be given of the use of this expression. The molecules of a gas, in addition to their kinetic energies, possess varying degrees of potential energy, mgh, depending upon their distance, h, from the surface of the earth.

[10] It will be recalled that the density, or mass per unit volume, of a material is given by the product mn, where n is the number of molecules per unit volume, and m is the mass of a single molecule.

[11] In terms of these two moments, the magnitude, p', of the electric moment of the molecule in the direction of $\boldsymbol{E}_1$ is given by

$$p' = p_p \cos\theta + p_i$$

where θ is the angle between $\boldsymbol{p}_p$ and $\boldsymbol{E}_1$.

The probabilities that a molecule will be found at heights $h = h$ and $h = 0$, respectively, are, from equation [4-59],

$$P(h) = \text{constant } e^{-mgh/kT} \qquad [4\text{-}60]$$

and

$$P(0) = \text{constant } e^{-0/kT} = \text{constant} \qquad [4\text{-}61]$$

The ratio of these probabilities is, of course, the ratio of the gas concentrations (n) at the two heights. Thus:

$$\frac{n_h}{n_0} = e^{-mgh/kT} \qquad [4\text{-}62]$$

or

$$n_h = n_0 e^{-mgh/kT} \qquad [4\text{-}63]$$

where the concentration n_h at any height is given in terms of the concentration n_0 at the surface of the earth.

Average Value of the Permanent Electric Moment in the Direction of the Applied Field

The same considerations can be applied to a system of electric dipoles that are in thermal equilibrium within a dielectric that is subjected to an electric field. Here, as mentioned above, the potential energy is given by $-\boldsymbol{E}_i \cdot \boldsymbol{p}_p$. The number of dipoles per unit volume whose direction lies within a solid angle $d\omega$ (lying between θ and $\theta + d\theta$) is given by:

$$dn_{\theta \to \theta + d\theta} = \text{constant } e^{E_i \cdot p_p/kT}\, d\omega \qquad [4\text{-}64]$$

$$= \text{constant } e^{E_i p_p \cos\theta/kT}\, d\omega \qquad [4\text{-}65]$$

If we consider a spherical surface of radius r with the dipole located at the center of the sphere, we can see that the elements of solid angle lying between θ and $\theta + d\theta$ is given by

$$d\omega = \frac{(2\pi r \sin\theta)(r\, d\theta)}{r^2} = 2\pi \sin\theta\, d\theta \qquad [4\text{-}66]$$

This result, when substituted in equation [4-65], gives the number of dipoles per unit volume lying in the angular range θ to $\theta + d\theta$, viz.,

$$dn_{\theta \to \theta + d\theta} = \text{constant } e^{E_i p_p \cos\theta/kT}\, 2\pi \sin\theta\, d\theta \qquad [4\text{-}67]$$

We now inquire into the mean electric moment in the direction of the field, that is, the average value of $p_p \cos\theta$.

Digression Concerning Average Values

Before using equation [4-67] to calculate the mean value of $p_p \cos\theta$, it might be helpful for us to recall, in a simple context, the method of calculating average values. Let us suppose that the heights of a group of persons are distributed as follows: 2 persons are 66″ tall, 1 is 69″, 3 are 71″, 2 are 72″, and 1 is 74″ tall. The average height, $\bar{h}$, is

$$\bar{h} = \frac{2(66) + 1(69) + 3(71) + 2(72) + 1(74)}{2 + 1 + 3 + 2 + 1}$$

The bar above a quantity, as in $\bar{h}$, is the standard symbol for an average value.

For a more general case, let us suppose that N_1 persons have a height h_1, N_2 a height h_2, and so on. The average height is then written:

$$h = \frac{N_1(h_1) + N_2(h_2) + \cdots}{N_1 + N_2 + \cdots} \qquad [4\text{-}68]$$

or

$$\bar{h} = \frac{\sum_1^n N_i h_i}{\sum_1^n N_i} \qquad [4\text{-}69]$$

If the range of heights is divided into a very large number of small subdivisions, the summations indicated in equation [4-69] take the form of integrations. Thus, if the number of persons whose height lies between h and $h + dh$ is given by dN_h, the average height becomes:

$$\bar{h} = \frac{\int_0^\infty h\, dN_h}{\int_0^\infty dN_h} \qquad [4\text{-}70]$$

Returning now to the problem of evaluating the average component of the electric moment in the direction of the applied field, we shall employ an expression of the type of equation [4-70]. Thus:

$$\overline{p_p \cos\theta} = \frac{\int_0^\pi (p_p \cos\theta)\, dn_\theta}{\int_0^\pi dn_\theta} \qquad [4\text{-}71]^{18}$$

[18] In this expression the number of moments per unit volume that lie in the angular range $\theta \longrightarrow \theta + d\theta$ (given by equation [4-67]) is written as dn_θ.

$$= \frac{p_p \int_0^{\pi} e^{E_i p_p \cos\theta / kT} \cos\theta \sin\theta \, d\theta}{\int_0^{\pi} e^{E_i p_p \cos\theta / kT} \sin\theta \, d\theta} \qquad [4\text{-}72]$$

or

$$\overline{p_p \cos\theta} = \frac{p_p \int_0^{\pi} \cos\theta \sin\theta \left(1 + \frac{E_i p_p}{kT} \cos\theta + \cdots\right) d\theta}{\int_0^{\pi} \sin\theta \left(1 + \frac{E_i p_p}{kT} \cos\theta + \cdots\right) d\theta} \qquad [4\text{-}73]$$

For weak fields or high temperatures, $E_i p_p \ll kT$, and all but the first non-vanishing terms in the numerator and denominator of equation [4-73] may be neglected. The first term in the numerator vanishes during the integration, while the second has the value $2E_i p_p / 3kT$. The first term in the denominator has the value 2. Thus:

$$\overline{p_p \cos\theta} \doteq \frac{p_p^2 E_i}{3kT} \qquad [4\text{-}74]$$

predicting, as expected, that if the field is small or the thermal motion is large (high temperature), the mean electric moment in the direction of the field approaches zero: the angular distribution of individual electric moments becomes almost uniform. We shall hereafter assume equation [4-74] to be exact.

Because the route to it has been a tortuous one, we might profitably review the significance of equation [4-74]. This relation applies to the permanent dipole moments, p_p, belonging to polar molecules. These, under the action of a local electric field, E_i, are still more or less randomly oriented, but nevertheless they display an average moment in the direction of the field. Equation [4-74] expresses the value of this average moment, which is seen to be strongly temperature-dependent.

Temperature Dependence of Polarizability

The *induced* electric moment, p_i, on the other hand, is *not* temperature-dependent. It is always parallel to the field and is proportional in magnitude to it. The induced moment of a molecule is, therefore, written as:

$$p_i = \alpha_i E_i \qquad [4\text{-}75]^{18}$$

It is necessary, in the light of equations [4-74] and [4-75], to rewrite equation [4-56], in which the permanent and induced electric moments were

[18] In the cases of p_i and α_i the subscript "i" stands for "induced," but in the case of E_i it stands for "inner."

lumped together. Equation [4-56] now takes the form:

$$p' = p_i + \overline{p_p \cos\theta} = \left(\alpha_i + \frac{p_p^2}{3kT}\right) E_i \qquad [4\text{-}76]$$

Thus, the molecular polarizability defined in connection with equation [4-56] is seen to consist of two distinct components:

$$\alpha = \alpha_i + \frac{p_p^2}{3kT} \qquad [4\text{-}77]$$

only one of which is a function of temperature.

4.12 Molecular Polarizability and Molecular Structure

Peter Debye (1884–), Dutch-born physical chemist, was the first to realize that important features of molecular structure could be deduced by observing the variation with temperature of the molecular polarizability of materials. The analysis leading up to equation [4-77] was developed (1912) and brilliantly exploited by him. An essentially similar type of calculation had been carried out earlier (1905) by Pierre Langevin to obtain the mean magnetic moment, in the direction of an applied magnetic field, of gas molecules that possess a permanent magnetic moment.

In order to understand the method by which these ideas have been applied to the study of molecular structure, let us rewrite the Clausius-Mossotti relation (equation [4-58]) in a form that reflects the complex character of α, viz.,

$$\frac{(k_e - 1)}{n(k_e + 2)} = \frac{p_p^2}{9\epsilon_0 kT} + \frac{\alpha_i}{3\epsilon_0} \qquad [4\text{-}78]$$

This equation has the form of the equation for a straight line, viz.,

$$y = mx + b \qquad [4\text{-}79]$$

where

$$y = \frac{(k_e - 1)}{n(k_e + 2)} \qquad [4\text{-}80]$$

$$x = \frac{1}{T} \qquad [4\text{-}81]$$

$$m = \frac{p_p^2}{9\epsilon_0 k} \qquad [4\text{-}82]$$

and

$$b = \frac{\alpha_i}{3\epsilon_0} \qquad [4\text{-}83]$$

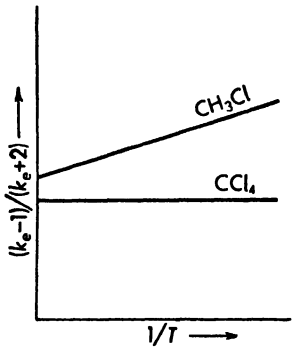


Fig. 4.15

Figure 4.15 shows the behavior of two gases for which y is plotted against x. As expected, the y *versus* x plot is, in each case, a straight line. Attention is first drawn to the CCl_4 line. Evidently, since the slope of the line is zero, the molecule of this gas has no permanent dipole moment. Otherwise $(k_e - 1)/n(k_e + 2)$ would show a temperature-dependence. One is thus able to conclude that the four chlorine atoms in the CCl_4 molecule are symmetrically placed with respect to one another and to the central carbon atom. For such an arrangement the center of gravity of the negative charge coincides with the center of gravity of the positive charge. On the other hand, the temperature dependence of $(k_e - 1)/n(k_e + 2)$ for CH_3Cl provides unequivocal evidence that this molecule possesses a permanent dipole moment which, incidentally, can be calculated from the slope of the line, using equation [4-82]. Some values of permanent dipole moments that have been determined in this fashion are shown in Table 2.

TABLE 2

Permanent Electric Moments of Some Molecules Measured in the Gaseous State

Substance	*Chemical formula*	p_p *(arbitrary units)*
Argon	A	0.0
Hydrogen	H_2	0.0
Hydrochloric acid	HCl	1.03
Carbon monoxide	CO	0.118
Water	H_2O	1.87
Ammonia	NH_3	1.44
Carbon tetrachloride	CCl_4	0.0

From Table 2 it is seen that the monatomic argon and the molecular H_2 (consisting of two identical atoms) are nonpolar. When, however, two different atoms are combined, as in HCl (Fig. 4.1) and CO, a slight charge dissymmetry is introduced. The atoms in the triatomic molecule H_2O are apparently not situated in a symmetric linear arrangement. Indeed, the result shown in Table 2 gave the clue to the fact that the water molecule is a triangular structure. Likewise, there is proof of dissymmetry in ammonia, but none in carbon tetrachloride.

Before leaving this subject it should be re-emphasized that equation [4-55] for the inner field was derived on the implicit assumption (except for cubic crystals) that the interaction energy between nearby dipoles in the dielectric is negligible. This assumption is justified for gases and weak solutions, but not for polar liquids or solids. For further details of this complicated subject, the interested reader is referred to recent monographs on the subject.[14]

4.13 Ferroelectricity[15]

The term *ferroelectricity* is applied to certain crystalline dielectric materials that exhibit a permanent electric polarization which is analogous in many ways to the permanent magnetization exhibited by ferromagnetic materials (§ 10.17). The prefix "ferro" arises from this analogy to ferromagnetism; it has nothing whatever to do with iron. Three outstanding ferroelectric substances, in the order of their discovery, are Rochelle salt, potassium dihydrogen phosphate, and barium titanate.

Under the action of an applied electric field a displacement occurs in the crystal structure. The polarization associated with this displacement creates an internal field which both increases and stabilizes the polarization, with the result that a portion of the polarization remains when the applied field is removed. The unusual behavior of barium titanate in this respect is shown in Figure 4.16, where the intensity E applied to the crystal is plotted as the independent variable and the resultant displacement D within the crystal is shown as the ordinate.

If the crystal is in the virgin state, the relationship between D and E, as E is increased, is shown by the line OA. Evidently, since this is not a straight line, the dielectric constant (D/E) does not have a fixed value but depends upon the value of the applied field.

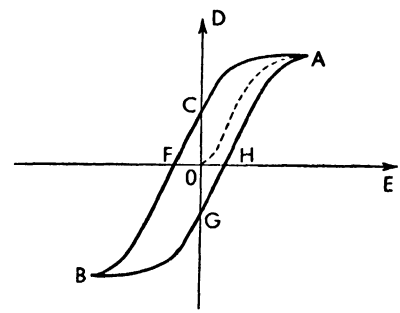


Fig. 4.16

For very small values of E, the resulting value of D is not large, as the first portion of the curve OA indicates. As E is further increased, however, the value of D increases very rapidly before leveling off near A as saturation is ap-

[14] H. Fröhlich, *Theory of Dielectrics*, Oxford, 1949; C. P. Smyth, *Dielectric Behavior and Structure*, McGraw-Hill, 1955.

[15] Ferroelectricity is sometimes (but not usually) known as Seignette electricity. Seignette, a pharmacist of La Rochelle (France), first produced Rochelle salt in 1672 for use as a laxative.

proached. In the steep portion of the curve the dielectric coefficient ($k_e = D/\epsilon_0 E$) attains exceptionally high values ($500 < k_e < 5000$), a behavior that is turned to good account in the construction of "miniature" capacitors. These are capacitors that achieve, in a small volume, values of capacitance that would entail a much greater physical size if conventional dielectrics, such as mica, were employed (see Table 1 and § 3.7). The variation of k_e with the applied field is, of course, a factor that limits the usefulness of this type of capacitor.

If the sample is in the state of polarization represented by the point A, the displacement does not fall to zero when the applied field is turned off. Rather, the sample follows the path AC and retains a residual displacement of amount OC. In fact, in order to remove the residual displacement, it is necessary to apply an opposing field of value OF, causing the sample to change from state C to state F. The value of this opposing field that reduces D to zero is known as the *coercive force*. When the opposing field is increased beyond this value, the change in the state of the sample is represented by motion along the line FB. When state B has been reached, a sudden reversal of the applied field carries the sample to state A along the route $BGIIA$. It is not possible, by simply removing the field, to bring the sample to its original nonpolarized condition. The closed loop representing the relationship between D and E is known as a *hysteresis loop*.

The fact that D does not drop to zero when E is removed qualifies barium titanate and other ferroelectrics for roles in electronic computers as storage or memory units, as well as in switching and amplifier circuits.

The preceding discussion of ferroelectricity has been a rather sketchy one. However, because of the close analogy between ferroelectricity and ferromagnetism, it is not necessary to describe both in detail. The latter is treated at some length in Chapter 10, and the reader who seeks further details of ferroelectric behavior is referred to that discussion. There (§ 10.17) he will correctly infer, for example, that ferroelectrics (ferromagnetics) lose their ferroelectric (ferromagnetic) properties above a certain temperature, T_C, known as the *Curie temperature* (for Rochelle salt, $T_C = 24°$ C; for potassium dihydrogen phosphate, $T_C = -150°$ C; for barium titanate, $T_C = 120°$ C). In addition, ferroelectrics (ferromagnetics) exhibit domain structure. The fuller discussion of hysteresis loops given in §§ 10–18 and 10–19 is applicable also to ferroelectricity.

An important distinction between the two effects is that although ferroelectricity is explicable in terms of classical physics (in this case the interaction of electric dipoles), the spontaneous magnetic polarization that characterizes a ferromagnetic can be rationalized only in terms of quantum mechanics.

4.14 Piezoelectricity and Pyroelectricity

Piezoelectricity is a property manifested by various crystalline materials, especially quartz and Rochelle salt, whereby mechanical pressure exerted upon the crystal results in the appearance of electric charge along its surfaces. The reverse effect, the distortion of the crystal, can be achieved by the application of electric charges. Piezoelectricity is due to electrical polarization resulting from mechanical stress. The effect was discovered in 1883 by Pierre Curie (1859–1906), the French physicist and chemist, who is better known for his work in radioactivity, in collaboration with his wife.

Both the piezoelectric effect and its converse are of immense importance. Because of the effect, a quartz crystal that is mechanically vibrating gives rise to an AC voltage across its face. As the frequency of mechanical vibration of the crystal depends upon the dimensions and other parameters of the crystal, the AC voltage thus generated possesses a very constant frequency. This is utilized in stabilizing the frequencies of electronic signal generators.

The reverse effect provides a convenient means of producing ultrasonic waves. If an AC voltage is applied to the faces of a crystal, and the frequency of the voltage is identical to the natural frequency of the crystal's mechanical vibration, the crystal can be made to vibrate violently and thus serve as a source of ultrasonic radiation. The crystal in this instance behaves as a *transducer:* it transforms electrical into mechanical energy.

It is observed also that when a crystal with piezoelectric properties is heated or cooled it develops charge at its surfaces—an effect known as *pyroelectricity*. These are piezoelectric charges resulting from the strain associated with thermal expansion or contraction of the crystal.

4.15 Problems

E (4.1) Charges of $+1$ microcoulomb are located at the points (0,0), (1,0), and (2,0) and charges of -1 microcoulomb are located at (4,0), (5,0), and (6,0). What is the dipole moment of this charge configuration? The coordinate distances are in meters.

[Ans.: $-12i \times 10^{-6}$ coulomb-meter.]

E (4.2) Four charges are arranged as shown in Figure 4.17. Calculate the field at the point P which is located a distance r from the center of the configuration of charges. $r \gg l$.

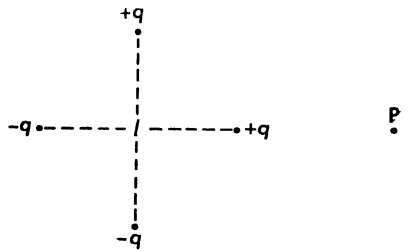


Fig. 4.17

E (4.3) Show that at the point $P(r,\theta)$ in free space the electric field of a dipole of moment p is given by

$$E = \frac{p}{4\pi\epsilon_0 r^3}(1 + 3\cos^2\theta)^{1/2}$$

M (4.4) Prove that the electrostatic energy between two parallel, coaxial dipoles (located in free space) of equal moment p is given by

$$\frac{-2p^2}{4\pi\epsilon_0 l^3}$$

where l is the distance between their centers. Assume l is large compared to the dimensions of the dipoles.

E (4.5) What is the magnitude of the force between the dipoles of Problem 4.4? [Ans.: $6p^2/4\pi\epsilon_0 l^4$.]

M (4.6) A dipole of moment $\boldsymbol{p}$ is located in a uniform electric field of strength $\boldsymbol{E}$. $\boldsymbol{p}$ is initially perpendicular to $\boldsymbol{E}$, and in this position the dipole may be assumed to possess zero potential energy. If the dipole is now rotated in the field, show that the potential energy associated with its new position is $-\boldsymbol{E}\cdot\boldsymbol{p}$.

E (4.7) Show that a dipole of moment $\boldsymbol{p}$ located in a uniform electric field of strength $\boldsymbol{E}$ experiences a torque

$$\boldsymbol{T} = \boldsymbol{p} \times \boldsymbol{E}$$

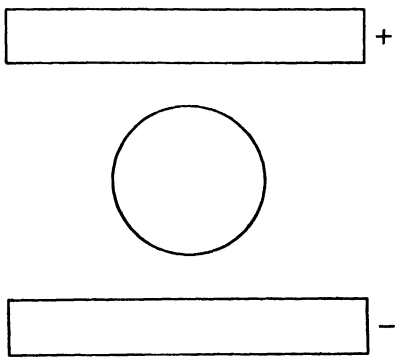
E (4.8) Bearing in mind that that work done in exerting a torque T through an angle $d\theta$ is given by $T\,d\theta$, use the result of Problem 4.7 to derive the result of Problem 4.6.

E (4.9) Show that the electric potential at a distance $\boldsymbol{r}$ from a portion of neutral but polarized dielectric is given by

$$V = \frac{1}{4\pi\epsilon}\int_v \frac{\boldsymbol{P}\cdot\boldsymbol{r}}{r^3}\,dv$$

where r is large compared to the dimensions of the sample of dielectric.

E (4.10) In Figure 4.18 the space between the two capacitor plates is filled with material of dielectric coefficient $k_e = 5$, except for the circular region shown, which is filled with a different material with $k_e = 2$. Sketch the shape of the field in and near the second material.



· Fig. 4.18

E (4.11) The "displacement" of polar molecules that occurs in a dielectric under the action of an applied field results in a certain energy loss through friction. How would you expect this "capacitor" loss to vary with the frequency of the applied AC voltage?

Chapter 5

Direct Currents

All parts of a perfect conductor are at the same potential. If an attempt is made to establish a difference in potential across such a conductor, there is an instantaneous flow of charge so as to equalize it. For nonperfect conductors it is feasible, by use of an appropriate device, to *maintain* a difference in potential, in which case the flow of charge is continuous. Heat is generated in this process.

5.1 Electric Current and Current Density

The *electric current*, I, through any cross-sectional area of a conductor is defined as the net charge that is transferred through it per second, that is,

$$I = \frac{dq}{dt} \qquad [5\text{-}1]$$

This transfer may result from the motion of either positive or negative charges, or both. The mks unit of current, the *ampere* (defined on page 219), represents the passage of *1 coulomb of charge per second.* In such a current, if each charge carrier contains a single elementary charge, the number passing the observation surface per second is 6.25×10^{18} (i.e., $1/1.6 \times 10^{-19}$). When the direction of charge motion does not change with time, the current is said to be *direct* (DC), in contrast to an *alternating* current (AC), where the direction of charge motion periodically reverses. In our discussion of electric currents, the symbol I will be

reserved for steady currents, and the symbol i for those whose magnitudes are functions of time.

"Direction" of a Current

Although current is a scalar quantity, reference is frequently made to the "direction" of the current. The "direction" referred to is, by convention, the direction of motion of positive charges. Thus, the direction of an electric current is from regions of higher potential to regions of lower potential. When the current is caused by the motions of electrons (as in metallic conductors), the direction of motion of the charge is opposite to the direction of the current. If all currents were caused by electron flow, a change in the conventional direction of current might be called for. As it is, however, some currents (e.g., in electrolytic solutions) result from the motion of positive ions, and so complete harmony between the direction of current and charge motion cannot be achieved.

Current Density

The current per unit area in the direction of positive charge motion (that is, through a unit surface that is normal to the direction of charge motion) is known as the *current density*, $\boldsymbol{J}$. In terms of the current density, the current $d\boldsymbol{I}$, through any element of surface $d\boldsymbol{s}$ (Fig. 5.1), is given by

$$d\boldsymbol{I} = \boldsymbol{J} \cdot d\boldsymbol{s} \qquad [5\text{-}2]$$

and the total current through any surface s by

$$\boldsymbol{I} = \int_s \boldsymbol{J} \cdot d\boldsymbol{s} \qquad [5\text{-}3]$$

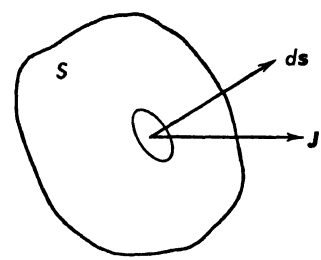


Fig. 5.1

In most cases of interest $\boldsymbol{J}$ and $d\boldsymbol{s}$ are either parallel or anti-parallel (e.g., as in a wire).

We shall now deduce an expression for the current density in terms of the motion of the individual charge carriers. Figure 5.2 represents a portion of

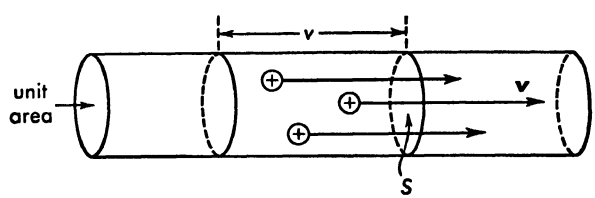


Fig. 5.2

cylindrical material of unit cross-sectional area through which positive charges, $+q$, are moving with uniform velocity. The current density is, by definition, the amount of charge per second which passes through any unit cross-sectional area (say, s) that is normal to the direction of charge motion. In this case, in one second, all the charges contained in the cylinder of length v, shown in Figure 5.2, pass through s. The current density is therefore

$$\boldsymbol{J} = nq\boldsymbol{v} \qquad [5\text{-}4]$$

where n is the number of charge carriers per unit volume of the material. It will be noted that negative charges moving to the left (in which case both q and $\boldsymbol{v}$ are negative) give rise to a current density whose direction is identical with that caused by positive charges moving to the right. Thus, in an electrolytic solution in which both positive and negative ions may be set in motion under the action of an electric field (the positive ions moving with the field and the negative ions moving against it) the total current density is the sum of two parallel vectors, viz.,

$$\boldsymbol{J} = \boldsymbol{J}' + \boldsymbol{J}'' = n'q'\boldsymbol{v}' + n''q''\boldsymbol{v}'' \qquad [5\text{-}5]$$

where the single-primed terms refer to charges of one sign and the double-primed terms to charges of the opposite sign.

■ *Equation of Continuity*

If the integration indicated in equation [5-3] is carried out over a closed surface an expression is obtained for the total current emerging from (or entering) the volume enclosed by the surface. If this current is not zero, it follows that the net positive charge within the enclosed volume is decreasing (or increasing). This statement is expressed symbolically as follows:

$$\int_s \boldsymbol{J}\cdot d\boldsymbol{s} = -\frac{\partial q}{\partial t} = -\frac{\partial}{\partial t}\int_v \rho\, dv = -\int_v \frac{\partial \rho}{\partial t}\, dv \qquad [5\text{-}6]$$

Here the total charge in the enclosed volume has been expressed as a volume integral and the partial derivative with respect to time has been written to emphasize that it is immaterial whether the time differentiation follows or precedes the space integration. By means of the divergence theorem (Appendix), equation [5-6] can be rewritten

$$\int_v \nabla\cdot\boldsymbol{J}\, dv = -\int_v \frac{\partial \rho}{\partial t}\, dv \qquad [5\text{-}7]$$

whence,

$$\nabla\cdot\boldsymbol{J} = -\frac{\partial \rho}{\partial t} \qquad [5\text{-}8]$$

This is the so-called *equation of continuity:* it simply expresses the fact that charge is conserved.

5.2 Electrical Conduction in a Metal: The Electron Gas

Electrons were discovered in 1897 as an outcome of the study of electrical discharges in gases. Soon afterwards, they were shown to be emitted by hot bodies (thermionic emission) and to be ejected from metals irradiated by light (photoelectric effect). The conclusion that they were the current carriers in a conductor was thereafter speedily drawn.

The electrons were soon also pictured as constituting in the metal an *electron gas.* On this basis, they were presumed to bounce around, suffering collisions and exchanging energy, with the result that their collective behavior could be predicted by methods that had been developed for the study of the statistical behavior of the molecules of a gas. For example, their velocities would be distributed according to the Maxwell-Boltzmann distribution law, and an increase in the temperature of the conductor would lead to an increase in the average kinetic energy of the electrons. These views have since been modified; in particular, the average energy of the electrons in a metal is now believed to be relatively insensitive to the temperature of the metal (see § 6.2). In addition, only a small fraction of the total number of electrons in a metal can give rise to an electric current. These are known as the *conduction electrons* (§ 6.3). On the other hand, the modern theory of solids has not invalidated the notion of an electron gas: it simply requires that statistics other than the Maxwell-Boltzmann statistics be employed to describe the gas (§§ 6.1 and 6.2).

5.3 Motion of Conduction Electrons under an Applied Field: Ohm's Law

Let us consider the motion of a conduction electron in an imperfect conductor, under the action of an applied electric field, $\boldsymbol{E}$. Since the electron is negatively charged, its motion is against the field. It is a jerky motion because of the collisions occurring *en route.* The time between collisions is given by

$$t = \lambda/u \qquad [5\text{-}9]$$

where λ is the *mean free path* and u is the *average random velocity* of the electrons. Between collisions the electron is accelerated under the action of the field, thereby acquiring a *forward velocity* v given by:

$$v = at = \left(\frac{-Ee}{m}\right) t = -E\frac{e\lambda}{mu} \qquad [5\text{-}10]$$

In each collision the electron is assumed to lose this forward velocity, thus

necessitating that it start again from scratch. The *average* forward velocity is, therefore, $v_a = v/2$.

The current density corresponding to this flow of electrons is seen from equation [5-4] to be

$$J = -nev_a = -\frac{ne}{2}v \qquad [5\text{-}11]$$

which becomes, upon substitution of equation [5-10],

$$J = \frac{ne^2\lambda}{2mu}E \qquad [5\text{-}12]$$

where n is the number of conduction electrons per unit volume of the conductor. We may rewrite equation [5-12] in the form

$$J = \frac{E}{\rho} \qquad [5\text{-}13]$$

where ρ $(= 2mu/ne^2\lambda)$ is the *resistivity* of the material supporting the current.

The current in the conductor through a cross-sectional area A that is normal to the direction of electron motion is given by

$$I = \mathbf{J}\cdot\mathbf{A} = JA = \frac{EA}{\rho} \qquad [5\text{-}14]$$

Further, if we consider an element of the conductor of length dl, across which, of course, the potential difference has the magnitude $dV = E\,dl$, we can rewrite equation [5-14] in the following useful form

$$I = \left(\frac{A}{\rho\,dl}\right)dV = \frac{dV}{dR} \qquad [5\text{-}15]$$

where

$$dR = \frac{\rho\,dl}{A} \qquad [5\text{-}16]$$

is known as the *resistance* of the element and depends, as indicated, upon its length, its cross-sectional area, and its *resistivity*. For a conductor of finite length equation [5-15] becomes

$$I = \frac{V}{R} \qquad [5\text{-}17]$$

where V is the total potential difference between its ends and R, the resistance of the conductor, is the integral of equation [5-16] over the length of the conductor.

The resistance of a conductor is a measure of the opposition which the conductor offers to the flow of current. Although we shall derive below the rules by which resistances are combined, the main discussion of resistance

per se is postponed until the next chapter. The mks unit of resistance, the volt/ampere, is known as the *ohm*. *An object has a resistance of 1 ohm if a potential difference across it of 1 volt gives rise to a current of 1 ampere.*

Equation [5-13] and its lineal descendant, equation [5-17], have been derived from first principles. The latter predicts a functional relationship between the current in a conductor and the potential difference between its ends. In particular, if the resistivity of the conductor were independent of the current, I would be directly proportional to V. Also, it would not be surprising to discover that n and λ (which partially comprise ρ) differ from one material to another, in which case the numerical value of the resistivity of a conductor would be one of its characteristics. Let us see how well these speculations are borne out.

Ohm's Law

The I *versus* V dependence for conductors was first investigated by Georg Simon Ohm (1787–1854), a high-school teacher in Cologne until his elevation, shortly before his death, to a professorial chair. Ohm's experiments were, of course, not guided by any such analysis as has led up to equation [5-17]. Instead, he simply plotted I as a function of V for a number of different conductors, and discovered that in each case, provided the temperature of the conductor remained unchanged, I was directly proportional to V, that is,

$$I = kV \qquad [5\text{-}18]$$

When the constant of proportionality k is written as $1/R$, as is the time-honored practice, this equation becomes identical with [5-17]:

$$I = \frac{V}{R} \qquad [5\text{-}19]$$

This empirically established result, in which the resistance is a constant, holds for metals and for certain other materials as well. It is known as *Ohm's law*, a name which is sometimes also applied to the equivalent statement (equation [5-13]) that the current density in a metal is proportional to the electric field.

The other speculation made above, that the resistivity is characteristic of a material, is also confirmed by experiment. This subject is discussed further in the next chapter.

Non-ohmic Materials

Materials for which the current is not proportional to the potential difference are known as *non-ohmic* materials. This category includes many materials

of great practical importance. The *I versus V* characteristics of some of them are discussed in § 6.7. In the meantime, during the balance of this chapter, we shall assume that we are dealing with ohmic materials only.

It should be reiterated that the ratio $R = V/I$ (equation [5-17]) provides the general definition of resistance. If R is independent of the value of I, the material is ohmic, but even if R is dependent upon the magnitude of I, it is still defined by the V/I ratio.

5.4 Refraction of Lines of Current Flow at a Boundary Between Conductors of Different Resistivity

We have already noted the relation between current and current density:

$$I = \frac{dq}{dt} = \int_s \boldsymbol{J} \cdot d\boldsymbol{s} \qquad [5\text{-}1]\ [5\text{-}3]$$

Let us suppose that the surface s is a closed one, enclosing a volume of conductor. When a steady state has been established, for which $dq/dt = 0$, we have

$$\int_s \boldsymbol{J} \cdot d\boldsymbol{s} = 0 \qquad [5\text{-}20]$$

Let the surface be a pillbox of the type shown in Figure 4.9, enclosing a portion of the interface between conductors of resistivity ρ_1 and ρ_2, respectively. Then, by a line of reasoning identical to that employed in § 4.7 in showing that the normal component of $\boldsymbol{D}$ is continuous across a charge-free boundary between two dielectrics of different ϵ, the reader can readily show from equation [5-20] that the normal component of $\boldsymbol{J}$ is continuous across the boundary between the two conductors. That is,

$$J_{1_n} = J_{2_n} \qquad [5\text{-}21]$$

In addition, by using the same argument as that for the dielectric case, the tangential component of the electric field is found to be continuous across the boundary. That is,

$$E_{1_t} = E_{2_t} \qquad [5\text{-}22]^1$$

or, substituting [5-13],

$$\rho_1 J_{1_t} = \rho_2 J_{2_t} \qquad [5\text{-}23]$$

---

[1] This argument is valid provided the electric field in the boundary region is electrostatic. Under these conditions

$$\oint \boldsymbol{E} \cdot d\boldsymbol{l} = 0 \qquad [2\text{-}26]$$

If the boundary region contains a source of electromotive force (§ 5.8), then the line integral of $\boldsymbol{E}$ around a closed path such as *cdef* in Figure 4.9 is no longer zero.

The case where $\rho_1 > \rho_2$ is shown in Figure 5.3. Evidently, from the figure,

$$\frac{\tan\theta_1}{\tan\theta_2} = \frac{(J_{1_t}/J_{1_n})}{(J_2./J_{2_n})} \qquad [5\text{-}24]$$

which becomes, upon substitution of equations [5-21] and [5-23],

$$\frac{\tan\theta_1}{\tan\theta_2} = \frac{\rho_2}{\rho_1} \qquad [5\text{-}25]$$

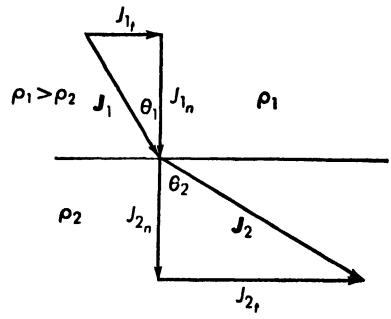


Fig. 5.3

Conditions at an Insulator-Conductor Boundary

An example of special interest is that in which one of the resistivities, say, ρ_1, is very much larger than the other. A case in point is a wire surrounded by an insulator. First we note, from equation [5-25], that $\tan\theta$, and hence θ, are small, indicating that J, is practically perpendicular to the interface, with the effect that J_{1_t} and E_{1_t} are practically zero. Hence, from equation [5-22], the tangential component of the field in the conductor, E_{2_t}, is essentially zero. Likewise, from $E_{2_n} = (\rho_2/\rho_1)E_{1_n}$ (equation [5-2']), the normal component of the field in the conductor is also seen to be small. In short, both components of the field in a good conductor are practically negligible.

5.5 Laplace's Equation for a Steady Current

A steady DC current represents a state of dynamic equilibrium in which the rate of inflow of charge to any volume of the conductor is exactly equal to the rate of charge outflow, that is, $\partial\rho/\partial t = 0$. From the equation of continuity (equation [5-8]) we therefore have

$$\nabla\cdot\boldsymbol{J} = 0 \qquad [5\text{-}26]$$

When Ohm's law is obeyed, we know in addition that

$$\boldsymbol{J} = \frac{\boldsymbol{E}}{\rho} \qquad [5\text{-}13]$$

where ρ, the resistivity,[‡] is a constant. Combining [5-26] and [5-13], we have

$$\nabla\cdot\boldsymbol{E} = 0 \quad \text{or} \quad \nabla^2 V = 0 \qquad [5\text{-}27]$$

[‡] The reader is cautioned that the symbol ρ is here serving double duty: in the equation of continuity, ρ is the volume density of charge, whereas in equation [5-13] it represents the resistivity.

Thus, for the case of steady current through an ohmic conductor the potential satisfies the same equation, Laplace's equation, as does the potential within an uncharged isotropic dielectric that is located in an electric field (see §§ 2.4 and 2.12). Consequently, since $\boldsymbol{J} = \boldsymbol{E}/\rho = -\nabla V/\rho$, we can obtain the current distribution in a conductor by solving Laplace's equation. For a given electrode arrangement, the lines of current flow in a uniform conductor would be identical to the lines of force in a uniform dielectric. This result explains why the equipotential lines and lines of force for a system of *fixed* charges can be obtained (§ 2.8) by employing an arrangement in which charges are in motion. The use of AC current does not invalidate this result since the necessary condition, $\partial\rho/\partial t = 0$, is still satisfied.

It has been pointed out, in discussing the usefulness of Laplace's equation in solving problems in electrostatics, that the solution of the equation involves a knowledge of the boundary conditions. A similar knowledge is needed for the solution of problems involving current flow. The usual boundary conditions are the following: the two ends of the conductor are equipotential surfaces ($V = V_1$ and $V = V_2$, respectively), while along the sides of the conductor, inasmuch as $\boldsymbol{E}$ is parallel to the surface of the conductor (see preceding subsection), the derivative of the potential with respect to the coordinate that is normal to the surface is zero. This latter boundary condition is often written $\partial V/\partial n = 0$, where n is the normal coordinate referred to.

5.6 Resistances in Series and Parallel

Series

The total resistance of a number of resistances in series is the sum of the separate resistances. This can be readily understood by reference to Figure 5.4.

Evidently, the entire current I is present in each of the separate resistances. The potential difference, or *voltage drop*, across R_i is seen from Ohm's law to be

$$V_i = IR_i \qquad [5\text{-}28]$$

Similarly, the total voltage drop across the entire resistance chain is

$$V = \sum_1^n V_i = I \sum_1^n R_i \qquad [5\text{-}29]$$

If we also write Ohm's law in terms of the total resistance, R, of the circuit:

$$V = IR \qquad [5\text{-}30]$$

we conclude, by comparing equations [5-29] and [5-30], that

$$R = \sum_1^n R_i \qquad [5\text{-}31]$$

as stated above.

Parallel

The reciprocal of the total resistance of a number of resistances in parallel is equal to the sum of the reciprocals of the separate resistances. This will be demonstrated by reference to Figure 5.5.

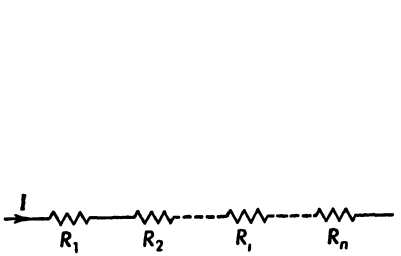


Fig. 5.4

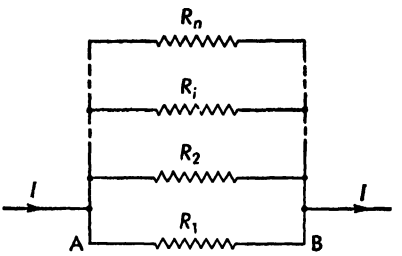


Fig. 5.5

In this case the total current I subdivides among the several resistances. I_1 amperes is the current in R_1, I_2 in R_2, and so on. The voltage drop V across each resistance, i.e., the potential difference between points A and B, is the same. Ohm's law for each separate resistance, therefore, takes the form:

$$I_i = \frac{V}{R_i} \qquad [5\text{-}32]$$

The total current in the entire resistance array is

$$I = \sum_1^n I_i = V \sum_1^n \frac{1}{R_i} \qquad [5\text{-}33]$$

from which, by comparison with equation [5-19], the relation between the total resistance R and the separate resistances R_i is seen to be

$$\frac{1}{R} = \sum_1^n \frac{1}{R_i} \qquad [5\text{-}34]$$

as stated above.

Further discussion of electrical resistance is delayed until the next chapter.

5.7 Energy Dissipation Resulting from DC Current Flow

The movement of electric charges in a conductor under the action of an electric field represents the doing of work or the expenditure of energy. This energy is imparted, in the first instance, to the conduction electrons and is then passed on by them, through collisions, to the crystal lattice of the conductor. Thus, the energy expended in accelerating the conduction electrons reappears

as heat in the conductor. Let us calculate the amount of heat that is produced in this manner.

The work done (§ 2.9) upon a charge of dq coulombs as it falls through a potential difference of V volts is $V\,dq$ joules. If this operation occurs in time dt, the work done per second—and, consequently, the rate of heat production—is

$$\text{Heat energy/sec} = V(dq/dt) \qquad [5\text{-}35]$$

$$= VI \qquad [5\text{-}36]^{3}$$

which becomes, upon substitution of $V = IR$,

$$\text{Heat energy/sec} = I^2R \qquad [5\text{-}37]$$

In the mks system, equation [5-37] is expressed in joules/second. By substituting the mechanical equivalent of heat (1 calorie = 4.19 joules), one obtains

$$\text{Heat energy/sec} = 0.24I^2R \text{ calories/sec} \qquad [5\text{-}38]$$

where I is in amperes and R is in ohms.

The result expressed above, that the rate of production of heat in a constant-resistance circuit is proportional to the square of the current, is known as *Joule's law*. James Prescott Joule (1818–1889), an English physicist, made valuable contributions to the field of heat, thermodynamics, and electricity. Joule (rhymes with "mule") was the first to determine the mechanical equivalent of heat.

Electrical Power

Power is the rate of doing work. Equations [5-35]–[5-38], giving the rate at which heat energy is produced when electric charges flow through an imperfect conductor, represent at the same time the rate at which work is done on the charges. Since this work is done by electrical forces, electrical energy is expended in the process. In these circumstances electrical power is said to be dissipated, that is,

$$\text{Electrical power loss} = VI, \qquad [5\text{-}39]^{4}$$

which becomes, upon substitution of $V = IR$,

$$\text{Electrical power loss} = I^2R \qquad [5\text{-}40]$$

---

[3] If the potential difference and the current are both fluctuating, equation [5-36] is rewritten Heat/sec = vi. This expression gives the *instantaneous* rate of heat production and is applicable, for example, to alternating currents, even when v and i are not in phase (see Chapter 13).

[4] The footnote that was made apropos of equation [5-36] is also applicable here.

This electrical power loss is colloquially referred to as the "I^2R loss" and results, as mentioned above, in the production of heat.

Whenever electrical power is dissipated in a circuit, there must exist elsewhere in the circuit a *source* of electrical power. Such a source may take a variety of forms, but in each case work is done on electric charges against electrical forces. As a result, electrical energy is produced. The amount of energy so produced is calculated also from equation [5-39], where I represents the amount of charge per second that is moved, by an external agent, through the potential difference V.

The mks unit of power is the *watt*. *The power dissipation (or production) is 1 watt when 1 joule of electrical energy is lost (or produced) per second.*

The first unit of power was the *horsepower*, defined by James Watt (1736–1819), Scottish inventor of the modern condensing steam engine. It was intended to represent the power of an average working horse, and is equivalent to 33,000 ft-lb of work per minute. Watt's observations of the rate at which horses could work were made on the strong dray horses employed by London breweries. Their capacity for work was about one and one-half times that of the normal horse. One horsepower is equivalent to 746 watts.

Electrical Energy

Although the joule or the watt-second is the mks unit of electrical energy, it is a rather small quantity. For this reason electrical energy is generally sold at so much per *kilowatt-hour*.

5.8 Electromotive Force

When a difference in potential exists between two points in a conductor, the conductor possesses electrostatic energy equal to the amount of work that was done on individual charges in the process of establishing the difference in potential. When such a condition exists some of the charges within the conductor immediately migrate so as to remove the difference in potential. In this operation electrostatic energy becomes transformed into kinetic energy of the charges. Subsequently, this kinetic energy is lost, through collisions, to the crystal lattice. The motion of the charges constitutes an electric current, but it will not be a continuous current unless the electrostatic energy of the system is continuously replenished by some external agent.

Any device for continuously transforming some other form of energy (mechanical, chemical, thermal, etc.) into electrical energy is known as a source of *electromotive force*, written as *emf* and designated by the symbol $\mathcal{E}$ (if steady)

or ε (if fluctuating). Thus, a source of emf, if located in a closed circuit, can maintain an electric current. The magnitude of the emf associated with such a source is the amount of work that is done (per unit charge) in transforming the other type of energy into electrical energy. The reader will realize from this that emf has the same dimensions as potential: in the mks system the unit of emf is the volt. Thus, *if the emf is* 1 *volt*, 1 *joule per coulomb of work is done during the energy-transformation process.*

Electromotive Force versus Difference in Potential

The fact that emf and potential are both expressed in the same units sometimes causes the one to be confused with the other. The following remarks should help to distinguish between the two.

The total electric field within a conductor may be written

$$\boldsymbol{E} = \boldsymbol{E}_{es} + \boldsymbol{E}_{emf} \qquad [5\text{-}41]$$

where $\boldsymbol{E}_{es}$ is the electrostatic field caused by the charge distribution within the conductor, and $\boldsymbol{E}_{emf}$ is the field produced by the source(s) of emf. In order to evaluate the line integral of $\boldsymbol{E}\cdot d\boldsymbol{l}$ around a closed circuit, we note first that

$$\boldsymbol{E}\cdot d\boldsymbol{l} = \rho \boldsymbol{J}\cdot d\boldsymbol{l} = \rho\, dl \frac{I}{A} = I\, dR \qquad [5\text{-}42]$$

where $d\boldsymbol{l}$ is assumed to be parallel to $\boldsymbol{J}$. Now, integrating around the circuit,

$$\oint \boldsymbol{E}\cdot d\boldsymbol{l} = I \oint dR \qquad [5\text{-}43]$$

or

$$\oint \boldsymbol{E}_{es}\cdot d\boldsymbol{l} + \oint \boldsymbol{E}_{emf}\cdot d\boldsymbol{l} = IR \qquad [5\text{-}44]$$

where R is the total resistance of the circuit. The first term on the left of equation [5-44] is zero, as for all electrostatic fields (equation [2-26]). The second term on the left is the electromotive force in the circuit. It is the total work which the source of emf does on a unit charge as the charge moves around the circuit. If the emf is not zero, then the right side of the equation is not zero, that is, a current exists in the circuit:

$$\mathcal{E} = IR \qquad [5\text{-}45]$$

This expression is similar in form to the expression $V = IR$ (equation [5-17]) by which resistance was defined. Referring to Figure 5.6, we shall now discuss the relationship between the two expressions.

Equation [5-45] refers to the circuit as a whole: it gives the total current in the circuit in terms of the total resistance ($r + R$) and the total circuital emf.

Thus,

$$\mathcal{E} = I(r + R) \qquad [5\text{-}46]$$

for the circuit of Figure 5.6. Equation [5-17], on the other hand, may deal only with some particular circuit element through which an electric current is passing. Disregarding the agent that is responsible for the current flow, it gives the potential difference across the element. This potential difference ($V = IR$) is known as the *potential drop*, or *IR drop*, across the element in question. In the circuit of Figure 5.1 there are two potential drops, namely,

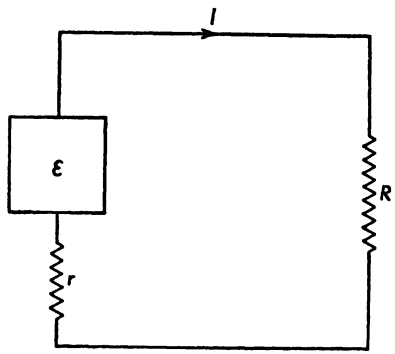


Fig. 5.6

$$V_r = Ir \qquad \text{and} \qquad V_R = IR \qquad [5\text{-}47]$$

Further, adding all such potential drops, we have

$$V = V_r + V_R = I(r + R) \qquad [5\text{-}48]$$

Comparing equations [5-46] and [5-48], we see that the total emf in the circuit is equal to the sum of the potential drops in the circuit. In general, for any closed circuit,

$$\sum_1^n \mathcal{E}_i = \sum_1^m I_j R_j \qquad [5\text{-}49]$$

This result, known as *Kirchhoff's second law* (§ 5.11), is simply an expression of the fact that electric potential is a single-valued function of position. In moving around a closed loop, one encounters certain potential rises (caused by the sources of emf) as well as certain potential drops (IR terms). When the loop has been circumnavigated, the total of the potential rises must, of course, be equal to the total of the potential drops.

In most cases, as in the above example, electromotive force gives rise to a difference in potential. Thus, in an electrolytic cell (to be discussed below), the emf of the cell maintains a positive charge on one plate of the cell and a negative charge on the other, and so creates a difference in potential. This difference in potential is a consequence of the emf and is not to be confused with the emf itself.

Some Sources of emf

The *Van de Graaff generator*, discussed in § 2.9 and illustrated in Figure 2.13, is a source of emf. In this instrument, once equilibrium has been established,

the charge that is carried from the sphere to the ground, either by leakage or along the acceleration tube, is immediately replaced by fresh charge that is borne upward on the endless belt. Here mechanical energy is transformed into electrical energy.

The *electrolytic cell* is another source of emf. The first such cell, invented by the Italian physicist, Count Alessandro Volta (1745–1827), consisted of discs of copper and zinc, separated by a piece of material that had been wetted with an electrolytic solution. The arrangement was equivalent to that shown in Figure 5.7.

The action of the electrolytic cell depends upon the tendency of certain metals, when placed in an electrolyte, to go into solution as positive ions. Zinc displays this tendency to a marked degree. Copper, on the other hand, goes into solution slightly, if at all. As a result, the zinc rod in Figure 5.7 is left with a large excess negative charge while the copper actually acquires a smaller excess positive charge caused by the passage to it of H^+ ions from the electrolyte. If the two rods are connected by an external wire, electrons flow through the wire from the zinc (lower potential) to the copper (higher potential). For every two electrons that depart from the zinc plate, an additional Zn^{++} ion goes into solution. Also, each electron on arrival at the copper plate is immediately neutralized by a H^+ ion from the electrolyte. Thus, the difference in potential is maintained between the plates and the cell serves as a source of emf. In this case chemical energy is transformed into electrical energy.

Thermocouples (see Chapter 7) and generators (in which an emf is induced when the magnetic flux through a loop changes—see Chapter 9) provide two further examples of sources of electromotive force.

Primary and Secondary Cells

The numerous types of electrolytic cells fall into either the *primary* or the *secondary* category. In both groups chemical energy is the source of the electrical energy that the cell supplies. Consequently, with use, some of the materials constituting the cell are depleted. In a primary cell the depleted materials must be replaced from time to time if the cell is to provide extended service. In a secondary or *storage* cell the depleted materials can be renewed by forcing a current backwards through the cell. In this act of *charging* the cell, the chemical reaction that led to the discharge of the cell is reversed.

The various types of electrolytic cells, and their merits and demerits, will not be discussed here, although some description of standard cells is given in § 6.14.

For certain applications several cells are grouped together to form *batteries*. Some such arrangements are discussed in § 5.10.

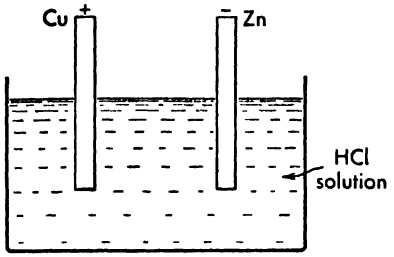


Fig. 5.7

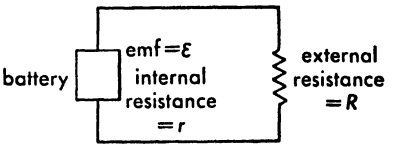


Fig. 5.8

Open-Circuited and Terminal Voltages of a Cell or Battery

In a battery which does not form part of a closed circuit, the potential difference between the plates is equal to the emf of the cell. This is the so-called *open-circuit voltage* of the battery. This is the potential difference that would be measured by a voltmeter, such as an electrostatic voltmeter (§ 3.11), that draws no current.

Such is not the case when the battery forms part of a closed circuit. Here one must reckon with the fact that the materials comprising the battery themselves offer resistance to the flow of current. A potential drop occurs across the internal resistance of the battery in the same manner as the potential drops across the other resistive elements in the circuit. Consequently, when a battery is delivering current to a load, as in Figure 5.8, the potential difference between its terminals, the so-called *terminal voltage*, is less than the battery's emf by the amount of the voltage drop across its internal resistance. It is readily shown (Problem 5.3) that the terminal voltage of a battery of emf ℰ volts and internal resistance r ohms is given by

$$\text{terminal voltage} = \mathcal{E}\left[\frac{R}{R+r}\right] \qquad [5\text{-}50]$$

when the battery is delivering current to a load of R ohms.

5.9 Condition for Maximum Power Transfer

An electric circuit normally contains a source of emf, which serves as the source of electrical energy, and a *load*, in which electrical energy is dissipated. Frequently, it is desirable to expend in the load as much of the available electrical energy as is possible. There is concern, therefore, with the efficiency with which electrical energy is transferred from the source to the load. This energy transfer is usually computed on a per-second basis, in which case it is properly referred to as *power transfer*. We shall now derive the condition for maximum power transfer from a source of emf to a load.

Let us consider the circuit shown in Figure 5.8. The source of emf is a battery of emf $\mathcal{E}$ volts and internal resistance r ohms. The only other circuit element is the load of resistance R ohms. As a result of charge flowing in the circuit, power is dissipated in the load, as intended, but also is dissipated in the internal resistance of the battery. When the largest fraction of the power that is dissipated in the circuit is dissipated in the load, maximum power transfer is said to be achieved.

The current in the circuit, given by equation [5-46], is

$$I = \frac{\mathcal{E}}{(r + R)} \qquad [5\text{-}51]$$

The power dissipation in the load, given by equation [5-40], is

$$P = I^2R = \mathcal{E}^2R/(r + R)^2 \qquad [5\text{-}52]$$

For maximum power transfer,

$$\frac{dP}{dR} = 0 = \frac{d}{dR}\left[\frac{\mathcal{E}^2R}{(r + R)^2}\right] \qquad [5\text{-}53]$$

Solving equation [5-53], one finds that

$$r = R \qquad [5\text{-}54]$$

Thus, the condition for maximum power transfer from the battery to the load is obtained when the resistance of the load is equal to the internal resistance of the battery.

The AC counterpart of this theorem, which is of particular importance in communications, will be derived in § 13.20.

5.10 Advantages of Various Cell Arrangements Within a Battery

When a number of cells are available for use they may be connected in series, in parallel, or in some series-parallel arrangement. As we have seen, the arrangement which provides maximum power transfer is that in which the internal resistance of the battery is matched to that of the load.

Series Connection

The series arrangement is shown in Figure 5.9, where each individual cell is indicated by a pair of parallel bars, the longer and thinner bar representing the positive terminal. The emf and internal resistance of a single cell are, respectively, $\mathcal{E}$ volts and r ohms. The total emf and internal resistance of the

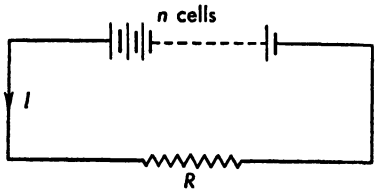


Fig. 5.9

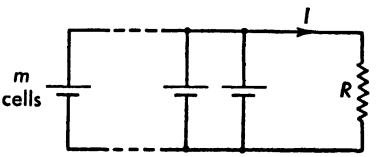


Fig. 5.10

battery are, therefore, $n\mathcal{E}$ volts and nr ohms, respectively. The current flowing through the load, given by equation [5-46], is

$$I = \frac{n\mathcal{E}}{nr + R} \qquad [5\text{-}55]$$

This arrangement is patently suited to deliver power to a load whose resistance is large compared to the internal resistance of a single cell.

Parallel Connection

In the parallel arrangement, shown in Figure 5.10, the negative terminals of the separate cells are connected together, as are also the positive terminals. The emf of the battery is, thus, the emf of a single cell. The internal resistance of the battery is r/m ohms. The current through the load is

$$I = \frac{\mathcal{E}}{\left[\frac{r}{m} + R\right]} \qquad [5\text{-}56]$$

This arrangement provides a battery of low internal resistance, which is useful when power is to be delivered to a load of small resistance.

Series-Parallel Connection

Let us suppose that a battery consists of m identical groups of cells connected in parallel, each group consisting of n identical cells in series. This battery will deliver (Problem 5.6) to a load of resistance R a current

$$I = \frac{n\mathcal{E}}{\left[\frac{nr}{m} + R\right]} \qquad [5\text{-}57]$$

This arrangement provides considerable flexibility in the problem of matching battery resistance to load resistance (Problem 5.7).

5.11 Kirchhoff's Laws

In calculating the currents in any but the simplest circuits, use is generally made of two laws which were first stated by the German physicist, Gustav Robert Kirchhoff (1824–1887). These are

(*1*) *The algebraic sum of all currents meeting at any point in a circuit is zero.* That is, at any point

$$\sum_1^n I_i = 0 \qquad [5\text{-}58]$$

Currents flowing toward the point are taken as positive, while those flowing away from the point are taken as negative. This law simply expresses the fact that when a steady state has been reached, as much charge is being carried away from the point as is being carried to it. It is an expression of the conservation of charge.

(*2*) *Around any closed loop in a circuit the algebraic sum of the emf's is equal to algebraic sum of the potential drops.* This law is the result already expressed in equation [5-49]:

$$\sum_1^n \mathcal{E}_i = \sum_1^m I_j R_j \qquad [5\text{-}49]$$

and expresses the fact that the potential differences created by the emf are exactly balanced by the potential drops throughout the circuit. The method by which the necessary algebraic summations are performed is explained in the illustrative example that follows.

Illustrative Example: Let us apply Kirchhoff's laws to the circuit shown in Figure 5.11. The directions of the currents in the three branches of the circuit have been arbitrarily assigned. Those currents that are actually opposite in direction to those assigned will be found, when solved for, to have negative signs.

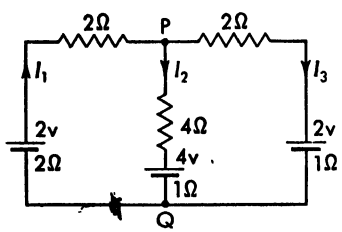


Fig. 5.11

First we note that the circuit contains three unknowns—the currents I_1, I_2, and I_3. Hence, in order to solve for these unknowns, we shall require three independent simultaneous equations involving them.

The first law, if invoked at the point P, gives

$$I_1 - I_2 - I_3 = 0 \qquad [5\text{-}59]$$

The same equation is obtained at the point Q. Two additional equations must be obtained.

The second law may be applied to any closed loop, of which there are three available in this example: the left loop, the right loop, and the outside loop. The two additional equations that are required are obtained by writing equation [5-49] for two of these

loops, naturally choosing (if there is any choice) those loops for which equation [5-49] contains the smallest number of terms. If equation [5-49] is written down for the third loop as well, it will be found to be the sum of the other two and is, therefore, not an independent equation.

In applying Kirchhoff's second law, one moves around the circuit, in, say, the clockwise direction, entering the emf's in equation [5-49] as positive or negative, depending upon whether their polarities are such as to give rise to clockwise or counter-clockwise currents. Similarly, each potential drop is entered as positive if the direction of the current is clockwise, and as negative if it is counter-clockwise. For the left loop, equation [5-49] becomes:

$$2 - 4 = (2 + 2)I_1 + (4 + 1)I_2 \qquad [5\text{-}60]$$

The corresponding equation for the right loop is

$$4 - 2 = (2 + 1)I_3 - (4 + 1)I_2 \qquad [5\text{-}61]$$

After collecting terms and rearranging where necessary, equations [5-59], [5-60], and [5-61] are found to comprise the following set of three simultaneous equations:

$$\begin{aligned} I_1 - I_2 - I_3 &= 0 \\ 4I_1 + 5I_2 &= -2 \\ \text{and} \qquad -5I_2 + 3I_3 &= 2 \end{aligned} \qquad [5\text{-}62]$$

which may be solved for I_1, I_2, and I_3. For example, using the determinant method of solution,

$$I_1 = \frac{\begin{vmatrix} 0 & -1 & -1 \\ -2 & 5 & 0 \\ 2 & -5 & 3 \end{vmatrix}}{\begin{vmatrix} 1 & -1 & -1 \\ 4 & 5 & 0 \\ 0 & -5 & 3 \end{vmatrix}} = \frac{0\begin{vmatrix} 5 & 0 \\ -5 & 3 \end{vmatrix} + 1\begin{vmatrix} -2 & 0 \\ 2 & 3 \end{vmatrix} - 1\begin{vmatrix} -2 & 5 \\ 2 & -5 \end{vmatrix}}{1\begin{vmatrix} 5 & 0 \\ -5 & 3 \end{vmatrix} + 1\begin{vmatrix} 4 & 0 \\ 0 & 3 \end{vmatrix} - 1\begin{vmatrix} 4 & 5 \\ 0 & -5 \end{vmatrix}} = -0.13 \text{ ampere} \qquad [5\text{-}63]$$

Thus, I_1 actually has direction opposite to that indicated in Figure 5.11, and has the magnitude 0.13 ampere.

Number of Equations Required for Solution of a Circuit

Let us generalize some of the comments made in the preceding discussion. But first let us introduce two useful terms. Any point at which three or more circuit elements are connected together is known as a *junction point*. Points P and Q in Figure 5.11 are junction points. The series of circuit elements between two junction points is known as a *branch*. The circuit of Figure 5.11 contains three branches—the left, the middle, and the right.

Suppose a circuit contains j junction points and b branches. The number of unknown currents (one in each branch) is b. Hence, the number of independent simultaneous equations required for the solution of the circuit is also b. Kirchhoff's first law applied to each of the junction points yields j equations, of which only $j - 1$ are independent. Kirchhoff's second law must therefore be written down for $b - (j - 1)$ closed loops.

5.12 Solution of Circuits by Method of Cyclic Currents

In solving a circuit by the method of cyclic currents, Kirchhoff's first law is automatically taken into account, thereby reducing the number of simultaneous equations that must be solved. The method will be illustrated by reference to the circuit (Fig. 5.12) that was solved in the preceding section.

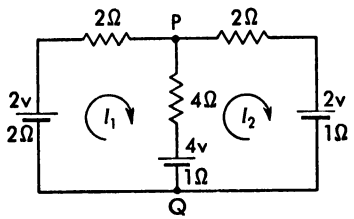


Fig. 5.12

First, a current is assigned to each of the sectional loops, as shown in Figure 5.11. Although it is not necessary to do so, it is convenient to assign the same direction, say, clockwise, to each of the sectional loops in turn.

When applying Kirchhoff's second law to any loop it must be remembered that in any portion of the loop that is also common to another loop the *total* current is the algebraic sum of the two loop currents. Thus, the current from P to Q is $I_1 - I_2$ amperes. With this in mind, let us apply the second law to the left loop,

$$2 - 4 = I_1(2 + 2) + (I_1 - I_2)(4 + 1) \qquad [5\text{-}64]$$

and, similarly to the right,

$$4 - 2 = (I_2 - I_1)(4 + 1) + I_2(2 + 1) \qquad [5\text{-}65]$$

Collecting terms and rearranging equations [5-64] and [5-65], one obtains

$$\begin{aligned} 9I_1 - 5I_2 &= -2 \\ -5I_1 + 8I_2 &= 2 \end{aligned} \qquad [5\text{-}66]$$

which, when solved, give the currents in the several parts of the circuit. In this instance there are only two equations to be solved simultaneously, compared with the three (equations [5-62]) that arose when the cyclic current method was not utilized.

As noted above, Kirchhoff's first law is incorporated in the method of cyclic currents. This is seen by applying the law to one of the junction points,

say P:

$$I_1 - I_1 + I_2 - I_2 = 0 \qquad [5\text{-}67]$$

The assigning of a cyclic current to each of the sectional loops, as in Figure 5.12, guarantees that each circuit element is traversed by at least one current. By applying Kirchhoff's second law to each of the sectional loops, we then have as many equations as we have unknowns. There is no uncertainty as to whether or not all equations are independent.

Illustrative Example: As a further example of the use of cyclic currents, let us calculate the readings of ammeter A and voltmeter V in the circuit of Figure 5.13. The voltmeter is assumed to have infinite resistance and the ammeter zero resistance.

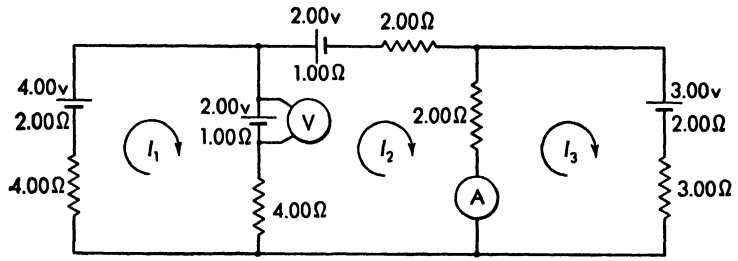


Fig. 5.13

Kirchhoff's second law applied to loops *1*, *2*, and *3* yields, respectively,

$$4 - 2 = 6I_1 + 5(I_1 - I_2) \quad \text{or} \quad 11I_1 - 5I_2 = 2$$

$$2 - 2 = 5(I_2 - I_1) + 3I_2 + 2(I_2 - I_3) \quad \text{or} \quad -5I_1 + 10I_2 - 2I_3 = 0$$

and $$-3 = 2(I_3 - I_2) + 5I_3 \quad \text{or} \quad -2I_2 + 7I_3 = -3 \qquad [5\text{-}68]$$

The solution of these equations is

$$I_1 = 0.185 \text{ ampere}$$

$$I_2 = 0.007 \text{ ampere}$$

and $$I_3 = -0.427 \text{ ampere}$$

Evidently, the reading of the ammeter A is $I_2 - I_3 = 0.434$ ampere.

The reading of the voltmeter V is the sum of two terms. The first of these is the potential difference caused by emf of the battery, namely, 2.00 volts. Then, in addition, account must be taken of the potential drop across the internal resistance of the battery, namely, $(I_1 - I_2)1.00 = 0.18$ volt. Because the direction of the current through the battery is from top to bottom, the upper end of the internal resistance of the cell is 0.18 volt positive with respect to the lower end. The voltmeter thus reads $2.00 + 0.18 = 2.18$ volts.

5.13 Some Electrical Circuit Theorems

Although Kirchhoff's laws provide the means to solve any electrical circuit, it is easier to handle certain problems by using one of several existing network theorems. The more important of these will now be described without formal proof.[5] Together with Kirchhoff's laws, these theorems are widely used in the solution of both DC and AC networks.

Superposition Theorem

The current that flows at any point in a circuit (or the potential difference that exists between any two points in the circuit), resulting from the simultaneous application of a number of emf's distributed throughout the circuit, is the sum of the component currents at the point (or the component potential differences between the two points) that would result from the individual emf's acting separately.

Kirchhoff's second law provides the justification for this theorem. As we have seen in equation [5-49], the branch currents and potential drops in a circuit are *linear* functions of the circuital emf's. It is therefore permissible to calculate the effect of each emf independently.

This theorem is particularly useful in computing, for a circuit that has already been solved, the effect of changing, adding, or removing an emf.

Reciprocity Theorem [6]

If an emf, $\mathcal{E}$, acting in a certain branch of a network, produces a current I in a second branch, then the same emf, $\mathcal{E}$, acting in the second branch, will produce the same current I in the first.

Thévenin's Theorem [7]

Any network containing one or more sources of emf and having two terminals behaves, insofar as a load connected across its terminals is concerned, as though the network were a simple source of emf $\mathcal{E}$ with internal resistance r, where $\mathcal{E}$ is the voltage that appears across its terminals when no load is connected, and r is the resistance that is measured between the terminals when all sources of emf (but not their internal resistances) are short-circuited.

We shall illustrate the use of Thévenin's theorem by solving the circuit

[5] The statements of the theorems are abridged from the *American Institute of Physics Handbook*, McGraw-Hill, 1957, p. 5–81.

[6] Readers interested in the proofs of the reciprocity and Thévenin's theorem may consult G. P. Harnwell, *Principles of Electricity and Electromagnetism*, McGraw-Hill, 1949; or W. T. Scott, *The Physics of Electricity and Magnetism*, Wiley, 1959.

[7] See footnote 6 above.

of Figure 5.11. Let us suppose that we wish to learn the current in the 2-ohm resistor in the upper left portion of the circuit. We then take as the two terminals the points A and B, as indicated in Figure 5.14. We imagine the 2-ohm resistor to be removed in order to calculate, first, the voltage appearing across AB and, second, the resistance between the two terminals. To facilitate this calculation, the circuit, with the 2-ohm resistor removed, is redrawn in Figure 5.15.

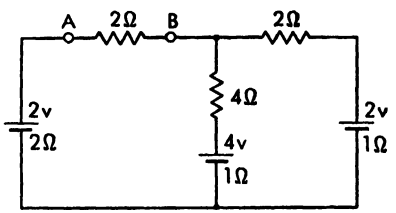


Fig. 5.14

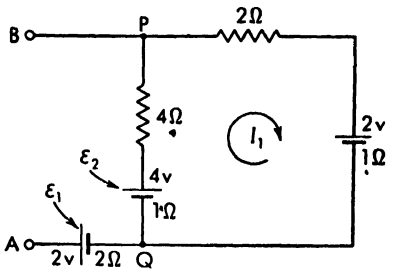


Fig. 5.15

The voltage V between B and A is the algebraic sum of the two emf's $\mathcal{E}_1$ and $\mathcal{E}_2$ minus the potential drop across the resistance in the section PQ. That is,

$$V = -2 + 4 - 5I_1 \qquad [5\text{-}69]$$

where I_1, it will be readily seen, is given by

$$I_1 = \frac{(4 - 2)}{8} = 0.25 \text{ ampere} \qquad [5\text{-}70]$$

Thus, the voltage between B and A is

$$V = 0.75 \text{ volt}$$

The resistance between B and A consists of a 2-ohm unit (the internal resistance associated with $\mathcal{E}_1$) in series with two parallel groups, of resistance 5 and 3 ohms, respectively. The effective resistance of this combination is easily calculated to be

$$r = 3.9 \text{ ohms}$$

Figure 5.15 may, therefore, be replaced by Figure 5.16. We may now imagine the 2-ohm resistor to be connected between B and A, in which case the current in it has the direction from B to A and the value

$$I = \frac{0.75}{3.9 + 2} = 0.13 \text{ ampere} \qquad [5\text{-}71]$$

in agreement with the value found by application of Kirchhoff's laws.[8]

[8] Although what has been done does not constitute a general proof of Thévenin's theorem, the fact that the network of Figure 5.15 can be represented, to an outside observer, by that of Figure 5.16 is proof that Thévenin's theorem is true for this particular case.

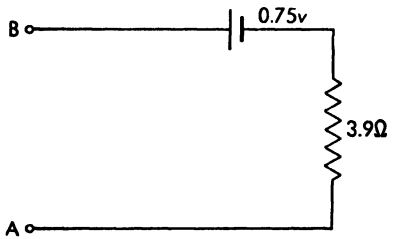


Fig. 5.16

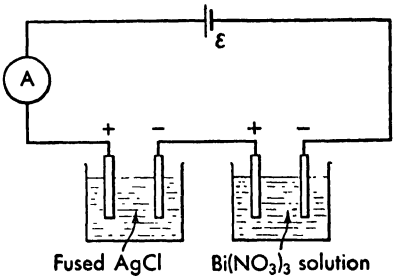


Fig. 5.17

The student could profitably calculate, using Thévenin's theorem, the current flowing in the central 4-ohm resistor in the circuit of Figure 5.11.

Thévenin's theorem is especially useful when studying the efficiency of power transfer. Maximum power transfer from a source of DC electrical energy to a load requires resistance matching of the load to the source. In cases where the source is a complicated network, the first step is to reduce it to its simple $\mathcal{E}$-plus-r equivalent, in the same way as we have reduced the circuit of Figure 5.15 to that of Figure 5.16. The degree of mismatch between the load and the source is then easily recognized.

5.14 Faraday's Laws of Electrolysis

So far in this chapter we have discussed what might be called metallic conduction. In this the current carriers are electrons and their passage produces no change in the chemical composition of the conductor. In nonmetallic conductors the situation is strikingly different. Here the current-carriers are usually atoms or aggregates of atoms comprising the conductor, and the passage of charge results in chemical decomposition. This type of conduction was first observed in the study of the electrical conductivity of solutions and was subsequently shown to occur in fused salts and certain crystals as well.

Let us consider the experiment illustrated in Figure 5.17, where the charge is passing through a sample of fused AgCl in series with a water solution of $Bi(NO_3)_3$. Conduction is possible only because Ag^+ and Cl^- ions are present in the fused silver chloride and Bi^{+++} and $(NO_3)^-$ ions are present in the water solution of bismuth nitrate.

In the fused silver chloride, Ag^+ ions are drawn to the *cathode* (or negative terminal), where they are neutralized and metallic silver is deposited. Cl^- ions move to the *anode* (or positive terminal), where they are neutralized and are then liberated as Cl_2 gas. In the bismuth nitrate solution, Bi^{+++} ions are discharged at the cathode with consequent deposition of metallic bismuth. The

$(NO_3)^-$ ions, on the other hand, react with H_2O, after neutralization, to liberate O_2 and form nitric acid, HNO_3. In brief, as a result of the passage of the same total charge, there is metallic deposition at two electrodes and liberation of gas at the other two. The two laws of electrolysis, first stated by Faraday, enable one to predict the amount of a given element that is liberated during electrolysis. These laws are:

(*1*) *The mass of an element liberated at an electrode during passage of a current is proportional to the total charge passing.* That is,

$$M \propto It$$

or

$$M = zIt \qquad [5\text{-}72]$$

Here z is the *electrochemical equivalent*, defined as *the mass of a substance that is liberated when unit current flows for unit time.* The electrochemical equivalent is, thus, *the mass liberated per unit charge.* For silver $z = 1.118 \times 10^{-3}$ gm/coulomb.[9]

(*2*) *The mass of an element liberated at an electrode is proportional to the gram equivalent weight of the element. The gram equivalent weight* is defined as *the gram atomic weight of the element divided by its valence.* Elements of valence 2, for example, form doubly-charged ions, with the result that only half as much material is liberated by the passage of a given charge as would be the case if the same atom were singly charged.

The atomic weights of the elements liberated in the apparatus in Figure 5.17 are Ag: 107.880, Cl: 35.457, Bi: 209.00, and O: 16. The masses of the four elements that are liberated are, therefore, in the ratio of 107.880 : 35.457 : 69.67 : 8.

Knowing the electrochemical equivalent of silver, we can readily calculate the same quantity for the other three elements involved: for example, for Cl,

$$z = \frac{35.457}{107.880} \times 1.118 \times 10^{-3} = 3.675 \times 10^{-4} \text{ gm/coulomb}$$

and so forth.

The Faraday

Let us consider the charge needed to deposit a certain mass of an element, rather than the mass deposited per unit charge, as above. In particular, let us calculate the charge required to liberate one gram-equivalent weight of the element. For any element, this charge is found by dividing the gram-equivalent

[9] Some of the language used below is essentially chemical language. Hence, to conform with chemical practice, the gram is taken as the unit of mass instead of the kilogram, as in the mks system.

weight by the electrochemical equivalent. Thus, for silver: charge/gm-eq.wgt. = 107.880/0.001118 = 96,500 coulombs: for chlorine: charge/gm-eq.wgt. = 35.457/0.0003675 = 96,500 coulombs.

The reader will realize that this agreement is no coincidence. The charge needed to liberate one gram-equivalent weight is the same for all elements. It is known as the *faraday* and is designated by the symbol F.

$$F = 96{,}500 \text{ coulombs} \qquad [5\text{-}73]$$

The Faraday and Avogadro's Number

Accurate knowledge of the value of the faraday and of the value of the elementary charge enables one to calculate *Avogadro's number*, N, the number of atoms per gram-atomic weight of an element.

For a monovalent element, such as silver, the gram-equivalent and the gram-atomic weights are identical. The charge carried per silver ion is $+e$. Thus, the total charge carried by a gram-equivalent weight of silver is

$$Ne = F \qquad [5\text{-}74]$$

from which one computes:

$$N = \frac{F}{e} = \frac{96.500}{1.6 \times 10^{-19}} = 6.02 \times 10^{23} \text{ atoms} \qquad [5\text{-}75]$$

This is one of the most important methods for the determination of Avogadro's number.

5.15 Determination of Current by Electrolytic Deposition: The Voltameter

The magnitude of a steady current may be established by measuring the amount of material that is liberated by it in a given time. The silver voltameter, shown in Figure 5.18, is an instrument devised for this purpose. It consists of a platinum bowl containing a solution of silver nitrate, into which a silver electrode dips. During charge flow, Ag^+ ions move to the platinum bowl where they are discharged, and metallic silver is deposited. Additional silver ions, as needed, go into solution from the silver anode.

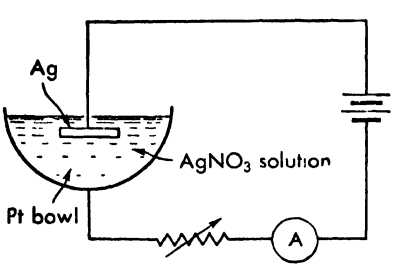


Fig. 5.18

An ammeter to be calibrated is in-

cluded in a series circuit containing the voltameter. After a steady current has existed for a certain known time, the platinum bowl is weighed in order to determine the mass of silver that has been deposited. The current is then calculated from equation [5-72].

For many years the silver voltameter, elaborately described by internationally established specifications, provided an important laboratory method for realizing the ampere.

5.16 Problems

E (5.1) The *Ayrton shunt*, shown in Figure 5.19, is a convenient device for reducing the sensitivity of a current-measuring instrument, such as a galvanometer. It consists of a resistor R which is tapped at some fraction α of its total value. Show that the galvanometer current I_g is related to the total current I, as follows:

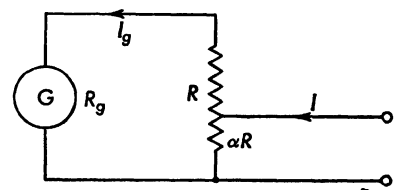


Fig. 5.19

$$I_g = \left[\frac{\alpha R}{(R + R_g)}\right] I$$

where R_g is the resistance of the galvanometer.

E (5.2) An insulated vessel contains a liquid in which a coil of resistance wire is immersed. An emf of 20 volts, applied to the ends of the coil, gives rise to a current of 2 amperes. When thermal equilibrium is established 5 gm of the liquid are boiled away every minute. Calculate the latent heat of vaporization of the liquid.

[Ans.: 115 cal/gm.]

E (5.3) A battery of emf $\mathcal{E}$ volts and internal resistance r ohms is connected to a load of R ohms. Show that the terminal voltage of the battery is $\mathcal{E}[R/(r + R)]$ volts.

E (5.4) A circuit consists of a 10-ohm resistor in series with 6 identical cells, each of emf 1.0 volt and internal resistance 5.0 ohms. Show that the difference in potential between the positive and negative poles of any one cell is 0.25 volt.

M (5.5) An enclosed battery consists of 12 identical cells in series. It is suspected that some of the cells are incorrectly connected. The battery is connected in series with a high-current meter and 2 additional cells that are similar to the others. When the additional cells support the battery, the resulting current is 50 percent greater

than when they oppose the battery. Show that this is evidence that one of the cells in the battery is connected with the wrong polarity.

E (5.6) Show that a battery consisting of m identical groups of cells connected in parallel, each group consisting of n identical cells in series, will deliver to a load of R ohms a current

$$I = \frac{n\mathcal{E}}{\frac{nr}{m} + R} \text{ amperes}$$

where $\mathcal{E}$ and r are the emf and internal resistance, respectively, of a single cell.

E (5.7) How should a battery of 36 identical cells, each having internal resistance of 1.4 ohms, be constructed so as to achieve maximum power transfer to a load resistance of 5.6 ohms?

E (5.8) A certain storage battery has a terminal voltage of 1.95 volts when delivering a current of 5.00 amperes, and 1.92 volts when delivering a current of 8.00 amperes. What is the potential difference across the battery terminals when it is being "charged" with 4.00 amperes current? [Ans.: 2.04 volts.]

E (5.9) Determine the current I in the circuit shown in Figure 5.20. [Ans.: 0.73 ampere.]

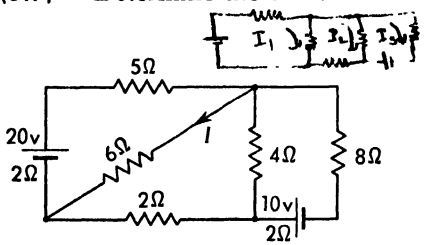


Fig. 5.20

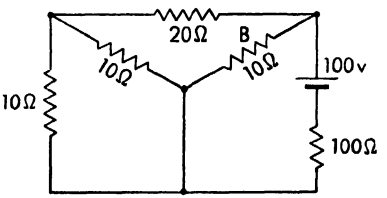


Fig. 5.21

E (5.10) Calculate the potential drop across resistance B in the circuit of Figure 5.21. [Ans.: 6.7 volts.]

M (5.11) Find the current in each branch of the circuit shown in Figure 5.22. Determine also the potential difference across each cell.

M (5.12) A battery consists of three cells in parallel, as shown in Figure 5.23. What is the potential difference between the terminals A and B of the battery? [Ans.: 2.11 volts.]

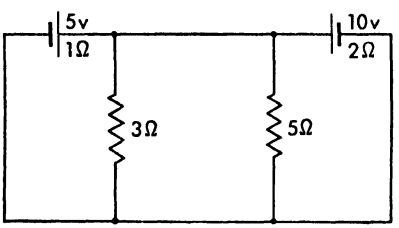


Fig. 5.22

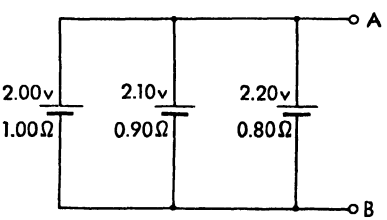


Fig. 5.23

H (5.13) Calculate the resistance between the points A and B in Figure 5.24. Also compute the resistance between A and C.

[Ans.: $8R/15$, $2R/3$.]

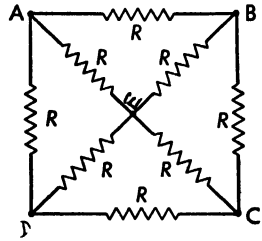


Fig. 5.24

M (5.14) Calculate the ammeter reading, A, in the circuit of Figure 5.25 and also the potential drops across R_1, R_2, C_1, and C_2.

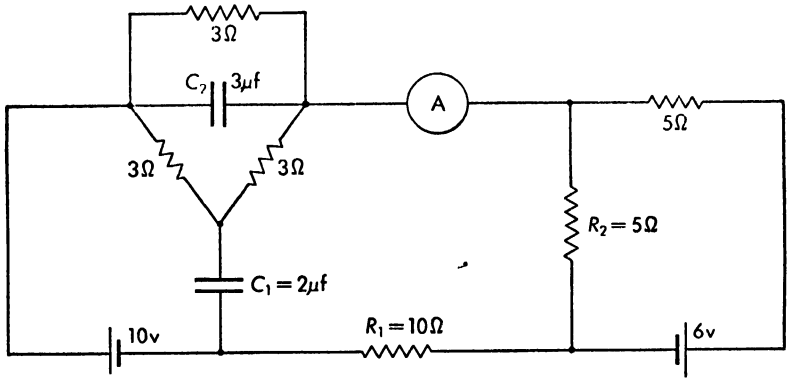


Fig. 5.25

E (5.15) Zinc has stable isotopes at mass numbers 64, 66, 67, 68, and 70. If 5 faradays of electricity deposit 163.45 gm of zinc, what is the valence of zinc?

E (5.16) How many ampere-days are needed to deposit a pound-equivalent weight of a metal in an electrolytic refining bath?

Chapter 6

Resistance

6.1 The Band Theory of Solids

The behavior of the valence electrons in an atom is strongly dependent upon the environment in which the atom is found.

In a free atom the electrons are arranged in shells around the nucleus, each occupying its own particular energy state. The electrons in the inner shells are strongly bound to the atom, while those in the outermost, or *valence*, shell may be removed from the atom with the expenditure of only a few electron-volts energy.

When atoms are close together, as in a crystalline solid, the electrons in the inner shells are not appreciably affected, but the charge clouds of the valence electrons overlap to a high degree. Individual valence electrons can no longer be thought of as belonging to individual atoms. Instead, they are shared on a more or less equal basis by all the atoms comprising the crystalline structure. Further, *bands* of energy, consisting of a distribution of energy levels, are now available to them, in place of the discrete energy states which they had previously occupied in the isolated atom. These *allowed* or *permitted* energy bands are normally separated by *forbidden* ranges of energy.

6.2 The Fermi-Dirac Distribution

A *free electron gas*, such as that formed by the conduction electrons in a conductor, possesses an energy distribution that is markedly different from that of an ordinary gas.

In the case of an ordinary gas in thermal equilibrium, the molecules have a range of energies given by the Maxwell-Boltzmann distribution function, shown in Figure 6.1 for two different temperatures. When the temperature of the gas is increased, a pronounced shift toward higher energies occurs in the distribution function. A rise in temperature therefore corresponds to a significant increase in the total energy of the gas. The student may recall in this connection that the specific heat of a monatomic gas at constant volume is $c_v = 3R/2$ [1] per mole.

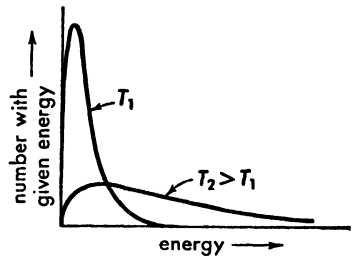


Fig. 6.1

Electron Spin

An important property of an electron, which we have not yet mentioned, is its *spin*, or *intrinsic angular momentum*. This may be caused by the electron spinning about an axis within itself, or by some other mechanism not yet suggested. Regardless of the mechanism, this intrinsic angular momentum is an indestructible property of electrons. It is present whether the electron has *orbital angular momentum* or not. In addition, there is associated with it a certain *magnetic moment*, which we shall have occasion to discuss further in Chapter 10.

Angular momenta are predicted by wave mechanics to have values of $\sqrt{I(I+1)}\,h/2\pi$, where I is integral or half-integral, and h is Planck's constant. The spin of the electron corresponds to $I = \frac{1}{2}$. All particles, such as the electron, which possess *half-integral spin* are known as *fermions*, and are described by the *Fermi-Dirac* statistics, rather than by the *Maxwell-Boltzmann* statistics, which describe the behavior of gas molecules. Incidentally, other fermions are the proton and the neutron, for each of which I is also equal to $\frac{1}{2}$. The characteristic feature of a system of fermions is that each energy level, when filled, contains two particles, one "spinning" in one direction and the other in the opposite direction—their spins are said to be *anti-parallel*. The two electrons in an atom of helium constitute such a pair, while an atom of neon has five such pairs of electrons, one in the first shell and the remaining four in the second. This description of fermions is covered by the statement, which will not be elaborated here, that fermions obey the *Pauli exclusion principle*.

Energy Distribution for a Gas of Fermions

The energy distribution for a gas of fermions is shown in Figure 6.2. The solid curve gives the distribution at absolute zero ($T = 0$). Here, all possible

[1] R is the *gas constant*, the factor of proportionality in the familiar equation of state ($pV = RT$) for an ideal gas.

energy states below $E = E_0$ are filled and all those above $E = E_0$ are empty. As the temperature is raised, some of the particles in the higher energy states are excited beyond $E = E_0$, giving rise to the slightly changed distribution shown by the dotted curve. There is on the whole, however, very little increase in the *total* energy of the system, even at elevated temperatures.

Electron Gas Is a Fermi Gas

The valence electrons in a solid comprise a *Fermi* gas and have an energy distribution given by Figure 6.2. Consequently, the specific heat at constant

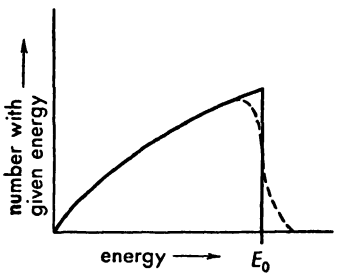


Fig. 6.2

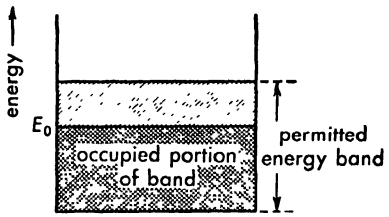


Fig. 6.3

volume of an electron gas is practically zero, in contrast to the value of $3R/2$ mentioned above for an ordinary monatomic gas. The energy E_0, known as the *Fermi energy*, is measured from the bottom of the last-occupied *energy band* (Fig. 6.3), in tacit recognition of the fact that deeper-buried electrons are not "free."[1]

At absolute zero, E_0 is the energy height of the filled portion of the band: at room temperatures, the energy of the last-populated level is only slightly above E_0 (cf. Fig. 6.2).

6.3 Insulators, Conductors, and Semiconductors

The band theory of solids provides the explanation for the differences in electrical conductivity among insulators, conductors, and semiconductors.

[1] Figure 6.3 is an example of an energy-level diagram. On such a diagram the various permitted energy states are normally shown as horizontal lines whose vertical positions represent the energies belonging to each. The electronic capacity of each level is not indicated on such a diagram. A band of energies would be represented by a very large number of closely spaced lines: in effect, a continuous energy distribution, as is shown in the diagram. Figure 6.3 represents Figure 6.2 (disregarding the number of electrons with each particular value of energy) turned through 90° and superimposed on a background of the available energy levels.

Insulators versus Conductors

In an insulator the valence electrons are just sufficient in number to fill one of the permitted bands of energy (Fig. 6.4). In these circumstances it

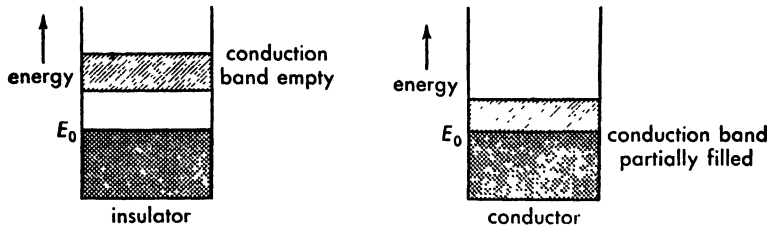


Fig. 6.4 Forbidden bands are unshaded; allowed bands are hatched, with occupied portions cross-hatched.

is not possible for an external field to give rise to a current. If one of the electrons in the band were to be accelerated, thereby gaining energy, it is inevitable, since the energy band is full, that the energy state formerly occupied by this electron would be occupied by a previously more energetic electron. The total electronic energy distribution would thus be identical with that existing before the application of the field. As many electrons would still be moving in one direction as in another, and the net current would be zero.

In a conductor, such as a metal, the last-occupied, allowed energy band is only partially filled (Fig. 6.4). This band is known as the *conduction band*, and it is entirely feasible to accelerate the electrons in it without the occurrence of a compensating effect of the type described above for insulators. For example, in the case of the alkali metals, where the number of valence electrons per atom is one, the conduction band is only half filled. To a good approximation, conduction electrons can be considered as quasi-free (as was done in § 5.2), although their *effective mass* may differ from that of free electrons.

Cause of Resistance in a Conductor

In the metallic crystal the atoms are arranged in a regular array known as the *crystal lattice*. If the lattice were perfect, conduction electrons would pass through it without encountering resistance, in the same way as light passes through a perfect crystal without scattering or absorption. The electrical resistance of an actual metal is due to deviations from a perfect lattice caused by *impurity atoms*, *lattice defects*, and, usually most important, by *thermal vibrations of the atoms comprising the lattice*. This matter will be further touched upon in § 6.5.

Semiconductors

A semiconductor is an interesting blending of insulator and conductor. It is an insulator at absolute zero but becomes partially conducting as the temperature is raised.

We shall consider first the so-called *intrinsic semiconductors*. In these materials the energy gap between the last filled band (at $T = 0$) and the next allowed band is small enough that, when $T > 0$, electrons can be excited across the gap by means of thermal energies alone (Fig. 6.5a). In the resulting electron

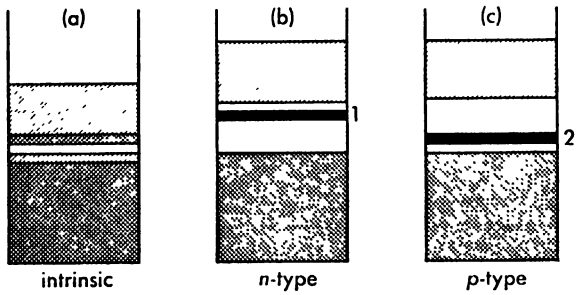


Fig. 6.5 Types of semiconductor. Forbidden bands are unshaded; allowed bands are hatched, with occupied portions cross-hatched. Levels *1* and *2* (blacked in) in (b) and (c), respectively, are energy levels provided by impurities.

distribution a few electrons are found at the bottom of an almost empty band and a few vacancies or *holes* at the top of an almost filled band. Both the excited electrons and the holes they leave behind give rise to conduction. The motion of a hole is similar to the motion of a bubble in a liquid. In the case of the bubble, a portion of the liquid is pulled by gravity into the volume occupied by the bubble, thus displacing the bubble upward, where the process is repeated. Similarly, when an electron moves, say, to the left, into a vacancy or hole, the hole is displaced to the right, where it is displaced further to the right by a second electron, and so on. Holes thus act as though they were positive charges, moving in the direction of the applied field. Conduction by electron-plus-hole motion is an example of the type of situation described by equation [5-5], viz.,

$$J = J' + J'' = n'q'v' + n''q''v'' \qquad [5\text{-}5]$$

where the current densities corresponding to the two motions are in the same direction.

In certain other types of crystals the semiconducting properties are explained by the presence of impurities. These impurity atoms provide new energy levels which lie between the last-filled and the next-empty (or conduction) bands of the matrix material, as shown in Figure 6.5b and c. In the

first of these, the impurity energy levels (1) lie close enough to the conduction band to be able to supply electrons to it, via thermal excitation. These impurity levels are known as *donor* levels. Conduction is then by means of electrons, and the semiconductor is known as an *n-type* (*n* for negative). In the second type, shown in Figure 6.5c, the impurity levels (2) lie close to the filled band and are able to accept electrons which are thermally excited from it. Thus, they provide *acceptor* levels, with the result that conduction is due to the motion of holes in the normally filled band. Such a semiconductor is known as a *p-type* (*p* for positive).

The conductivity evinced by impurity-type semiconductors is strongly dependent upon the crystal structure of the matrix material and the amount and type of impurity that it contains. Also, because it depends upon thermal excitation all semiconductor conductivity is strongly temperature-dependent.

6.4 Resistivity and Its Use in the Calculation of the Resistance of Ohmic Materials

It was shown in § 5.3 that the resistance which a sample of material offers to the flow of charge is proportional to the length of the sample and inversely proportional to its area. Also, of course, it depends upon the specific conducting properties of the material itself. These dependences are expressed as follows:

$$R = \rho \int \frac{dl}{A} \qquad [6\text{-}1]$$

where dl and A are the length and cross-sectional area of an element of the sample under consideration, and ρ is the *resistivity* of the material comprising it. For ohmic materials, as indicated in equation [6-1], ρ is constant, independent of the magnitude of the current which the material is carrying. It is seen from equation [6-1] that resistivity in the mks system has units of ohm-meters and is numerically equal to the resistance in ohms between opposite faces of a cube of material 1 meter on a side (see also below).[3]

The reciprocal of resistance is known as *conductance* and is designated by the symbol G. The unit of conductance in the mks system is, appropriately enough, the *mho*—an etymological triumph that ranks with one of the early

[3] Electrical engineers frequently express the resistance of wires in *ohms per mil foot*. In this system ρ is the *resistance in ohms of a round wire of length 1 foot and diameter 0.001 inch, or 1 "mil."* Such a wire is said to have a cross-sectional area of 1 *circular mil*. The cross-sectional area of a round wire of diameter d inches is then $(d \times 10^3)^2$ circular mils. For a certain alloy, for example, $\rho = 275$ ohms per (circular) mil foot. Then the resistance of 100 yards of wire (diameter 0.010 inch) of this alloy is

$$R = \frac{\rho \times l \text{ (in feet)}}{A \text{ (in circular mils)}} = \frac{275 \times 300}{(10)^2} = 825\Omega$$

(now outmoded) names given to ∇, "atled." The *conductivity* of a material, designated by the symbol σ, is defined as *the reciprocal of its resistivity*. Equation [6-1] may now be rewritten as

$$G = \frac{1}{R} = \frac{\sigma}{\int \frac{dl}{A}} \qquad [6\text{-}2]$$

Clearly, σ has mks units of mhos/meter.

The concept of conductance is especially useful in dealing with conductors that are in parallel. Since the reciprocal of the total resistance of such a group is the sum of reciprocals of the separate resistances, it follows that the total conductance is the sum of the separate conductances. On the other hand, conductances in series add according to the reciprocal rule.

Table 3 shows the wide differences in resistivity that exist between good metallic conductors and good insulators. The resistivities of mica and silver,

TABLE 3

Resistivities of Some Typical Conductors and Insulators

Material	*Resistivity, ρ (ohm-m) at 20° C*
Metals	
Silver	1.6×10^{-8}
Copper	1.7×10^{-8}
Aluminum	2.8×10^{-8}
Tungsten	5.5×10^{-8}
Steel (4% Si)	6.2×10^{-7}
Manganin	4.4×10^{-6}
Nichrome	1.0×10^{-6}
Insulators	
Mica	9×10^{12}
Quartz crystal II	1×10^{12}
Quartz crystal I	3×10^{14}
Diamond	3×10^{11}–5×10^{12}

for example, differ by a factor of 5×10^{20}. On the other hand, among metallic conductors themselves the range is only a couple of orders of magnitude.

Calculation of the Resistance of Conductors Having Simple Geometric Shapes

By means of equation [6-1] the total resistance of many shapes of conductors can readily be calculated. One first chooses an element of the conductor whose

cross-sectional area is normal to the current density. The resistance of such an element has the value

$$dR = \rho \frac{dl}{A} \qquad [6\text{-}3]$$

The total resistance of the conductor is then the integral of equation [6-3] evaluated between the proper limits.

The calculation of the resistance between opposite ends of a rectangular conductor (Fig. 6.6) is especially simple. In practice, the end faces of such a conductor would be clamped between blocks of low resistivity material so that the current density within the conductor would have the direction shown. Evidently, the total resistance is

$$R = \int dR = \frac{\rho}{A} \int_0^l dl = \rho \frac{l}{A} \qquad [6\text{-}4]$$

When this expression is rewritten in the form

$$\rho = \frac{RA}{l} \qquad [6\text{-}5]$$

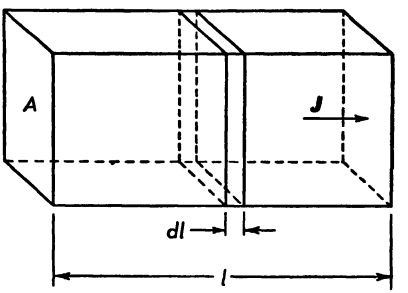


Fig. 6.6

it confirms the statement made earlier, that the resistivity of a material is numerically equal to the resistance between opposite faces of a unit cube of the material.

Problems 6.3 and 6.4 provide examples of the use of equation [6-3] in calculating the resistance of other-shaped conductors.

6.5 Variation of Resistance with Temperature

Metals

The comment has already been made in § 6.3 that a perfect metal crystal would display infinite conductivity. Through such a lattice the waves belonging to the moving conduction electrons (recall equation [2-52]) would pass without interference.

Unfortunately, perfect crystals do not exist. In actual metal crystals, vibrations of the crystal lattice constitute the chief source of interference. Electrons collide with the vibrating lattice, losing energy to it, being reflected or refracted by it, and so forth. An elegant way of describing these events is to say that the electrons collide with phonons, these being the energy quanta that are associated with lattice vibrations.

When $T \gg 0$ the energy associated with vibrations of the crystal lattice is

directly proportional to the absolute temperature. It is, therefore, to be expected that the resistance due to this effect will have a similar dependence upon temperature. The observed fact that the resistance of pure metals is approximately proportional to T is evidence that the contributions of other sources of interference, such as impurities, dislocations, and imperfections, are small.

The structure of alloys is often disordered. This disorder results, as expected, in a larger resistivity than is characteristic of a pure metal. The contribution of lattice vibrations is, in this case, a small fraction of the total resistance, which is consequently not strongly dependent upon temperature.

Insulators and Semiconductors

The band structure of insulators and semiconductors is such that at 0° K the conduction band is bereft of electrons. In semiconductors, at room temperature, electrons are raised to the conduction band as a result of thermal energy acquired from the lattice vibrations. It is to be expected, therefore, and it is found, that the resistance of semiconductors *decreases* with increasing temperature. At sufficiently elevated temperatures the lattice vibrations are energetic enough even to excite some electrons into the conduction band of insulators. For many insulators the resistance is relatively insensitive to temperature in the room-temperature range. But when the temperature is raised to the point at which an effect is observed, the resistance always decreases with increasing temperature.

Temperature Coefficient of Resistance

The variation with temperature of resistance is described by a *temperature coefficient*, designated by the symbol α and defined by the following relationship:

$$\alpha = \frac{R_{t_2} - R_{t_1}}{R_{t_1}(t_2 - t_1)} \qquad [6\text{-}6]$$

Here R_{t_2} and R_{t_1} are the resistances of the sample at t_2° C and t_1° C, respectively. The temperature coefficient of resistance is thus the fractional change in resistance per degree Centigrade change in temperature. It is always evaluated at some particular base temperature, t_1, usually 0° C or 20° C, and in a certain stated temperature range. For example, if α is evaluated in the range 0–100° C, we have

$$\alpha_0 = \frac{R_{100} - R_0}{100 R_0} \qquad [6\text{-}7]$$

where the symbol α_0 indicates that 0° C is the base temperature. The re-

sistance at any other temperature t is given by

$$R_t = R_0[1 + \alpha_0(t - 0)] \qquad [6\text{-}8]$$

Equation [6-8] is correct only if the relationship between resistance and temperature is linear. For pure metals it is approximately linear in the range $0° \text{ C} < t < 200° \text{ C}$, and equations of the form of [6-8] can be used to compute reasonably accurate values of resistance within the temperature range over which α has been evaluated. For alloys the variation of resistance with temperature is erratic, and an equation of the form of [6-8] is valid only over a small temperature range.

TABLE 4

Values of α_0 for Some Pure Metals and Some Alloys

Metal	α_0
Silver	0.0040
Gold	0.0038
Zinc	0.0037
Platinum	0.00367
Nichrome	0.0004
Manganin	0.00001

The values of α_0 in Table 4 bear out the earlier assertion that the resistance of many pure metals is approximately proportional to the absolute temperature. Exact proportionality would require α_0 to have the value $\frac{1}{273} = 0.00366$. This relationship does not hold even approximately at low temperatures.

We have noted above that the resistance of an alloy is relatively insensitive to lattice vibrations and should, therefore, show little temperature dependence. In keeping with this argument, the α_0's for the alloys listed in Table 4 are appreciably smaller than those for pure metals. The two alloys listed, nichrome and manganin, are useful in the construction of resistances whose values must show little change with temperature. Manganin (84 percent copper, 12 percent manganese, 4 percent nickel) is particularly well suited for this purpose.

A more exact expression than [6-8] for the resistance-temperature curve is given by the second (or higher) degree equation,

$$R_t = R_0(1 + \alpha t + \beta t^2 + \cdots) \qquad [6\text{-}9]$$

We shall make use of this expression in discussing the platinum resistance thermometer (§ 6.11).

6.6 Superconductivity

It was discovered in Leiden in 1913 by the Dutch physicist Kamerlingh Onnes (1853–1926) that the resistivity of mercury completely vanishes below 4.2° K (Fig. 6.7). The transition region between normal and superconducting mercury is very narrow (~0.05° K). One of the most striking experiments done by Kamerlingh Onnes was to cool a ring of superconducting material through its transition temperature while it was in a magnetic field. The removal of the magnetic field induced in the superconducting ring a current which persisted undiminished for days.

Twenty-odd metallic elements and a large number of compounds and alloys are now known to display this bizarre effect, each below its own *critical temperature.* Unexpectedly, it is found that all superconducting metals are only fair conductors of electricity at room temperature. None of the very good conductors, such as silver or copper, has yet been made superconducting. The highest transition temperature that has been observed so far is 18° K—for a tin-niobium alloy. The time is not far distant when superconductors, maintained at the necessary temperature by liquid helium, will find practical use, especially to transmit weak electrical signals without loss of energy.

The full explanation for superconductivity has yet to be given. Undoubtedly, the fact that normally mediocre conductors become superconducting while normally good conductors do not is a vital clue to it. Because the interaction between electrons and lattice vibrations is greater for a poor than for a good conductor, it is thought that this interaction is responsible for the effect.

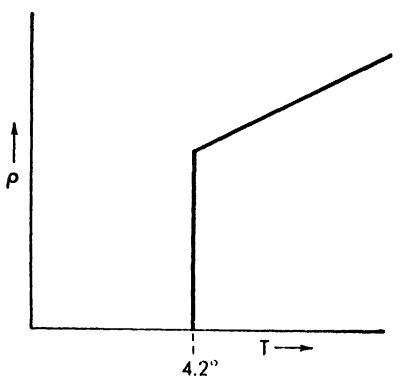


Fig. 6.7

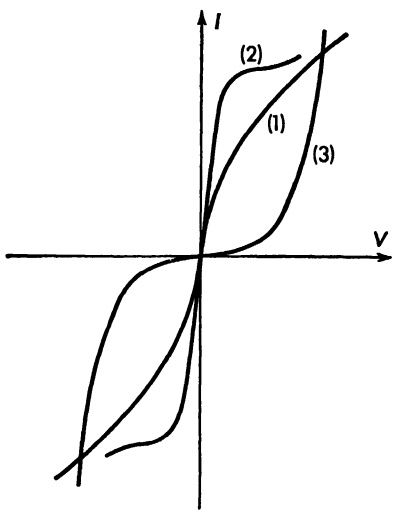


Fig. 6.8

Presumably it is possible, at low temperatures, for the lattice vibrations and the electron waves to be synchronized in such a way that the resistance to electron flow disappears. The subject, however, is beyond the scope of this book.

6.7 Nonohmic or Nonlinear Resistances

The resistance of a circuit element is defined as

$$R = \frac{V}{I} \qquad [6\text{-}10]$$

where V is the potential difference across the element and I is the current which it carries. If R is a constant, the element is said to be ohmic: that is, a plot of I versus V is a straight line of slope R (obeys Ohm's law). Metallic elements are ohmic, provided the temperature remains constant.

When R is not a constant, the element is said to be *nonohmic* or *nonlinear*. Such elements are both common and important, although the reverence paid to Ohm's law in elementary courses often conceals this fact. The resistance behavior of nonohmic elements is usually presented as an I-versus-V plot, known as a *characteristic*. When V is positive, I is positive, and the characteristic lies in the first quadrant; when V is reversed, I is negative (if there is any value), and the characteristic lies in the third quadrant. If these first- and third-quadrant characteristics are similar in form, the element is said to be *bilateral*, or *symmetric*. If the two characteristics are dissimilar in form, the element is said to be *unilateral*, or *asymmetric*. We shall discuss, in a qualitative way, a few representative examples of these two classes of nonlinear circuit elements.

Some Examples of Bilateral Nonlinearity

Although the I-versus-V characteristic of a metal is a straight line at constant temperatures, this condition is often neither desirable nor possible. For example, the metal comprising the *filament of an incandescent lamp* is deliberately raised to a temperature well above that of the room. This is done by passing an electric current through it. Two effects then occur. First, because of the resistance of the filament, heat is produced in the amount of $0.24\ I^2R$ calories per second, causing the temperature of the filament to rise. Second, because of the temperature rise, the resistance of the filament increases. Consequently, as the potential across the lamp is increased, the current increases less than proportionately. The resulting characteristic, which is symmetric, is shown as Curve *1* of Figure 6.8. For each value of current through the filament a state of dynamic equilibrium is reached in which as much heat

is lost from the filament (by radiation and by conduction through the inert gas within the glass envelope of the lamp) as is supplied by the current.

The slope of the characteristic can be greatly altered from that of Curve *1* by proper choice of the filament material and the surrounding gas. The so-called *ballast lamp*, whose filament is made of an alloy of iron and is surrounded by an atmosphere of hydrogen, has a characteristic of the form of Curve *2*. Over a rather broad range, the current through the lamp is practically independent of the potential across it. When such a lamp is placed in series with a load and a fluctuating voltage supply, the current in the circuit—and, consequently, the potential drop across the load—is maintained approximately constant. Ballast lamps are used in this way to reduce the effect of line-voltage fluctuations.

The characteristic shown as Curve *3* belongs to the material known as *thyrite*, formed by firing at high temperature a pressed-together mixture of silicon carbide and a ceramic binder. The current through this material varies as a power of the applied potential, viz.,

$$I = kV^n \qquad [6\text{-}11]$$

where k is a constant and n has values ranging between 3.5 and 7. If $n = 7$, a doubling of the voltage corresponds to a 128-fold increase in current. Suppose a thyrite resistor is placed in parallel with a load, R_L, and the combination is placed in series with another resistor, R. When the voltage across this entire unit changes, practically all the voltage change appears across R. Thus, the current through R_L is stabilized despite voltage fluctuations.

Thermistors are formed by embedding certain metallic (Mn, Fe, Co, Ni, etc.) oxides in ceramic binders and heating them at temperatures sufficiently high to form hard beads. The temperature coefficient of resistance of a thermistor is negative (as it is for other semiconductors—see § 6.5), but is unusually large, of the order of a few percent per degree Centigrade. The unusual shape of its *I*-versus-*V* characteristic (Problem 6.7) is explicable in terms of its negative temperature coefficient. Because of their large temperature coefficients, thermistors are used as thermometers and for temperature control. A novel use is that of absorbing microwave energy in, say, a wave guide. This absorption is most efficient when the resistance of the thermistor matches the effective resistance of the source of microwave energy. The current through the thermistor can be adjusted until the matched condition has been established. Thermistors are also used as time delays: when the current passing through a thermistor is suddenly altered, a period of time elapses before the new equilibrium temperature is reached and some associated thermal switch is actuated.

Some Examples of Unilateral Nonlinearity

A *rectifier* is a circuit element that offers a high resistance to the passage of charge in one direction, but a low resistance to the passage of charge in the opposite direction. A *vacuum-tube rectifier*, or vacuum *diode*, consists of two electrodes in a vacuum envelope. One electrode is heated to the point at which it emits electrons: these serve as the current-carriers through the tube. The other electrode, which is simply a plate, serves as the collector of electrons. When the second electrode is positive with respect to the first, charge flows through the tube, but not when the potential is reversed. In the one case, the I-versus-V characteristic lies in the first quadrant, but in the second case it practically coincides with the negative V axis. Many *gas-filled diodes* are also in use, ranging from the *tungar rectifier* (filled with argon at a pressure of several centimeters of mercury), which is used for low-voltage applications such as battery charging, to the several varieties of *phanotron* (containing mercury vapor or inert gases at pressures less than 0.1 mm of mercury pressure), which is used for medium- to high-voltage applications. Their characteristics are similar in form to those of vacuum diodes. In all rectifiers the negative portion of the characteristic can be made to depart from the negative V axis by making the plate so negative with respect to the emitter that a discharge occurs.

In *metallic rectifiers*, the contact between a metal and a semiconductor serves as a highly asymmetric resistance. Copper oxide on copper or selenium on iron are two such combinations which are capable of passing large currents in the forward direction.

In the *crystal diode*, the rectifying junction is the point of contact between a sharp metal point (the "cat's whisker") and a crystal of silicon or germanium. Because of their low capacitance, these elements are useful in communications work, where frequencies are high but currents are small. The shape of the characteristic of a type IN34 germanium crystal diode is shown in Figure 6.9.

Finally, a *voltage regulator* (VR) *tube* is a two-element, cold-cathode tube containing inert gas at low pressure. When the potential between the plate and cathode reaches a certain value, known as the *striking potential*, a glow discharge is established in the tube. The voltage drop across the tube is thereafter practically independent of the current carried by the tube as long as the glow covers only part of the cathode. This unique characteristic extends typically over the range

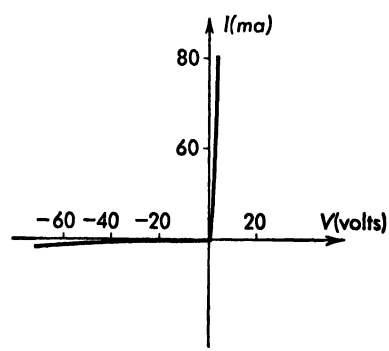


Fig. 6.9

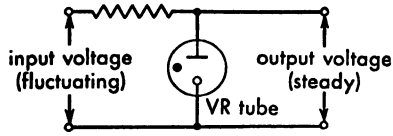


Fig. 6.10

$5 < I < 30$ ma, and makes the tube extremely useful for voltage stabilization, as shown in Figure 6.10. The exact voltage drop across the tube depends upon the cathode material and the gas, but usually lies in the range $75 \leq V \leq 150$ volts.

6.8 Voltmeters and Ammeters: Their Use in the Determination of Resistance

In Chapter 11 the various types of voltmeter and ammeter will be described. In the meantime, in order to discuss their use in the determination of resistance, a few remarks will be made concerning the action of these instruments.

Ammeter

The ammeter measures the total current flowing through it. It is, therefore, inserted in the circuit as a *series* element, in the manner indicated in Figure 6.11. The ideal ammeter would have zero resistance and could be inserted in the circuit without in any way disturbing it. Actual ammeters have some resistance but, in general, are constructed to have as little as possible.

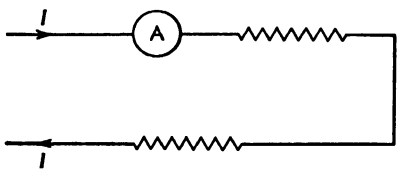


Fig. 6.11

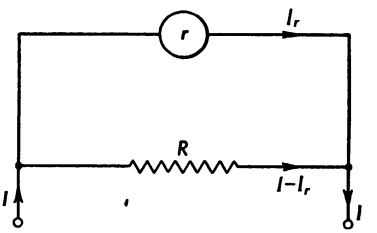


Fig. 6.12

A range of DC ammeters is usually obtained by appropriately shunting a certain basic meter. Let us suppose (Fig. 6.12) that the basic meter has a resistance of r ohms and a full-scale reading of I_r amperes. When this is shunted by R ohms, the basic meter registers full-scale deflection when a total current of $I - I_r$ amperes passes through the shunt. Applying Kirchhoff's second law to the loop consisting of the basic meter and its shunt, we find

$$I_r r - (I - I_r)R = 0 \qquad [6\text{-}12]$$

or

$$I = I_r \left[\frac{r + R}{R}\right] \text{ amperes} \qquad [6\text{-}13]$$

For example, a meter that reads 50 microamperes full-scale and has a resistance of ~2000 ohms is sometimes used as the basic element. From equation [6-13]

it is seen that a shunt of 0.1 ohms transforms this meter into one that reads 1 ampere full-scale. Evidently, it is an easy matter to obtain high-amperage meters that have negligible resistance. On the other hand, the resistance of low-amperage meters is often not negligible and allowance must be made for it.

Voltmeter

The voltmeter records the potential difference between two points, and is inserted in the circuit in *parallel* with the circuit element across which the potential difference which is to be measured exists (Fig. 6.13). Ideally, it should have an infinite resistance, in which case none of the current flowing through the circuit would be sidetracked through it. Actual instruments, of course, fall short of this perfection.

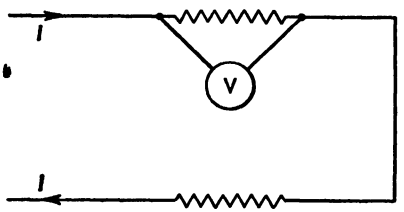


Fig. 6.13

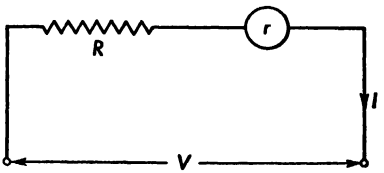


Fig. 6.14

The usual type of voltmeter consists of a sensitive ammeter in series with a large resistance (Fig. 6.14), known as a *multiplier*. The small current that flows through this combination when a potential difference is applied across it is proportional to the potential difference, that is

$$I = \frac{V}{r + R} \qquad [6\text{-}14]$$

Ranges of DC voltmeters are usually constructed from the same basic meter that is used for ammeters. Thus, a 100-volt voltmeter is formed from the 50-microampere, 2000-ohm meter mentioned above by adding to it a series resistor of 2×10^6 ohms.

A figure of merit for a voltmeter is its *ohms per volt*, defined as the *resistance of the voltmeter divided by its full-scale voltage reading*.[4] The voltmeter just referred to is a 20,000 ohms/volt instrument (that is $[2 \times 10^6]/100$), and is a better-than-average instrument. Some voltmeters available commercially have a 100,000 ohms/volt rating and consequently have very little shunting effect on the portion of the circuit across which they are connected.

Brief reference should be made to two other types of voltmeter, the electrostatic voltmeter and the vacuum-tube voltmeter. The former, which has al-

[4] Clearly, the ohms per volt rating is the reciprocal of the current necessary to produce full-scale deflection of the basic meter.

ready been discussed in § 3.11, has an infinite resistance, but is not useful for the measurement of potential differences that are less than 100 volts. The vacuum-tube voltmeter normally has a resistance of 10^7 ohms, even for low-voltage full-scale readings. Thus, for a 1-volt full-scale reading, it is a 10^7 ohms/volt instrument. This value decreases with increasing full-scale voltage reading.

Voltmeter-Ammeter Method for Determination of Resistance

A rough-and-ready method for determining the value of a resistance is to measure the voltage drop across it as it carries a known current (Fig. 6.15). If the voltmeter were an ideal one, possessing infinite resistance, the total current, I, measured by the ammeter, would pass through the unknown resistance, R. The value of R would then be given by

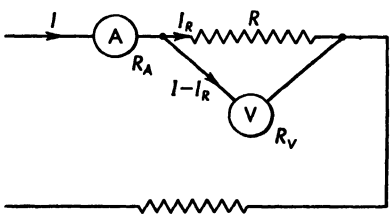


Fig. 6.15

$$R = \frac{\mathrm{V}}{I} \text{ ohms} \qquad [6\text{-}15]$$

where V is the voltmeter reading.

In practice, because of the finite resistance of the voltmeter, a portion of the total current passes through it, as indicated in Figure 6.15. In terms of the symbols shown in Figure 6.15, the value of the unknown resistance is given by

$$R = \frac{\mathrm{V}}{I_R} \text{ ohms} \qquad [6\text{-}16]$$

The current, I_R, that flows through the resistance is obtained by applying Kirchhoff's second law to the V-R loop:

$$I_R R - R_V(I - I_R) = 0 \qquad [6\text{-}17]$$

or

$$I_R = I\frac{R_V}{R + R_V} \qquad [6\text{-}18]$$

Upon substitution of equation [6-16], equation [6-18] yields:

$$R = \frac{\mathrm{V}R_V}{(IR_V - \mathrm{V})} \qquad [6\text{-}19]$$

from which the value of R may be written in the form

$$R = \frac{\dfrac{\mathrm{V}}{I}}{\left[1 - \dfrac{\mathrm{V}}{IR_V}\right]} \qquad [6\text{-}20]$$

It can be seen from equation [6-20] that when $R_V \gg R$ the fractional error introduced by ignoring the shunting effect of the voltmeter is $\sim R/R_V$. When $R_V = \infty$, equation [6-20] reduces to [6-15], describing the ideal case.

6.9 Wheatstone Bridge

The Wheatstone bridge, devised by Sir Charles Wheatstone (1802–1875), English physicist and inventor, provides a convenient method for comparing an unknown resistance with one whose resistance is known. It is the most commonly used method for resistance determination.

One of the principal merits of the Wheatstone bridge is that it constitutes a *null* method. In such a method, the circuit parameters are adjusted to secure a *zero* reading on the indicating meter. In this way, any lack of linearity that the meter may possess is not a cause for concern. In addition, the sensitivity of null (or balance) methods is generally much greater than those of deflection methods.

The Wheatstone bridge, shown in Figure 6.16, requires three known resistances, A, B, and R, the last of which is continuously variable; X is the unknown resistance. The meter G is a very sensitive current meter, known as a galvanometer (see Chapter 11).

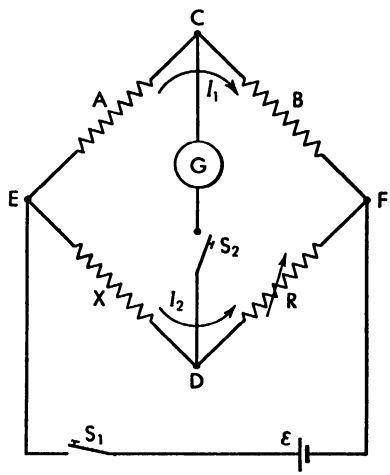


Fig. 6.16

Switch S_1 is first closed and then R is adjusted to bring points C and D to the same potential, as indicated by the fact that zero current passes through the galvanometer when switch S_2 is closed. When *balance* has been achieved, the same current, I_1, flows through both A and B; likewise, X and R carry the same current, I_2, as indicated in the figure. Applying Kirchhoff's second law to the loops ECD and FCD, we obtain the two relationships:

$$I_1A - I_2X = 0, \quad \text{or} \quad I_1A = I_2X \qquad [6\text{-}21]$$

and

$$I_1B - I_2R = 0, \quad \text{or} \quad I_1B = I_2R \qquad [6\text{-}22]$$

Dividing [6-21] by [6-22] and rearranging, we obtain the familiar bridge relationship:

$$X = R\left(\frac{A}{B}\right) \qquad [6\text{-}23]$$

Commercial bridges provide a means of varying the ratio A/B, usually by factors of 10. This makes it possible to determine the values of unknown resistances that differ greatly in magnitude from the variable one.

As mentioned earlier, when the Wheatstone bridge is in use, the switch S_1 should be closed first. In this way any transient currents (Chapter 12) that may arise because of inductance in the resistance coils will die out before the galvanometer is brought into the circuit. Until approximate balance is obtained, S_2 is closed for only very brief periods in order to avoid harming the galvanometer. The galvanometer may be further protected by a low-resistance shunt which is switched out of the circuit as balance is approached.

Sensitivity of a Wheatstone Bridge

Although the battery emf and the constants of the galvanometer do not appear in the balance equation [6-23], they partially determine the sensitivity of the bridge inasmuch as the galvanometer deflections, when the bridge is slightly out of balance, are dependent upon them. The balance point is found by adjusting R to a too-high value, as evidenced by a perceptible galvanometer deflection, then to a correspondingly too-low value, as evidenced by an equal but opposite deflection, and finally choosing R to be the average of these two slightly-off-balance values. The magnitude of these slightly-off-balance galvanometer deflections is crucial to the sensitivity of the bridge; the larger they are, the better. Since any current flowing through the galvanometer is always proportional to the battery emf, we may conclude that the sensitivity of a bridge is directly proportional to the battery emf. (One is limited, in this regard, however, by the important practical consideration that excessive or prolonged currents through the bridge will raise the temperature of the resistance elements and alter their resistance values.) In addition, from a consideration of the conditions for maximum power transfer (§ 5.9), one may say that, if other things are equal, the bridge sensitivity is greatest when the galvanometer resistance is equal to the resistance that would be measured between C and D if the galvanometer were removed. Furthermore, it hardly needs saying that the sensitivity of the bridge increases with the sensitivity of the galvanometer. If one has a choice of galvanometers, resistors, or both to be used for the measurement of a particular resistance, the relative sensitivities of the various possible arrangements can be calculated from an expression (equation [6-31]) to be derived below. Fortunately, the sensitivity is not a rapidly varying function of the bridge parameters.

■ *Conditions for Maximum Sensitivity of a Wheatstone Bridge*

In order to analyze the problem of bridge sensitivity, let us derive the expression for the galvanometer current when the bridge is slightly out of balance. The symbols to be used in this analysis are shown in Figure 6.17. The battery is assumed to have zero internal resistance.

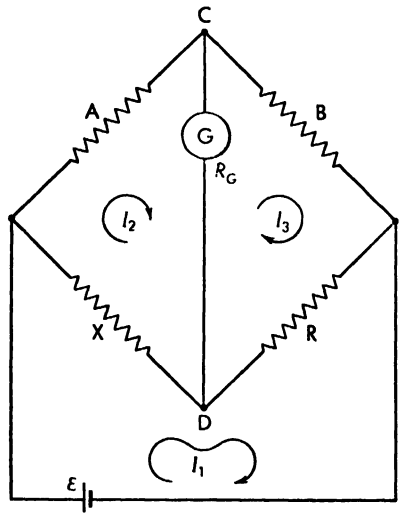


Fig. 6.17

Applying Kirchhoff's second law to this network, we obtain the three simultaneous equations:

$$\begin{aligned} \mathcal{E} &= (I_1 - I_2)X + (I_1 - I_3)R \\ 0 &= I_2A + (I_2 - I_3)R_G + (I_2 - I_1)X \\ 0 &= I_3B + (I_3 - I_1)R + (I_3 - I_2)R_G \end{aligned} \qquad [6\text{-}24]$$

which, when rewritten, become

$$\begin{aligned} I_1(X + R) - I_2X - I_3R &= \mathcal{E} \\ I_1X - I_2(A + X + R_G) + I_3R_G &= 0 \\ I_1R + I_2R_G - I_3(B + R + R_G) &= 0 \end{aligned} \qquad [6\text{-}25]$$

and

The solutions of these equations are

$$\begin{aligned} I_1 &= \frac{\mathcal{E}[(A + X + R_G)(B + R + R_G) - R_G^2]}{\Delta + R_G(A + B)(X + R)} \\ I_2 &= \frac{\mathcal{E}[RR_G + X(B + R + R_G)]}{\Delta + R_G(A + B)(X + R)} \\ I_3 &= \frac{\mathcal{E}[XR_G + R(A + X + R_G)]}{\Delta + R_G(A + B)(X + R)} \end{aligned} \qquad [6\text{-}26]$$

and

where $\Delta = ABX + BXR + XRA + RAB$

From equations [6-26] the galvanometer current is calculated to be

$$I_G = I_3 - I_2 = \frac{\mathcal{E}(AR - BX)}{\Delta + R_G(A + B)(X + R)} \qquad [6\text{-}27]$$

From this we see that when $AR = BX$ the galvanometer current is zero. This is the balance equation (equation [6-23]), which we have already derived above in a more direct manner.

Equation [6-27] enables us to investigate the conditions under which the bridge is most sensitive. First, however, a criterion must be agreed upon by which the sensitivity, S, can be reckoned. Two factors are involved—the sensitivity of the galvanometer itself and the magnitude of the galvanometer current (ΔI_G) that results when the balancing resistor (R) undergoes a fractional change ($\Delta R/R$) from its balanced value. It will be shown in § 11.2 that for galvanometers of similar type and similar coil dimensions the current sensitivity is proportional to the square root of the galvanometer resistance. Thus, we may write

$$\text{Bridge sensitivity, } S = k\sqrt{R_G}\, R \left(\frac{\partial I_G}{\partial R}\right)_{I_G=0} \qquad [6\text{-}28]$$

where k is the factor of proportionality relating the current sensitivity of the galvanometer to $\sqrt{R_G}$.

Using equation [6-27] to evaluate equation [6-28], we obtain

$$S = \frac{k\sqrt{R_G}\, RA\mathcal{E}}{\Delta + R_G(A + B)(X + R)} \qquad [6\text{-}29]$$

where use has been made of the fact that the bridge is essentially balanced, so that $AR \doteq BX$. Evidently, the sensitivity can be increased by increasing $\mathcal{E}$, that is, by increasing the currents in the bridge. As mentioned earlier, however, this would raise the temperatures of the resistors involved and would alter their resistance values. Let us, therefore, limit the current through the unknown resistance, X, to some permissible value, I_1'. Then, at balance [6]

$$\mathcal{E} = I_1'(R + X) \qquad [6\text{-}30]$$

When this expression is substituted in equation [6-29], and when B is everywhere replaced by AR/X, we obtain

$$S = \frac{kXI_1'}{\left[\dfrac{(X + A)}{\sqrt{R_G}} + \sqrt{R_G}\left(1 + \dfrac{X}{R}\right)\right]} \qquad [6\text{-}31]$$

The sensitivity can be maximized by proper choice of R_G, as follows,

$$\frac{\partial S}{\partial R_G} = 0 = \frac{-kXI_1'\left[-\frac{1}{2}R_G^{-\frac{3}{2}}(X + A) + \frac{1}{2}R_G^{-\frac{1}{2}}\left(1 + \dfrac{X}{R}\right)\right]}{\left[\dfrac{X + A}{\sqrt{R_G}} + \sqrt{R_G}\left(1 + \dfrac{X}{R}\right)\right]^2} \qquad [6\text{-}32]$$

[6] What follows is taken principally from P. Vigoureux and C. E. Webb, *Principles of Electric and Magnetic Measurements*, Prentice-Hall, 1936.

whence,

$$R_G = \frac{(X + A)R}{(X + R)} \qquad [6\text{-}33]$$

This, substituted in [6-31], gives the maximum sensitivity that the proper choice of R_G will provide:

$$S_{max} = \frac{kXI_1'}{2\sqrt{(X + A)\left(1 + \frac{X}{R}\right)}} \qquad [6\text{-}34]$$

We shall discuss equation [6-33] before drawing conclusions from equation [6-34]. It can easily be shown (Problem 6.11) that the quantity $R(X + A)/(X + R)$ appearing in equation [6-33] is the resistance measured between C and D with the galvanometer removed. Equation [6-33] thus documents the statement made earlier that maximum sensitivity coincides with a resistance match between the galvanometer and the rest of the circuit. We also see from equation [6-33] that, when the four arms of the bridge are equal, $R_G = X$.

Turning now to equation [6-34], we see that when the four arms of the bridge are equal, $S_{max} = k\sqrt{X}\, I_1'/4$. S_{max} can be increased by employing unequal arm arrangements, but the increases are small, even for large changes. For example, when R is increased by a factor of ten, and B is correspondingly increased to maintain balance, S_{max} increases by only one-third. If R_G be ten times its optimum value, the reduction in sensitivity for an equal arm bridge is only 40 percent. In short, the sensitivity is not a rapidly changing function of either R_G or the relative sizes of the arm resistances.

Since bridge measurements are made, in practice, with equipment that happens to be available, the conditions for maximum sensitivity can rarely be achieved. Notwithstanding, equations [6-31], [6-33], and [6-34] are useful guides to the best deployment of the available apparatus.

In conclusion, attention is drawn to an additional useful guiding principle. It can be demonstrated from equation [6-34] (Problem 6.12) that, when two arms of a bridge contain low resistances and the other two arms contain high resistances, the galvanometer should be placed between the junction of the two small resistances and the junction of the two large resistances, rather than between the two other corners of the bridge network.

The Slide-Wire Bridge

In a form of Wheatstone bridge much used in elementary laboratories two of the known resistances, A and B, together comprise a uniform resistance wire of constant length (Fig. 6.18).

The resistance wire, usually 1 m in length, extends between E and F and is divided into its two portions by the sliding contact C. Hence, the ratio A/B is continuously variable and is given by the ratio of the lengths EC/CF of the uniform resistance wire. It can readily be shown (Problem 6.13) that this bridge is most accurate when $EC = CF$.

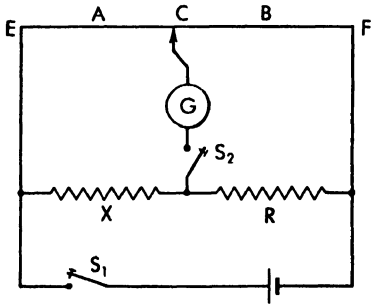


Fig. 6.18

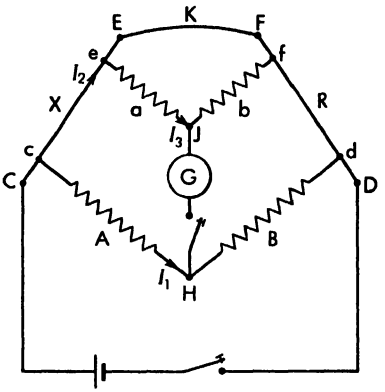


Fig. 6.19

Very High Resistances

When accurately known resistances of sufficient magnitude and galvanometers of sufficient sensitivity are not available for the Wheatstone-bridge determination of a very high resistance (say, $X > 10^8$ ohms), the rate at which a charged capacitor discharges through the resistance in question can serve as the basis for its determination (§ 12.2).

6.10 The Kelvin Double Bridge

When low resistances are determined by means of the Wheatstone bridge, the resistance of connecting wires and imperfect contacts at junction points may lead to serious error. In addition, when low resistance elements comprise the Wheatstone bridge, the potential drops involved are so small that detectable galvanometer deflections only occur when the bridge is far off balance. These difficulties are avoided in the Kelvin bridge (Fig. 6.19), a modification of the Wheatstone bridge that is especially adapted for the accurate comparison of low resistances (say, $X < 1$ ohm). The bridge was invented by Lord Kelvin.

X is the unknown low resistance and R is a continuously variable, standard, low resistance. A, B, a, and b have resistances much larger than X or R. The ratios a/b and A/B are usually variable in decade steps.

The contact resistance problem is dealt with in the following manner: X is a portion, of length ce, of a wire or rod. Electrical connections are made to the ends of this rod at C and E, respectively. Any contact resistance associated with these junctions does not comprise a portion of X. Additional electrical connections are made to the rod at c and e, respectively. Contact resistances associated with these junctions comprise parts of a and A, respectively, and are

negligible by comparison with the total resistances involved. Similar arguments apply to the contact resistances at F, D, f, and d.

Most of the current flowing through X reaches R via the low-resistance route *eEKFf*. The fact that the resistance of this route is not known does not concern us because it does not affect the bridge balance. This we shall now see in the course of deriving the working equation for the bridge. At balance, applying Kirchhoff's second law in turn to the loops *ceJH* and *fdHJ*, we have

$$AI_1 = XI_2 + aI_3 \qquad [6\text{-}35]$$

and

$$BI_1 = RI_2 + bI_3 \qquad [6\text{-}36]$$

Dividing [6-35] and [6-36], we obtain

$$\frac{A}{B} = \frac{X + ar}{R + br} \qquad [6\text{-}37]$$

where, for convenience, r is written for the ratio I_3/I_2. This ratio r will be small. Now solving [6-37] for X, we find:

$$X = \frac{A}{B}(R + br) - ar \qquad [6\text{-}38]$$

or

$$X = \frac{A}{B}R + rb\left(\frac{A}{B} - \frac{a}{b}\right) \qquad [6\text{-}39]$$

The final term, already small, is made to vanish by making $\frac{A}{B} = \frac{a}{b}$. The unknown resistance is therefore computed from the usual bridge equation,

$$X = \frac{A}{B}R \qquad [6\text{-}40]$$

Commercial models of the Kelvin bridge provide a range of fixed values for the ratio A/B.

6.11 Platinum Resistance Thermometer

The suggestion to employ in thermometry the variation in resistance of a metal with temperature was made in 1871 by Sir William Siemens, German-born British electrical engineer. Platinum is the most suitable metal for this purpose. In addition to its principal attribute, that it does not combine with oxygen, even at elevated temperatures, its resistance versus temperature curve is closely parabolic over a wide temperature range. It can therefore be described by a second-order equation of the type of [6-9].

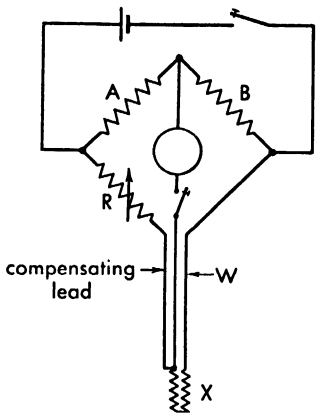


Fig. 6.20

A sample of pure, well-annealed, platinum wire is made the fourth arm of a Wheatstone bridge, as shown in Figure 6.20. When this element is raised above room temperatures, the resistance of the lead-wire, W, connecting it to the bridge also increases. This change is canceled out by including in the variable arm of the bridge a similar lead-wire (which experiences an identical temperature change) and by making $A = B$.

Before it can be used to determine an unknown temperature, the resistance thermometer must be calibrated. This is done by determining its resistance at three standard temperatures, in order to evaluate the three constants in the resistance-temperature relation (equation [6-9]). Suitable standard temperatures are, for example, the ice point (0° C), the steam point (100° C), and the freezing point of cadmium (320.9° C). R_0, α, and β are obtained from the three simultaneous equations,

$$R_0 = R_0$$

$$R_{100} = R_0[1 + \alpha(100) + \beta(100)^2] \qquad [6\text{-}41]$$

and

$$R_{320.9} = R_0[1 + \alpha(320.9) + \beta(320.9)^2]$$

An unknown temperature lying within the calibration range may now be readily determined from a knowledge of the resistance displayed by the platinum element when exposed to it. When calibrated between the melting point of mercury (−38.87° C) and the boiling point of sulphur (444.60° C), a platinum resistance thermometer yields a temperature at any point within this range that is correct to within 0.02–0.03° C. More than usual care must be taken to keep the bridge current low and so avoid Joule (I^2R) heating of the elements.

The International Temperature Scale of 1949 is defined in the temperature range $-182.97° \text{ C} < t < 630.5° \text{ C}$ in terms of a platinum resistance thermometer.

6.12 Resistance Standards

Almost without exception, resistance standards are made of manganin, in the forms of sheet material for low resistances or of wire for higher values. Manganin has two outstanding qualifications: its temperature coefficient of resistance is only a few parts per million at room temperatures and its thermal emf against copper, the usual lead-wire, is only 2–3 μV per degree Centigrade.

On the other hand, manganin is strain-sensitive and it oxidizes to some extent at ordinary temperatures.

The *primary* standard of resistance in the United States is a group of ten 1-ohm resistors. These are wound of No. 12 manganin wire, annealed *in vacuo* at 550° C, and are sealed in air in vacuum-tight containers. The average change per resistor since the group was established in 1933 has been 12 microhms. The values of other resistors can be ascertained with respect to these to a few parts in 10^7.

Laboratory standards of resistance are usually the four-terminal type shown in Figure 6.21. Two of the terminals, V' and V'', are for connecting a volt-

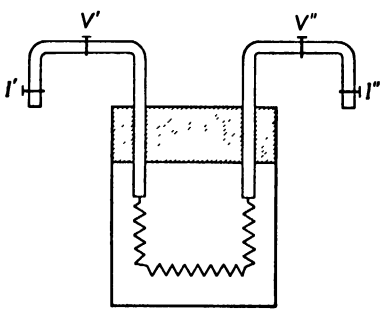


Fig. 6.21

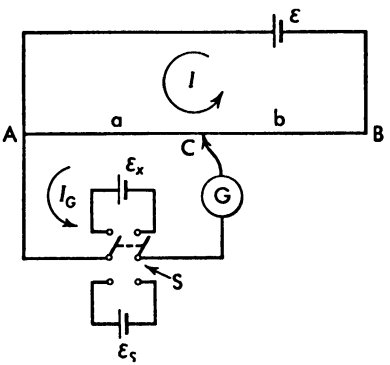


Fig. 6.22

meter; the other two, I' and I'', are for making the current connections. The resistance between the potential connections is accurately given by the stated value, usually to ≤ 0.02 percent.

The "standard" resistors most frequently encountered in elementary laboratories are the *decade-box* or *plug box* varieties. Here, variations in switch and contact resistances lead to an uncertainty of ~0.1 percent. In modern units of this type the individual resistance coils are wound in such a way as to possess zero self-inductance—they are said to be *noninductively* wound.

6.13 The Potentiometer

The potentiometer is a device for comparing an unknown potential with a known one. In order to elucidate the principle of the potentiometer, let us consider the primitive slide-wire version shown in Figure 6.22. AB is a uniform wire of resistance R, to which sliding contact is made at C. $\mathcal{E}$ is the emf of the battery supplying current to the slide-wire, $\mathcal{E}_s$ is a known emf, and $\mathcal{E}_x$ is an

emf whose value is to be determined. S is a double-pole double-throw switch, by means of which either $\mathcal{E}_s$ or $\mathcal{E}_x$ may be thrown into the branch of the circuit that contains the galvanometer G.

Let us suppose that S is thrown to the lower position. Applying Kirchhoff's law to the loop in which the galvanometer current I_G is flowing, we have

$$\mathcal{E}_s = (I - I_G)a \qquad [6\text{-}42]$$

By adjusting the contact point C, a position is found (provided $\mathcal{E} > \mathcal{E}_s$) for which $I_G = 0$. At this point, for which $a = a_s$, from [6-42],

$$\mathcal{E}_s = Ia_s \qquad [6\text{-}43]$$

Likewise, when S is thrown to the upper position, a balance point is found for $a = a_x$. Here

$$\mathcal{E}_x = Ia_x \qquad [6\text{-}44]$$

Dividing [6-42] by [6-44] and rearranging, we may express the unknown emf in terms of the known emf and the two lengths of resistance wire for which balances were obtained, viz.,

$$\mathcal{E}_x = \frac{a_x}{a_s}\mathcal{E}_s \text{ volts} \qquad [6\text{-}45]$$

It will be realized that the polarities of the emf's $\mathcal{E}$ and $\mathcal{E}_s$ (or $\mathcal{E}_x$) must be such as to send currents in opposite directions in the branch containing the galvanometer. At balance these opposite currents are equal in magnitude.

6.14 Standard Cells

By means of the potentiometer, unknown potentials and emf's can be compared with a standard potential, which is invariably the emf of an electrolytic cell. The preparation and care of *standard cells* has been the subject of much study.

The international standard of potential is the *Weston cadmium cell.* The positive electrode of this cell is mercury and the negative electrode is a mercury-cadmium amalgam containing 12.5 percent cadmium. The electrolyte is a saturated solution of cadmium sulphate. The emf of a saturated cadmium cell is given by the formula (internationally adopted in 1908):

$$\mathcal{E}_t = \mathcal{E}_{20} - 4.0 \times 10^{-5}(t - 20) - 9.5 \times 10^{-7}(t - 20)^2 + 10^{-8}(t - 20)^3 \qquad [6\text{-}46]$$

where $\mathcal{E}_t$ and $\mathcal{E}_{20}$ are the emf's at $t°$ C and 20° C, respectively.

The National Bureau of Standards in Washington maintains a group of 47 such cells, twelve of which have been in use since 1906. They are kept at

a constant temperature of 28.00 ± 0.01° C and intercompared periodically. An intercomparison in 1955 revealed that the United States and French standards differed by only 0.7 microvolts.

The usual laboratory standard in North America is an *unsaturated* cadmium cell. These cells can be shipped without damage and their temperature coefficient of emf (~ -5 microvolts per degree Centigrade) is lower than that of the saturated type. The emf of such cells, however, decreases with time at a rate of 50–100 μV per year. Also, the emf of each cell must be ascertained individually by the manufacturer and supplied with the cell.

Care of Standard Cells

Among the injunctions given by F. K. Harris[6] for the care of standard cells are the following.

a. Sudden or large changes in temperature should be avoided.
b. When the cell is in use, all parts of it should be at the same temperature.
c. Temperatures above 40° C and below 4° C should be avoided.
d. Currents drawn from a cell should never exceed 100 microamperes and should flow for only a few seconds at a time. In most standard cells intended for student use, a 10,000-ohm resistor is built into the unit, thus guaranteeing that excessive current will never be drawn.
e. An unsaturated cell should be discarded whenever its emf falls below 1.0183 volts. It usually lies in the range $1.0185 < \mathcal{E} < 1.0197$ volts.

6.15 Direct-Reading Potentiometers

The potentiometer circuit shown in Figure 6.22 and discussed in § 6.13 can be embellished slightly to provide a direct-reading instrument, a typical example of which is shown in Figure 6.23.

In this variety of direct-reading potentiometer, the uniform resistance wire of Figure 6.22 is replaced by two resistances, A and B, in series. Resistance A consists of fifteen 10-ohm resistors, while B consists of a uniform wire of total resistance 10 ohms. Current is supplied to A and B from the battery of emf $\mathcal{E}$, the variable external resistance R being a series member of this circuit. With S_2 in the left position, R is adjusted until the current through $A + B$ is 0.01 amperes, at which time the voltage drop across them is 0.01 volt/ohm. This process of adjustment constitutes *standardizing* the potentiometer, and is accomplished by the following method.

[6] *American Institute of Physics Handbook*, McGraw-Hill, 1957.

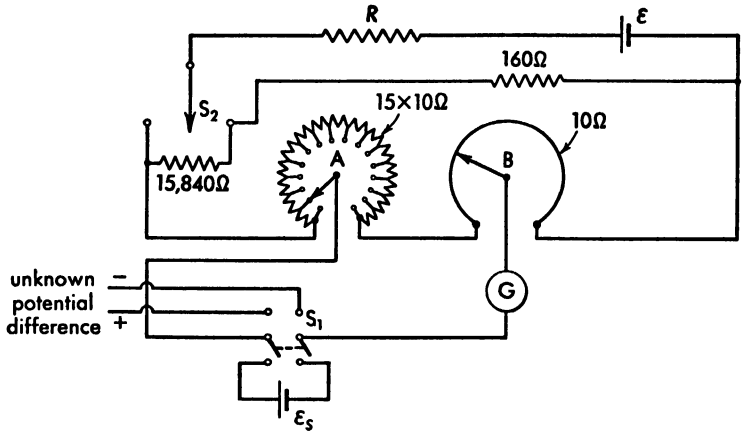


Fig. 6.23

The variable contacts on A and B are set to give, in the galvanometer circuit, a resistance that is numerically one hundred times the certified standard cell emf, i.e., approximately 101.8 ohms. Switch S_1 is then thrown to the standard-cell, or standardizing, position, and the external resistance R is adjusted to secure a balance (that is, to make the galvanometer current zero). The potential difference across the portion of $A + B$ that is in the galvanometer circuit is then equal to the open-circuit standard-cell voltage and is, in fact, 0.01 volt/ohm. Elsewhere in A and B, since the current is constant throughout, the same volt/ohm drop prevails. The potentiometer is now standardized and R is not again adjusted unless it is necessary to check and restore the standardization.

An unknown potential difference lying in the range 0–1.6 volts can now be quickly determined. It is thrown into the galvanometer circuit by means of switch S_1, after which the variable contacts on A and B are adjusted to produce a balance. Evidently, the unknown potential difference is 0.01 times the value of this balancing resistance. A and B are equipped with dials which give the value of this potential difference directly.

If desired, the current through $A + B$ may be reduced by two orders of magnitude[7] by throwing switch S_2 to the right position. The total resistance of the $A + B$ branch is now 16,000 ohms and the current is consequently $\frac{1}{100}$ of its former value. In this way, the potentiometer can be used to measure with accuracy very small emf's, such as those associated with the thermoelectric effect. The student will notice that, because of the presence of a 160-ohm resistor in the branch parallel to that containing $A + B$, a change in the position of switch S_2 does not alter the load on the battery $\mathcal{E}$.

[7] The term "order of magnitude" is synonymous with "factor of 10" and is quantitatively exact. It is sometimes incorrectly used as meaning "approximately a factor of 10."

The preceding discussion is essentially a description of the "Student's Potentiometer" manufactured by the Leeds and Northrup Company. Several more-elaborate direct-reading instruments are available but are best studied at the time of use through their instruction manuals.

6.16 Some Applications of Potentiometers

The potentiometer, by virtue of its capacity to measure potential differences, can be used to perform a number of related services. Some of these are briefly described below.

Calibration of a Voltmeter

In the circuit of Figure 6.24, the voltmeter reading can be set to any

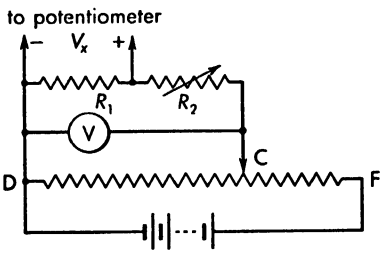


Fig. 6.24

desired value, V, by moving the contact, C. R_1 and R_2 are accurate resistors, the latter being variable. The voltage drop across R_1 is

$$V_x = \frac{R_1}{R_1 + R_2} V \text{ volts} \qquad [6\text{-}47]$$

and can be altered, by varying R_2, to a value that is within the range of an available potentiometer. The voltmeter is calibrated by adjusting C to obtain representative meter readings, and determining the actual voltages thereof by measuring V_x and using [6-47].

Comparison of Resistances

An unknown resistance can be compared with a known one by sending a common current through the two and comparing, with a potentiometer, the voltage drops across them. These voltage drops are proportional to the two resistances.

Calibration of an Ammeter

The ammeter to be calibrated is connected in series with a standard resistor, R, a variable resistor, and a battery, as shown in Figure 6.25. By means

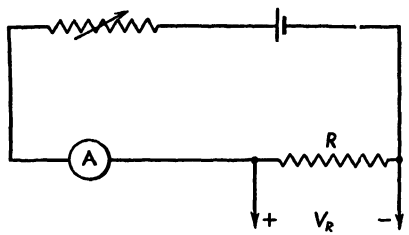


Fig. 6.25

of the variable resistor, the current through the circuit is adjusted to coincide with desired ammeter readings. The actual currents corresponding to these readings are calculated from $I = V_R/R$, where V_R is determined with a potentiometer.

6.17 Relation Between Electrical and Thermal Conductivity: The Wiedemann-Franz Law

Heat flow is described in a manner strikingly similar to that describing the flow of electricity. The thermal analogue to electric intensity is the *temperature gradient*, dt/dl. Also, the amount of heat, H, passing through a surface of area A per second corresponds to the electric current. For the electrical case we have already seen that, for an element of a uniform conductor,

$$I = \sigma A \frac{dV}{dl} \qquad \text{[5-15]} \quad \text{[6-2]}$$

"Ohm's law" for heat flowing in a uniform conductor is, therefore,

$$H = kA \frac{dt}{dl} \qquad \text{[6-48]}$$

where k, the *thermal conductivity*, is the amount of heat that passes from one side to the other of a unit cube of material when the temperature difference between two sides of the cube is 1 degree.

It has been known for a long time that good conductors of electricity are at the same time good conductors of heat. Indeed, it was discovered during the latter part of the nineteenth century that the ratio of thermal to electrical

conductivity is approximately constant for many metals. A few values of this ratio (in arbitrary units) at 0° C, are shown below.

	(k/σ)
Copper	6.71
Gold	7.09
Lead	7.15
Silver	6.86
Tin	7.35

The fact that this ratio is constant is known as the *Wiedemann-Franz law.* It strongly suggests that conduction electrons are responsible for heat as well as for electrical flow.

One can treat the transfer of heat energy by electrons by essentially the same method that was employed in § 5.3 to discuss the motion of conduction electrons under the action of an applied field. When this is done, an expression is got for k in which n, the number of conduction electrons per unit volume, and λ, the mean free path of a conduction electron, appear in the numerator as they do in the expression for $\sigma = 1/\rho$ got from equation [5-13]. Thus, in the ratio k/σ, the quantities n and λ, which are characteristic of the particular conductor, disappear and the ratio is independent of the nature of the material, in agreement with the Wiedemann-Franz observation.

6.18 Problems

E (6.1) Arsenic atoms are sometimes added to germanium crystals in order to produce a semiconductor. Germanium, in the fourth group of the periodic table, has a chemical valence of 4 and crystallizes in a cubic structure with each atom forming covalent bonds with its four nearest neighbours. Arsenic has a chemical valence of 5. When an atom of arsenic is substituted for one of the germanium atoms, will the resulting semiconductor be n- or p-type?

E (6.2) A wire having a resistance of 5 ohms is drawn out until its length is four times its former length. Assuming that the resistivity remains constant, calculate the resistance of the longer wire.

H (6.3) The periphery of a copper disk 50 cm in radius and 10^{-3} mm in thickness is maintained at a potential of 50 volts. A thin rod 1 cm in radius is soldered to the disk at its center (at right angles to the plane of the disk) and maintained at a potential of 49 volts. If the resistivity of copper is 1.7×10^{-8} ohm-m, compute the current through the disk. [Ans.: 94 amperes.]

H (6.4) In Figure 6.26 the shaded sections are nichrome sheets of thickness t and the unshaded material is copper of the same thickness. Calculate the resistance of the copper section, that is, of the portion between the boundaries A and B. $\left[\text{Ans.: } \frac{\pi\rho}{2t \ln (b/a)}.\right]$

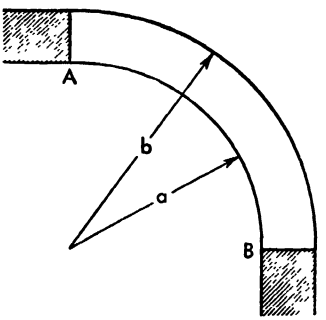


Fig. 6.26

E (6.5) Show that α_0 and α_{20} are related as follows:

$$\alpha_{20} = \frac{\alpha_0}{1 + 20\alpha_0}$$

E (6.6) Assume that the resistance-temperature curve for tungsten is linear with $\alpha_{20} = 0.006$ per degree Centigrade. The current drawn by a tungsten lamp when it is first turned on is twelve times the final operating current. Calculate the operating temperature of the tungsten filament in the lamp. [Ans.: ~1850° C.]

M (6.7) A *thermistor* is a semiconductor whose resistivity varies *greatly* with temperature ($\alpha_0 \approx -0.04$). The voltage across a thermistor, as the current is increased, follows the course shown in Figure 6.27. Account qualitatively for the shape of the I-versus-V characteristic.

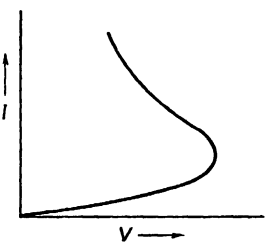


Fig. 6.27

E (6.8) A voltmeter is connected across the terminals of a resistor that is connected in series with an ammeter. If the ammeter reads 0.5 amperes and the voltmeter (resistance 10,000 ohms) reads 100 volts, what is the percentage error introduced in calculating the value of the resistor by assuming the voltmeter and ammeter are ideal instruments? [Ans.: 2 percent.]

E (6.9) Suppose the voltmeter in the previous problem were connected across the ammeter-resistor combination instead of across the resistor alone. For the same ammeter and voltmeter readings, what value of ammeter resistance leads to the same percentage error as in Problem 6.8? [Ans.: 4 ohms.]

E (6.10) In Problems 6.8 and 6.9, suppose that the ammeter is an exceptionally good one (in the sense that its resistance is unusually low), whereas the voltmeter has a poorer-than-average ohms/volt rating. If, under these circumstances, no account is to be taken of the non-ideal characteristics of the two instruments, which method of connection is to be preferred?

E (6.11) (Based on small-type material in § 6.9.) A Wheatstone bridge (Fig. 6.17) can be considered as an energy source, with terminals C and D, that is delivering power to a load consisting of the galvanometer (re-

sistance, R_G). The bridge has maximum sensitivity when the load is matched to the energy source. Assuming the bridge to be balanced and the battery to have zero internal resistance, show that maximum bridge sensitivity occurs when

$$R_G = \frac{(X + A)}{(X + R)} R \qquad [6\text{-}33]$$

M (6.12) (Based on small-type material in § 6.9.) Show from equation [6-34] that, for maximum sensitivity in a nonequal-arm bridge, the galvanometer should be connected between the junction of the two high-resistance arms and the junction of the two low-resistance arms.

M (6.13) Assume that in a slide-wire bridge (Fig. 6.18) the error in resistance determination arises from the uncertainty in measuring the lengths A and B. Obviously, the most accurate arrangement of such a bridge is that for which a given change in the length of A (or B) represents the least change in X. Show that this condition holds when $A = B$.

E (6.14) The ratio A/B (Fig. 6.16) can usually be varied in commercial Wheatstone bridges by factors of ten between the values 10^3 and 10^{-3}. If R can be changed in 1-ohm steps between 0 and 10^4 ohms, to what precision may the resistance of a nominally 10-ohm resistor be determined? Assume that the accuracy is not limited by uncertainties in the values A/B or R.

E (6.15) What is the voltage between A and B in Figure 6.28: (a) as determined by a potentiometer; (b) as determined by a 0–2 volt voltmeter, with a rating of 1000 ohms/volt?

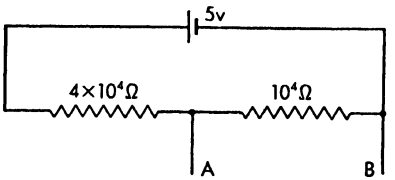


Fig. 6.28

E (6.16) A DC voltmeter has three ranges, 0–1, 0–10, and 0–100 volts. Its resistance on the highest scale is 10^4 ohms. When connected across a battery it reads, on the 0–10 scale, a voltage of 5.90 volts. The same battery has an emf of 6.00 volts as determined by a potentiometer. What is the internal resistance of the battery? [Ans.: 17 ohms.]

E (6.17) The value of a resistance may be determined by the so-called "voltmeter method," shown in Figure 6.29. With switch S open the voltmeter (resistance R_V ohms) reads V_1 volts, with the switch closed it reads V_2 volts. Show that the value of the unknown resistance is given by $R = \dfrac{(V_2 - V_1)R_V}{V_1}$ ohms.

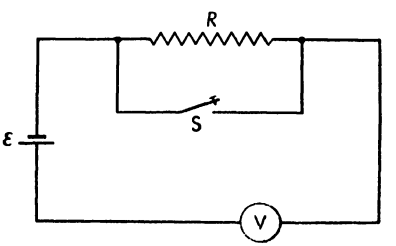


Fig. 6.29

Chapter 7

Thermoelectricity

The adjective "thermoelectric" describes phenomena in which heat energy is transformed into electrical energy, or vice-versa. If the interchange can proceed in either direction, it is said to be *reversible;* otherwise, it is *irreversible.*

The production of heat because of the passage of an electric current, the so-called I^2R or *Joule* heat, is the most familiar of the irreversible thermoelectric phenomena. This effect proceeds in one direction only; reversal of the current has no effect upon it, and it cannot be used to transform heat energy into electrical energy.

The reversible thermoelectric phenomena are three in number. These occur at junctions of dissimilar conductors and/or throughout conductors that possess a temperature gradient. In the order of their discovery they are the *Seebeck effect*, the *Peltier effect*, and the *Thomson effect*. The term *thermoelectric effects* customarily refers only to these reversible events.

7.1 Seebeck Effect

In 1826 Thomas Johann Seebeck (1770–1831), a German physicist, published the results of experiments done four years earlier that opened up a new branch of electricity. He observed that an electric current is present in a series circuit of two dissimilar metals, provided the junctions of the two metals are at different temperatures. This effect is pictured in Figure 7.1.[1] The thermoelectric current increases as $t_2 - t_1$ increases.

[1] A pair of metals that will give rise, when connected as in Figure 7.1, to a thermoelectric current is known as a *thermocouple.*

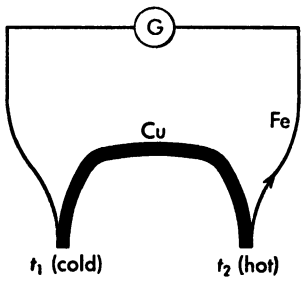


Fig. 7.1

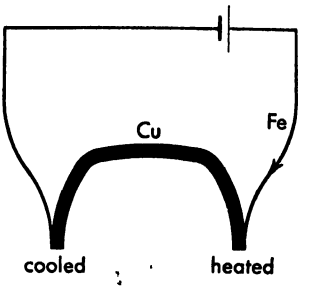


Fig. 7.2

Seebeck investigated the thermoelectric properties of a large number of metals and arranged them in a *thermoelectric series.* When a circuit is formed of two metals in the series, the thermal emf is greater the further the metals are apart in the series. Also, the direction of the current, at the hot junction, is from the metal occurring earlier in the series to the metal occurring later in it. The following is a selection from Seebeck's series: Bi, Ni, Co, Pd, Pt, Cu, Mn, Hg, Pb, Sn, Au, Ag, Zn, Cd, Fe, Sb, Te. The position of a metal in this series, as Seebeck himself observed, is dependent upon the temperature and also somewhat upon its purity. The thermal emf for the couple Cu and Fe shown in Figure 7.1, although not atypical, is only 1.22 millivolts for $t_2 - t_1 = 100°$ C.

We shall later see that the Seebeck effect is not independent of, but is rather a combination of, the Peltier and Thomson effects.

7.2 Peltier Effect

In 1834 Jean Peltier (1785–1845), a French watchmaker who had retired at the age of thirty to devote himself to science, discovered that junctions of dissimilar metals were heated or cooled, depending upon the direction in which a current passed through them. This reversible effect was found to be proportional to the first power of the current, not to its square, as is the irreversible generation of heat caused by resistance throughout the circuit.

As an example of the Peltier effect, let us replace the galvanometer in Figure 7.1 with a battery which gives rise to a current whose direction is opposite to the thermoelectric current in Figure 7.1. The arrangement is shown in Figure 7.2. Under these conditions it is observed, as indicated in the figure, that the right-hand junction is heated, showing that electrical energy is being transformed into heat energy. Meanwhile, heat energy is transformed into electrical energy at the left junction, thereby causing it to be cooled. When the current is reversed, heat is absorbed at the right junction and produced at the left one.

It should be noted that the heating and cooling of the two junctions in

Figure 7.2, caused by the Peltier effect, provide the conditions necessary for the Seebeck effect of Figure 7.1. The direction of the resulting Seebeck current is opposite to the battery current. If this thermoelectric current did not oppose, but aided, the battery current, it would result in the heated (cooled) junction being still further heated (cooled), thereby resulting in a still greater current, and so on, without limit.

Absorption or production of heat at a current-carrying junction is an indication that an emf exists at the junction. Suppose the direction of the main current is the same as that produced by the *junction emf*, π, as in Figure 7.3. Each current carrier, q, is accelerated as it crosses the junction, gaining energy:

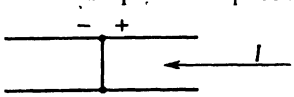


Fig. 7.3

$$W = \pi q \qquad [7\text{-}1]$$

This energy must be absorbed from the environment, in which event the junction is cooled. When Figure 7.2 is interpreted in the light of this argument, it is seen that at each junction the copper is positive with respect to the iron.

The heat absorbed per second at a junction carrying a current of I amperes is given by

$$\text{Heat absorbed/sec} = \pi_{ab} I \qquad [7\text{-}2]$$

where the current is from metal a to metal b. The junction emf, π_{ab}, is known as the *Peltier Coefficient;* it is the heat energy absorbed per second at the junction per unit current, that is, the heat energy absorbed per unit charge. π_{ab} is positive if metal a is positive with respect to metal b (thus $\pi_{\text{Cu,Fe}}$ is positive). The magnitude of π_{ab} is a function of the temperature of the junction. For identical temperatures $\pi_{ab} = -\pi_{ba}$. Thus, if the direction of the current in equation [7-2] is reversed, the heat absorbed per second is

$$\text{Heat absorbed/sec} = \pi_{ba} I \qquad [7\text{-}3]$$

which should be, and is, opposite in sign to equation [7-2].

7.3 Thomson Effect

In 1851 William Thomson (later Lord Kelvin) was led by thermodynamic reasoning, which we shall not reproduce here, to conclude that sources of emf exist in a thermoelectric circuit in addition to those located at the junctions. In particular, he predicted that an emf would arise within a *single* conductor whenever a temperature gradient was present. The truth of this prediction can be demonstrated by the experiment illustrated in Figure 7.4.

In this experiment a current passes through an iron rod which is bent into a U-shape. Resistance coils, R_1 and R_2, are wound about the two sides of the U, as shown. These form two arms of a balanced Wheatstone bridge. The bottom of the U is then heated. This establishes two temperature gradients—a positive one extending from A to C, and a negative one extending from C to B. As a result of this operation, the bridge becomes unbalanced in such a direction as to indicate that the resistance of R_1 has increased more than that of R_2. Evidently, heat has been liberated at R_1 and absorbed at R_2.

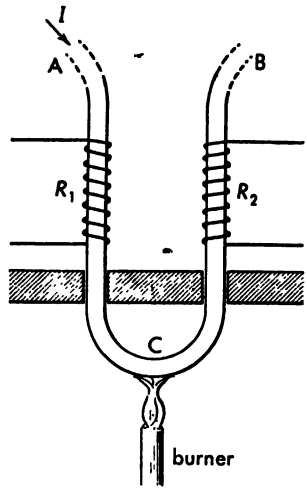


Fig. 7.4

As in the preceding section, absorption of heat is to be taken as evidence for an emf that is acting in the same direction as that of the current, that is to say, electrical energy is being supplied to the circuit at the expense of heat energy of the environment. Such is the case in the section AB. Likewise, in the section AC, the current is opposed by an emf, with consequent transformation of electrical energy into heat energy. Thus, in iron, the Thomson emf would give rise to a current *in the iron* from hot to cold regions. Many metals, including bismuth, cobalt, nickel, and platinum, in addition to iron, exhibit this same property, which is referred to as the *negative Thomson effect.* Another group of metals, including antimony, cadmium, copper, and silver, display a *positive Thomson effect;* in these, the direction of the Thomson emf is such as to support a current within the metal from cold to hot regions. In one metal, lead, the Thomson effect is zero. In certain metals the effect reverses sign as the temperature is raised or as the crystal structure is altered.

The magnitude of the Thomson emf for a given material, a, is expressed in terms of the *Thomson coefficient,* σ_a, which has dimensions of emf/degree. Thus $\sigma_a\, dt$ is the emf that exists between two points whose temperatures differ by $dt°$ C. Hence, the heat absorbed per second between two points at temperatures t_1 and t_2, respectively, when a current of I amperes passes through the material, is given by

$$\text{Heat absorbed/sec} = I\int_{t_1}^{t_2} \sigma_a\, dt - I^2R \qquad [7\text{-}4]$$

The Thomson coefficient, σ_a, is positive (negative) for materials exhibiting the positive (negative) Thomson effect. The second term in [7-4] is, of course, simply the Joule heat that is always liberated when a current flows through an imperfect conductor. It has no relation to the Thomson effect, but is included in [7-4] for completeness.

It appears that the total Thomson emf along a conductor depends only upon the temperatures of the two ends, and not in any way upon the particular manner in which the temperature gradient varies. This empirically observed fact is known, after its discoverer, as the *law of Malus* (1851).

7.4 Seebeck emf Is the Algebraic Sum of Peltier and Thomson emf's

Let us now reconsider the circuit of Figure 7.1 with a view to establishing the relation between the Seebeck, Peltier, and Thomson effects. For the time being we shall ignore the presence of the galvanometer, but we shall later justify this disregard.

The current observed by Seebeck corresponds to a total emf in the circuit which we shall call the *Seebeck emf*, $\mathcal{E}_{ab}$, where a and b refer to the two metals constituting the thermocouple. $\mathcal{E}_{ab}$ (in this case $\mathcal{E}_{Cu,Fe}$) is the resultant of four component emf's: (a) the Peltier emf, $(\pi_{Cu,Fe})_{t_2}$, at the hot junction, (b) the Peltier emf, $(\pi_{Fe,Cu})_{t_1} = -(\pi_{Cu,Fe})_{t_1}$, at the cold junction, (c) the Thomson emf, $\int_{t_2}^{t_1} \sigma_{Fe}\, dt$, of the iron, and (d) the Thomson emf, $\int_{t_1}^{t_2} \sigma_{Cu}\, dt$, of the copper. Because iron (negative) and copper (positive) have Thomson coefficients of opposite sign, the two Thomson emf's act in concert. The Peltier emf's, however, oppose one another. We can therefore write:

$$\mathcal{E}_{Cu,Fe} = (\pi_{Cu,Fe})_{t_2} - (\pi_{Cu,Fe})_{t_1} + \int_{t_1}^{t_2} \sigma_{Cu}\, dt + \int_{t_2}^{t_1} \sigma_{Fe}\, dt \qquad [7\text{-}5]$$

or

$$\mathcal{E}_{Cu,Fe} = \left| \pi_{Cu,Fe} \right|_{t_1}^{t_2} + \int_{t_1}^{t_2} (\sigma_{Cu} - \sigma_{Fe})\, dt \qquad [7\text{-}6]$$

Equation [7-6] expresses the relationship between the Seebeck emf and its constituents. Since the Seebeck emf is the resultant emf in a thermocouple circuit, we shall hereafter refer to it simply as the *thermal emf*.

Law of Intermediate Temperatures

It can be seen from equation [7-6] that the magnitude of the thermal emf for a thermocouple consisting of a given pair of metals depends only upon the temperature, t_1, of the reference junction and the temperature difference, $t_2 - t_1$, between it and the hot junction. We can rewrite equation [7-6] in the following form

$$\mathcal{E}_{\mathrm{Cu,Fe}} = \left| \pi_{\mathrm{Cu,Fe}} \right|_{t}^{t_2} + \left| \pi_{\mathrm{Cu,Fe}} \right|_{t_1}^{t} + \int_{t}^{t_2} (\sigma_{\mathrm{Cu}} - \sigma_{\mathrm{Fe}})\, dt + \int_{t_1}^{t} (\sigma_{\mathrm{Cu}} - \sigma_{\mathrm{Fe}})\, dt \qquad [7\text{-}7]$$

when t is some temperature intermediate between t_2 and t_1. The first and third terms on the right of equation [7-7] give the thermal emf that results when the thermocouple is operated between temperatures t and t_2, while the second and fourth terms give the emf when it is operated between temperatures t_1 and t. Let us rewrite equation [7-7] to express this idea:

$$\mathcal{E}_{\mathrm{Cu,Fe}} \Big|_{t_1}^{t_2} = \mathcal{E}_{\mathrm{Cu,Fe}} \Big|_{t_1}^{t} + \mathcal{E}_{\mathrm{Cu,Fe}} \Big|_{t}^{t_2} \qquad [7\text{-}8]$$

Evidently, $t_2 - t_1$ may be divided into any number of temperature intervals with the same result.

Equation [7-8] provides the basis for the *law of intermediate temperatures:* the thermal emf for a thermocouple with junction temperatures t_1 and t_2 is the sum of the emf's for any number of successive temperature ranges into which $t_2 - t_1$ may be divided. The symbolic statement of the law is as follows:

$$\mathcal{E}_{ab} \Big|_{t_1}^{t_2} = \mathcal{E}_{ab} \Big|_{t_1}^{t_t} + \mathcal{E}_{ab} \Big|_{t_s}^{t_j} + \cdots + \mathcal{E}_{ab} \Big|_{t_n}^{t_2} \qquad [7\text{-}9]$$

Variation of Thermal emf with Temperature

It is found experimentally that for pure metals the relationship between the emf and the temperature difference, $t_2 - t_1$, between hot and cold junctions is approximately a parabolic one,

$$\mathcal{E}_{ab} \doteq \alpha t + \beta t^2 \qquad [7\text{-}10]$$

where $t = t_2 - t_1$. The temperature, t_1, of the cold junction is known as the *reference temperature* and, in practice, is usually 0° C.

Figure 7.5 shows the $\mathcal{E}_{ab}$ versus t curve for a copper-iron thermocouple ($t_1 = 0°$ C). The *neutral temperature* (~270° C) is found from equation [7-10] by setting $d\mathcal{E}_{ab}/dt = 0$. For most pure metal thermocouples the neutral temperature is much higher than for copper-iron.

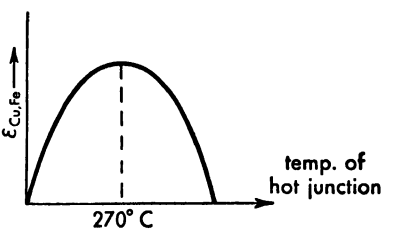


Fig. 7.5

Thermocouples made of alloys often display a practically linear relationship between $\mathcal{E}_{ab}$ and t over wide temperature ranges.

7.5 Thermoelectric Power

The rate of change of $\mathcal{E}_{ab}$ with t is defined as the *thermoelectric power*, θ. From equation [7-10] (assuming it to be exact) we see that:

$$\theta = \frac{d\mathcal{E}_{ab}}{dt} = \alpha + 2\beta t \qquad [7\text{-}11]$$

The thermoelectric power is thus seen to be a linear function of temperature, a property that makes it a particularly useful figure of merit in describing thermocouples. The inclusion of the word "power" in the name for θ is unfortunate since the thermoelectric "power" does not have units of power at all, but rather volts/degree.

7.6 Effect of Other Metals in a Thermocouple Circuit: Law of Intermediate Metals

Actual thermocouple circuits usually contain other metals in addition to those constituting the thermocouple. For example, there are the metals comprising the circuital portion of the galvanometer or potentiometer, as well as the metals associated with switches, connecting wires, etc. It is necessary to investigate what effect, if any, these additional metals have upon the magnitude of the thermal emf. We shall consider first the effect of a third metal, and then generalize the result.

Let us suppose that the third metal, say Pt, is introduced into the Cu-Fe thermocouple circuit in the manner shown in Figure 7.6. The Seebeck emf for the circuit is now

$$\mathcal{E}_{\mathrm{Cu,Fe,Pt}} = (\pi_{\mathrm{Pt,Cu}})_{t_1} + (\pi_{\mathrm{Cu,Fe}})_{t_2} + (\pi_{\mathrm{Fe,Pt}})_{t_3} + \int_{t_1}^{t_2} \sigma_{\mathrm{Cu}}\, dt + \int_{t_2}^{t_3} \sigma_{\mathrm{Fe}}\, dt + \int_{t_3}^{t_1} \sigma_{\mathrm{Pt}}\, dt \qquad [7\text{-}12]$$

Now let $t_1 = t_3$, i.e., assume that the two terminals are at the same temperature. Equation [7-12] reduces to

$$\mathcal{E}_{\mathrm{Cu,Fe,Pt}} = (\pi_{\mathrm{Pt,Cu}} + \pi_{\mathrm{Fe,Pt}})_{t_1} + (\pi_{\mathrm{Cu,Fe}})_{t_2} + \int_{t_1}^{t_2} (\sigma_{\mathrm{Cu}} - \sigma_{\mathrm{Fe}})\, dt \qquad [7\text{-}13]$$

In order to simplify this expression, let us suppose for the moment that the entire circuit is at the same temperature, t_1. Under these conditions no current flows in the circuit. Since there are no temperature gradients, the Thomson emf's are all zero. It therefore follows that:

$$(\pi_{\mathrm{Pt,Cu}})_{t_1} + (\pi_{\mathrm{Cu,Fe}})_{t_1} + (\pi_{\mathrm{Fe,Pt}})_{t_1} = 0 \qquad [7\text{-}14]$$

When this expression is substituted into equation [7-13] the expression for the total emf in the circuit becomes

$$\mathcal{E}_{\mathrm{Cu,Fe,Pt}} = \left| \pi_{\mathrm{Cu,Fe}} \right|_{t_1}^{t_2} + \int_{t_1}^{t_2} (\sigma_{\mathrm{Cu}} - \sigma_{\mathrm{Fe}})\, dt \qquad [7\text{-}15]$$

which is identical with the total emf (equation [7-6]) of the circuit containing copper and iron alone.

We see from this that a third metal, whose terminals are at the same temperature, can be introduced into a thermocouple circuit without affecting the magnitude of the thermal emf. Similarly, additional metals with isothermal terminals have no effect. This important result, known as the *law of intermediate metals*, makes it possible to include in a thermocouple circuit whatever lead-wires, meters, and other accessories are required. The requirement that the terminals must be at the same temperature is not an inconvenient restriction. In practice, additional metals are inserted in the circuit in the manner shown in Figure 7.7, where t_3 is room temperature. This obviates the necessity for maintaining the additional-metal junctions at the reference temperature.

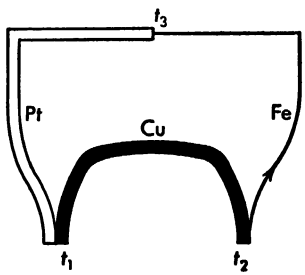


Fig. 7.6

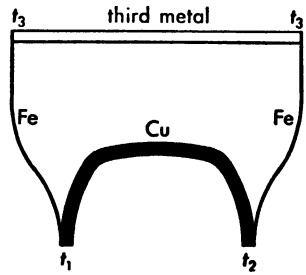


Fig. 7.7

Method of Tabulating Thermal emf's

Let us suppose that thermocouples are composed, in turn, of two of the three metals Cu, Fe, and Pt. The thermocouples operate between the temperatures t_1 and t_2. For the Cu-Fe couple, the thermal emf is given by equation [7-5], and for each of the other two couples by equations of the same type. If these three equations are written down and added, making use of equation [7-14], it is found that

$$\mathcal{E}_{\mathrm{Cu,Fe}} + \mathcal{E}_{\mathrm{Fe,Pt}} + \mathcal{E}_{\mathrm{Pt,Cu}} = 0 \qquad [7\text{-}16]$$

or

$$\mathcal{E}_{\mathrm{Cu,Fe}} = \mathcal{E}_{\mathrm{Cu,Pt}} - \mathcal{E}_{\mathrm{Fe,Pt}} \qquad [7\text{-}17]^{\ddagger}$$

‡ The sign convention is the following: When $\mathcal{E}_{ab}$ is positive in a simple thermocouple circuit the resultant emf is such as to produce a current, at the cold junction, from metal a to metal b. Thus, $\mathcal{E}_{ab} = -\mathcal{E}_{ba}$.

This relationship makes it possible to tabulate the thermoelectric properties of metals by listing the emf that arises when each is combined with some standard metal. This is done in Table 5, in which platinum is taken as the standard and the reference junction is at 0° C.

TABLE 5

Thermal emf's of Metals Relative to Platinum

(Millivolts; cold junction at 0° C)

Temp. of Hot Junction (°C)	*Copper*	*Iron*	*Constantan*	*Pt-Rh* (10%)	*Chromel*	*Alumel*
100	0.76	1.98	−3.51	0.643	2.81	−1.29
200	1.83	3.69	−7.45	1.436	5.96	−2.17
300	3.15	5.03	−11.71	2.316	9.32	−2.89
400	4.68	6.08	−16.19	3.251	12.75	−3.64
500	6.41	7.00	−20.79	4.221	16.21	−4.43
600	8.34	8.02	−25.47	5.224	19.62	−5.28
700	10.49	9.34	−30.18	6.260	22.96	−6.18
800	12.84	11.09	−34.86	7.329	26.23	−7.08
900	15.41	13.10	−39.45	8.432	29.41	−7.95
1000	18.20	14.64	−43.92	9.570	32.52	−8.79
1100	—	—	—	10.741	35.56	−9.58
1200	—	—	—	11.935	38.51	−10.34
1300	—	—	—	13.138	41.35	−11.06
1400	—	—	—	14.337	44.04	−11.77

SOURCE: *American Institute of Physics Handbook*, McGraw-Hill, 1957.

The thermal emf of any thermocouple is obtained from Table 5 by combining the two appropriate entries, as indicated by equation [7-17]. Thus, for example, when the hot junction is at 1000° C, the thermal emf for a chromel-alumel thermocouple is 32.52 − (−8.79) = 41.31 millivolts.

7.7 Thermocouple Thermometers

The principal application of thermocouples has been in the field of thermometry. One of the junctions is kept at the reference temperature while the other is placed in the region whose temperature is to be determined. The unknown temperature is then deduced from the resultant thermal emf, which

is determined with a potentiometer in the manner indicated in Figure 7.8. Among the advantages offered by a thermocouple are its small size, low heat capacity, and rapid response.

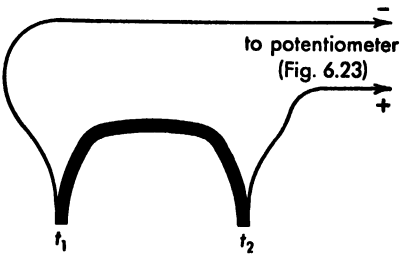


Fig. 7.8

The thermocouple used for very accurate work is composed of pure platinum and a platinum (90 percent) -rhodium (10 percent) alloy. The International Temperature Scale of 1948 is defined in terms of this thermocouple in the temperature range $630.5° < t < 1063.0°$ C. As with the platinum resistance thermometer, it is necessary to calibrate the thermocouple. $\mathcal{E}_{ab}$ and t are assumed to be related as follows:

$$\mathcal{E}_{ab} = a + bt + ct^2 \qquad [7\text{-}18]$$

The constants a, b, and c are found by measuring $\mathcal{E}_{ab}$ for the three standard temperatures—the freezing point of antimony (630.5° C), the freezing point of silver (960.8° C) and the freezing point of gold (1063.0° C). The cold junction is maintained at the ice point.

A thermocouple widely used for general thermometry consists of two nickel-iron alloys, *chromel* and *alumel*. As their names suggest, the former contains chromium and the latter aluminum. It is found that $\mathcal{E}_{ab}$ does not depend upon t in a parabolic manner; rather

$$\mathcal{E}_{ab} = \alpha t + \beta t^2 + \gamma t^3 \qquad [7\text{-}19]$$

If α, β, and γ are evaluated using $t = 100°$ C, $t = 200°$ C and $t = 300°$ C, the couple will be found to be in error by 1° at 50° C, by 0.5° at 150° C, and by 0.2° at 250° C. It must be carefully calibrated over the whole range for which it is to be used. Representative emf's for the chromel-alumel couple, taken from Table 5, are the following: 100° C: 4.10 mV, 200° C: 8.13 mV, 300° C: 12.21 mV, 500° C: 20.64 mV, 1000° C: 41.31 mV, and 1400° C: 55.81 mV. Apparently, $\mathcal{E}_{ab}$ is linear with t to a fair approximation.

Another alloy, *constantan*, is often used in low-temperature work in conjunction with either copper or iron. For many ordinary applications of thermocouples the reference junction is at room temperature rather than at 0° C.

7.8 Thermopiles

An interesting application of the thermocouple is its use as a detector of radiation. Figure 7.9 shows the essential features of a *thermopile*. It consists of many thermocouples in series so as to multiply the thermoelectric effect. One set of junctions is protected from the radiant energy while the other is placed so as to receive it. The small thermoelectric current is measured with a galvanometer. Although the thermopile has a sluggish response to changes in the energy incident upon it, it is very sensitive. It has been used to study the radiation from stars.

7.9 Thermocouple Meters

One of the important types of current meter, illustrated in Figure 7.10, employs the thermocouple principle. The current to be measured passes

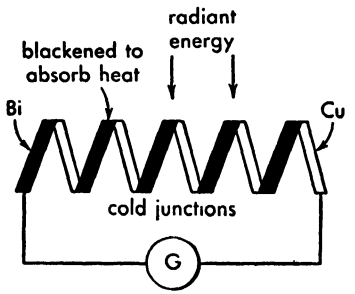


Fig. 7.9

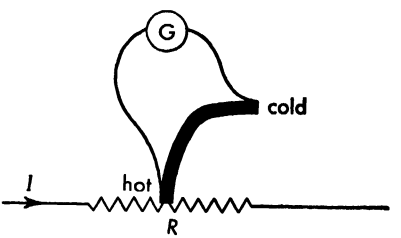


Fig. 7.10

through a resistance where heat is generated in the amount of I^2R joules/sec. The "hot" junction of a thermocouple (whose reference junction is at room temperature) is in thermal contact with this resistance, with the result that a thermoelectric current flows through the galvanometer.

This type of meter has three specific features:

a. It is immaterial whether the input current is AC or DC.

b. Since the thermoelectric current is roughly proportional to the temperature of the hot junction, which in turn is approximately proportional to I^2, the galvanometer reading is roughly proportional to the square of the current being measured.

c. It is particularly useful in the measurement of radio frequency currents, and needs no special calibration for this task, since it can be made to represent an essentially resistive load to the incoming current.

7.10 Thermoelectric Generators and Refrigerators

Generators

Attempts to construct thermoelectric generators using *metallic* thermocouples have not been very successful. In the first place, the thermoelectric power of the most favorable couple, bismuth and antimony, is only slightly greater than 0.1 mV per degree Centigrade. Achievement of appreciable voltages therefore involves the use of a large number of thermocouples and large temperature differences between alternate junctions. In the second place, and much more important, the *efficiency* of a metallic thermocouple is no more than about 1 percent. This is to be compared with ~25 percent for a coal-fired electric generator. The low efficiency arises from the fact that practically all the thermoelectric energy is expended in overcoming the resistance of the thermocouple wires. If the resistance is reduced by use of short thick wires, heat is conducted too quickly from the hot to the cold junctions.

A thermocouple, to be suitable as a generator, must possess high thermoelectric power, high electrical conductivity, and low thermal conductivity. As we have said, metals do not fulfil these requirements—their thermal conductivity is too large and bears a more or less constant relationship to their electric conductivity (Wiedemann-Franz law). On the other hand, in *semiconductors*, the electrical and thermal conductivities can be separately controlled. Also, the thermoelectric power is often substantially higher than for the best metals. Recent work suggests that thermoelectric generators employing semiconductors will soon make their commercial appearance.

A word will be said concerning the method by which the electrical and thermal conductivities are individually altered. In a semiconductor heat is transmitted mainly by lattice vibrations, whereas the electrical conductivity of a *p*- or *n*-type material is largely determined by the number of impurity atoms. The lattice may be slightly distorted by dissolving into it an element or compound that crystallizes in a similar lattice. This slight distortion is sufficient to lower significantly the thermal conductivity but does not interfere (since the lattice periodicity is the same) with the passage of the conduction electron waves. The electrical conductivity is then increased by adding suitable impurities. This cannot be done without limit, since the thermoelectric power in general decreases with conductivity. A compromise between these two effects must be struck.

A promising thermocouple for the generation of electricity is formed of *p*- and *n*-type bismuth telluride (Bi_2Te_3). It appears that this couple will display an efficiency of ~7 percent with $t_2 - t_1 = 300°$ C. The thermo-

electric power is ~0.25 mV per degree Centigrade. This efficiency figure is already within hailing distance of conventional generators and will undoubtedly be further improved.

Russian engineers have developed a low-power generator for use in isolated communities. The thermocouples are built into the shade of a kerosene lamp, which is in use anyway, and produce sufficient electricity to operate a radio.

Experiments have also been made on the use of semiconductor thermocouples as the energy-transformers in devices designed for the generation of electricity from solar energy.

Refrigerators

The Peltier cooling of one of the thermocouple junctions, caused by current passing through it, is being applied to refrigeration problems. For example, with the *p*- and *n*-type Bi_2Te_3 junction mentioned above, temperature differences of 60° C have been obtained between the hot and cold junctions. In Russia, this development has apparently reached the commercial stage. One refrigerator has been described in which a 40-liter chamber has been maintained at 0° C with the expenditure of 55–70 watts.

7.11 Source of Thermal emf

The explanation for the existence of thermal emf's lies in the fact that the conduction electrons in a metal carry *both* heat *and* electricity. As a consequence, there is an interdependence between thermal and electrical effects. The rate of flow of charge, for example, depends upon the temperature gradient as well as upon the potential gradient. Indeed, a current may result from a temperature gradient only, in which case one can ascribe to the temperature gradient a certain effective emf.

The Peltier effect is easily understood on this basis. The conduction electrons possess greater energy in one metal than they possess in the second. As they pass from the one metal to the other, this difference in energy is given up or absorbed, as the case may be.

In the Thomson effect there is a gradient in the electron density within the metal, corresponding to the temperature gradient that exists. Electrons diffuse from regions of high electron density to regions of lower density, giving rise to the observed thermoelectric current. The manner in which the electron density varies with temperature depends, in a way that cannot be explained here, upon the position of the Fermi level in the conduction band. For lead,

there is no variation. For other metals, the electron density sometimes increases with temperature, sometimes decreases. Thus, both positive and negative Thomson effects are observed.

7.12 Problems

E (7.1) When the hot junction of a certain thermocouple (cold junction at 0° C) is raised to 50° C, the thermal emf is 2 mV; when raised to 100° C, it is 3.5 mV. Compute (a) the neutral temperature of the junction; (b) the inversion temperature, at which the emf changes sign, and, (c) the thermoelectric power at 100° C.

[Ans.: (a) 225° C.]

E (7.2) If the cold junction of a copper-constantan thermocouple is kept at 100° C and the other junction at 600° C, what is the resulting thermal emf? Does this differ in magnitude from the emf for cold- and hot-junction temperatures of 0° C and 500° C?

Chapter 8

Moving Charges and Magnetic Fields

In this chapter we begin the study of magnetism, recognizing from the start that all magnetic effects are associated with the motion of electric charges. Before the acceptance of this, the modern point of view, it was assumed that elementary magnets, each consisting of a north pole and a south pole, were actually to be found in nature. These magnetic dipoles were discussed in terms similar to those which we have used in discussing electric dipoles. In this way the subject of *magnetostatics* developed, which was in many respects similar to electrostatics.

Although artificial, magnetostatics is still an important subject of study, provided its artificialities are realized. Indeed, it will be treated at some length in Chapter 10. First, however, we shall turn our attention to the question which is fundamental to all magnetism: the force that is exerted by one electric current (or charge in motion) upon another.

8.1 Force Between Electric Currents

The magnetic effect of an electric current was discovered in 1819 by Hans Christian Oersted (1777–1851), Danish physicist and chemist. He found that a compass needle placed under an electric current takes up a direction perpendicular to that of the current. This observation, which initiated the subject of electromagnetism, was confirmed by André Marie Ampère (1775–1836),

French physicist and mathematician, who rapidly explored, both theoretically and experimentally, many of its implications.

Having witnessed the action of an electric current upon a magnet, it was natural for Ampère to look for the reaction of the magnet upon the current. This he did, establishing that the direction of the force exerted upon a current by a magnetic field (the magnetic field of the earth) is consistent with Newton's third law ("to every action there is always opposed an equal reaction"). He then looked for and found the action of one electric current upon another. Thereafter, on the basis of a brilliant series of experiments, he inferred the law of force between two current elements. We shall now present the general statement of this law without proof. In fact, there is no "proof" of it. Its correctness lies in the fact that experimental evidence substantiates it.

Consider the two current elements dl_i and dl_j (Fig. 8.1), carrying the steady currents I_i and I_j, respectively. The force between these two current elements depends upon the following factors:

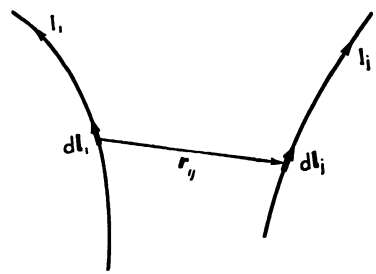


Fig. 8.1

a. It depends upon the nature of the medium.

b. It varies directly as the magnitude of each current.

c. It varies inversely as the square of the distance between the two current elements.

d. It depends upon the lengths and orientations of the two current elements.

The mathematical expression for the force exerted by dl_i upon dl_j is given by

$$dF_{ij} = C[I_i I_j]\left[\frac{1}{r_{ij}^2}\right]\left[dl_j \times \left(dl_i \times \frac{r_{ij}}{r_{ij}}\right)\right] \qquad [8\text{-}1]$$

Let us make a term-by-term examination of this rather formidable expression. The constant C contains not only the factor of proportionality, which depends upon the system of units employed, but it incorporates also the effect of the medium, mentioned above as factor (*a*). Factor (*b*) appears as the product $I_i I_j$. Factor (*c*) appears as the term $1/r_{ij}^2$. The final term represents factor (*d*). The parentheses in the final term enclosing $(dl_i \times r_{ij}/r_{ij})$ indicate that this cross product is to be taken first, after which the cross product between dl_j and the resulting vector quantity is to be taken. The vector (r_{ij}/r_{ij}) is simply a unit vector in the r_{ij} direction.

Equation [8-1] can be arranged to give

$$dF_{ij} = \frac{C[I_j\, dl_j] \times [I_i\, dl_i \times r_{ij}]}{r_{ij}^3} \qquad [8\text{-}2]$$

In this form the law of force between current elements has much in common with the law of force (Coulomb's law—equation [1-2]) between point charges, although the direction of the force, as expressed in equation [8-2], is not so easy to visualize as was the case with the Coulomb force. Before proceeding further with [8-2] let us recall the course of action that was followed in our study of electrostatics.

Although we began our study of electrical forces by calculating directly the forces between individual charges, we soon found it convenient to perform this calculation by means of an intermediary, the electric field. Once the electric field is known for a given charge configuration, the electrical force exerted by the configuration upon any other charge is given by the product of the field and the charge (*Eq*). Similarly, in calculating magnetic effects, rather than computing directly the force exerted by one current upon a second, using equation [8-1], it is usually convenient to employ the intermediary of the magnetic "field." [1] With this in mind we rewrite equation [8-2] in the form

$$d\boldsymbol{F}_{ij} = I_j d\boldsymbol{l}_j \times d\boldsymbol{B}_i \qquad [8\text{-}3]$$

where

$$d\boldsymbol{B}_i = \frac{CI_i\, d\boldsymbol{l}_i \times \boldsymbol{r}_{ij}}{r_{ij}^3} \qquad [8\text{-}4]$$

is the *magnetic induction* or the *magnetic flux density* of the current element $d\boldsymbol{l}_i$ at a vector distance $\boldsymbol{r}_{ij}$. It represents the magnetic effect or "field" that is produced by the first current element. Knowing it, the force that is experienced by the second current element is then easily calculated from equation [8-3].

8.2 Force Exerted by a Magnetic Field upon a Charge in Motion

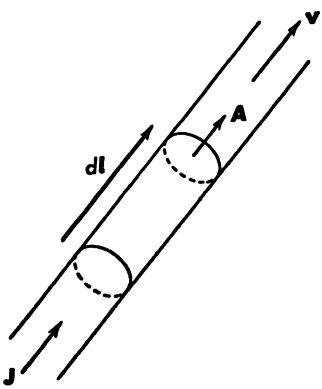


Fig. 8.2

Before proceeding to calculate $\boldsymbol{B}$ for typical current configurations, we shall extend equation [8-3] to include the case of a charge in motion in a magnetic field. This we shall do by visualizing the electric current flowing in the current element $d\boldsymbol{l}$ as a group of charges moving with velocity $\boldsymbol{v}$ (Fig. 8.2).

The force experienced by this current element in a field of induction $\boldsymbol{B}$ is given by equation [8-3], viz.,

$$d\boldsymbol{F} = I\, d\boldsymbol{l} \times \boldsymbol{B} \qquad [8\text{-}5]$$

[1] The reason for the quotation marks will be understood after reading § 8.3.

The current density, J, in the conductor is given by

$$J = nqv \qquad [5\text{-}4]$$

where n is the number of charge carriers per unit volume and q is the charge carried by each. The total current flowing in the conductor is

$$I = J \cdot A = nqv \cdot A \qquad [8\text{-}6]$$

Substituting equation [8-6] in [8-5] we find the force exerted on the current element dl to be

$$dF = (nqv \cdot A)\, dl \times B \qquad [8\text{-}7]$$

or

$$dF = (nq\, dl \cdot A)v \times B \qquad [8\text{-}8]$$

Here $dl \cdot A$ is the volume of the current element and, consequently, $n\, dl \cdot A$ is the number of charge carriers contained within it. Hence, it follows that the force experienced by a single charge carrier in the current element is

$$F \text{ (on charge } q) = \frac{dF}{(n\, dl \cdot A)} = qv \times B \qquad [8\text{-}9]$$

where v is the velocity of the charge and B is the field of induction in which the charge moves.

Equation [8-9] may be regarded as a third form of the law of force between currents. We thus have the three versions of the law: (*a*) equation [8-1] giving the force exerted by one current element upon a second, (*b*) equation [8-3] giving the force experienced by a current element when it is in a field of induction B, and (*c*) equation [8-9] giving the force experienced by a charge in motion in a field of induction B. Equation [8-9] holds for *any* charge in motion, whether or not it is confined to a conductor. Since the force in [8-9] is always normal to the direction of motion of the charge, the speed of the charge is not altered by the force.

Lorentz Force Equation

Equation [8-9] permits us to write down the *total* force acting upon a charge in a region that is pervaded by electric and/or magnetic fields. Thus

$$F = Eq + qv \times B \qquad [8\text{-}10]$$

where E, the electric field strength, has been defined in § 1.4 as the electric force per unit charge. Equation [8-10] is of fundamental importance: it is known as the *Lorentz force equation.* It serves as the basis for Problems [8.8], [8.9], and [8.11].

In the latter part of the chapter we shall study further the forces between electric currents, and those experienced by currents and moving charges in a

magnetic field. Prior to that, however, we shall calculate the magnetic induction produced by a variety of current configurations. This will give us some idea of the manner in which magnetic fields are established, and should give substance to our subsequent discussion of magnetic forces.

8.3 Magnetic "Fields"

As mentioned above, we designate by the symbol $\boldsymbol{B}$ the quantity known as the *magnetic induction*, or the *magnetic flux density*. It is the property of the magnetic field that determines the force that is exerted upon a current or a moving charge in the field. Later we shall also refer to the *magnetic intensity*, or *field strength*, $\boldsymbol{H}$, and the *permeability*, μ. These three quantities are related as follows:

$$\boldsymbol{B} = \mu \boldsymbol{H} \qquad [8\text{-}11]$$

corresponding to the electrical analogue

$$\boldsymbol{D} = \epsilon \boldsymbol{E} \qquad [4\text{-}21]$$

This correspondence, a legacy from magnetostatics, is misleading. In a dielectric medium the electrical force exerted upon a free charge is proportional to $\boldsymbol{E}$, not $\boldsymbol{D}$. On the other hand, in a magnetic material, the magnetic force that is exerted upon a moving charge is proportional to $\boldsymbol{B}$, not $\boldsymbol{H}$. Thus, in the sense that $\boldsymbol{E}$ is the electric field for free charges, $\boldsymbol{B}$ is the magnetic "field" for moving charges. Unfortunately, $\boldsymbol{H}$ bears the title that describes $\boldsymbol{B}$'s function.

We shall make a point, therefore, of calculating the magnetic induction $\boldsymbol{B}$, rather than the magnetic intensity $\boldsymbol{H}$, for the several current configurations that are studied in this chapter.

8.4 The Magnetic Induction of a Current Element: Ampère's Law

The magnetic induction, $d\boldsymbol{B}$, of a current element has been defined above in equation [8-4]:

$$d\boldsymbol{B} = \frac{C I \, d\boldsymbol{l} \times \boldsymbol{r}}{r^3} \qquad [8\text{-}4]$$

This expression gives the value of the induction at the point P (Fig. 8.3), arising from the current I in the element of a conductor of vector length $d\boldsymbol{l}$. C is a factor of proportionality whose value depends upon the system of units that is being employed and upon the nature of the medium in which the induction is produced.

Equation [8-4], which gives the strength of the magnetic "field" produced by a current element, is usually known as Ampère's law, despite the fact that it only gives part of the story of the force between current elements. The reasonableness of the law is apparent when one examines the magnitude of dB,

$$dB = \frac{CI\,dl \sin\theta}{r^2} \qquad [8\text{-}12]$$

Thus, dB increases linearly with the current I, it follows an inverse square law (i.e., $\propto 1/r^2$), and it is proportional to the component ($dl \sin\theta$) of the element in the plane normal to the radius vector. The *total* induction of a current-carrying conductor is found by appropriate integration of equation [8-4] or equation [8-12].

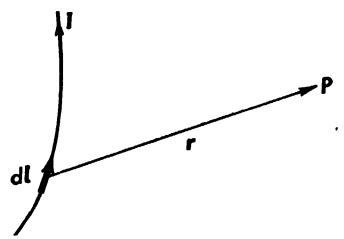


Fig. 8.3

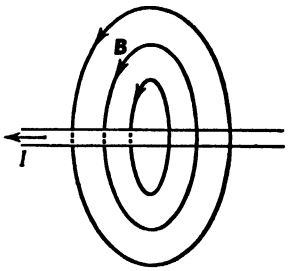


Fig. 8.4

The Right-hand Rule for the Direction of B

The direction of $d\boldsymbol{B}$ at the point P is given by $d\boldsymbol{l} \times \boldsymbol{r}$. Evidently, $d\boldsymbol{B}$ is normal to both $d\boldsymbol{l}$ and $\boldsymbol{r}$: in Figure 8.3 it is directed into the paper.

The direction of the induction produced by a current is also given by the *right-hand rule*, which may be stated as follows: if a current-carrying conductor is grasped in the right hand with the thumb pointing in the direction of the (conventional) current, the fingers will point in the direction of $\boldsymbol{B}$. The magnetic lines of force, to which $\boldsymbol{B}$ is everywhere tangent, thus form concentric circles around the conductor (Fig. 8.4) and cause, as Oersted discovered, a compass needle located nearby to take up a direction perpendicular to that of the current or parallel to that of the lines of force.

Ampère's Law in the mks System of Units

In the mks system of units the magnetic induction or flux density, $\boldsymbol{B}$, is expressed in *webers per square meter*.[1] By convention the factor of proportionality,

[1] In the next chapter we shall introduce the concept of *magnetic flux*, Φ_m, which is the surface integral of $\boldsymbol{B}$, and which is expressed in *webers*. This accounts for our occasional use of the term *flux density* for $\boldsymbol{B}$, and for the units of webers/m^2.

C, in equation [8-4] is written as

$$C = \frac{\mu}{4\pi} \qquad [8\text{-}13]$$

Thus, the mks expression of Ampère's law is

$$dB = \frac{\mu I}{4\pi} \frac{d\boldsymbol{l} \times \boldsymbol{r}}{r^3} \text{ webers/m}^2 \qquad [8\text{-}14]$$

where, as usual, I is the current expressed in amperes. The factor 4π in the denominator, although appearing unnecessarily to complicate equation [8-14], leads to the maximum *over-all* simplicity in the expressions relating magnetism and electricity. The use of μ in equation [8-14] and in other equations to follow implies that the surrounding medium is either free space or a homogeneous, isotropic material medium occupying the entire space pervaded by the magnetic field. As in the electrostatic case, cavities or other discontinuities in the medium alter the simple conditions described by equation [8-14].[8]

In free space $\mu = \mu_0$ and has the value

$$\mu_0 = 4\pi \times 10^{-7} \text{ webers/ampere-meter} \qquad [8\text{-}15]$$

This is not an arbitrarily assigned value, but follows from the definition of the ampere (§ 8.13). The ratio $k_m = \mu/\mu_0$ is known as the *magnetic coefficient* or *relative permeability*.

8.5 Magnetic Induction at the Center of a Circular Current: The Electromagnetic System of Units

The magnetic induction at the center of a circular current (Fig. 8.5) is readily calculated by use of Ampère's law, in the form, say, of equation [8-14]. Since $d\boldsymbol{l}$ and $\boldsymbol{r}$ are normal to one another, the magnitude of the induction at O arising from the current element $d\boldsymbol{l}$ is given by

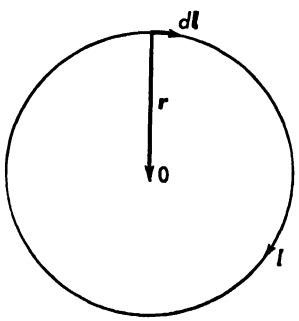


Fig. 8.5

$$dB = \frac{\mu I\, dl}{4\pi r^2} \qquad [8\text{-}16]$$

The total induction at O is then

$$B = \oint dB = \frac{\mu I}{4\pi r^2} \oint dl = \frac{\mu I}{2r} \text{ webers/m}^2 \qquad [8\text{-}17]$$

[8] See "Report of Coulomb's Law Committee of the American Association of Physics Teachers," *American Journal of Physics*, **18**, p. 15 ff. (1950).

pointing into the page (i.e., in the direction of $d\boldsymbol{l} \times \boldsymbol{r}$). If there are N closely grouped turns of wire, rather than the single one shown in Figure 8.5, the magnetic induction at O is

$$B = \frac{\mu NI}{2r} \text{ webers/m}^2 \qquad [8\text{-}18]$$

The Electromagnetic System of Units

Although the mks system of units has official sanction and is becoming widely used, there is still a large body of scientific literature in which magnetic effects are discussed in *electromagnetic* units. The unit of current in the electromagnetic system is defined in terms of the induction which it produces at the center of a circular loop.

Let us calculate the total induction at the center of the circular current of Figure 8.5, using the general expression for $d\boldsymbol{B}$ (equation [8-4]) rather than the specifically mks one which was used above. We find

$$B = \oint dB = C\frac{2\pi I}{r} \qquad [8\text{-}19]$$

In the electromagnetic system of units the constant in the above equation is itself called the permeability. It is a different quantity, of course, from the permeability in the mks system and, for that reason, will be designated by a distinctive symbol, namely, μ_{em}. Magnetic induction in the electromagnetic system is expressed in *gauss*. Equation [8-19], giving the value of the induction at the center of a circular current, may now be rewritten in terms of electromagnetic units:

$$B = \mu_{em}\frac{2\pi I}{r} \text{ gauss} \qquad [8\text{-}20]$$

The electromagnetic unit of current is defined in such a way as to make $\mu_{em} = 1$ for free space. Thus, the *electromagnetic unit of current is that current which, when present in a circular loop of radius 1 cm, gives rise to a magnetic induction at the center of 2π gauss.*[4] This current is known variously as the *absolute electromagnetic unit of current*, the *abampere*, and the *cgsm unit* of *current*. The corresponding charge, known as the *emu of charge*, the *abcoulomb*, or the *cgsm unit of charge*, is that quantity of charge which passes in 1 second when the current is 1 abampere.

The mks and cgsm units of current and charge are related as follows:

$$1 \text{ ampere} = 0.1 \text{ abamperes}$$

and

$$1 \text{ coulomb} = 0.1 \text{ abcoulombs} \qquad [8\text{-}21]$$

[4] The unit of magnetic induction in the cgsm system, the gauss, is defined in Chapter 10 and bears the following relationship to the corresponding mks unit:

$$1 \text{ weber/m}^2 = 10^4 \text{ gauss}$$

These relationships, it will be borne in mind, do not constitute the definitions of the ampere and the coulomb; these are defined in § 8.13.

Reference was made in Chapter 1 to a third system of units, the electrostatic or cgse system, which grew out of the study of electrostatic phenomena. Both the electromagnetic and the electrostatic systems are cgs (cm-gm-sec) systems, but the magnitudes of unit current and unit charge in the two systems differ vastly:

$$1 \text{ abampere} = 3 \times 10^{10} \text{ statamperes}$$

and

$$1 \text{ abcoulomb} = 3 \times 10^{10} \text{ statcoulombs} \qquad [8\text{-}22]$$

The conversion factor, written here as 3×10^{10}, is the velocity of light, c, in the cgs system, a more accurate value for which is $c = 2.9979 \times 10^{10}$ cm/sec. The observation that the ratio of the cgsm and cgse units of charge closely approximates the velocity of light was made in 1856 by Weber and Kohlrausch. It was in search of the explanation of this coincidence that Maxwell developed the electromagnetic theory of light, thereby establishing not only that the relationship was exact but also the theoretical basis for it. Other relationships between cgse and cgsm quantities are established in Problem 8.1.

Because of the widespread use of electromagnetic units and the known reluctance of individuals to give up familiar usages, even in the interests of uniformity, some of the more important results that will be developed in this text are expressed in cgsm units in addition to the standard mks form. This should help to make existing literature comprehensible and should also facilitate the transition to uniform mks practice. Persons wishing to climb immediately onto the mks bandwagon may simply ignore the electromagnetic equivalents.

8.8 Magnetic Induction Along Axis of a Circular Current

Ampère's law will now be used to calculate the magnetic induction at the point P (Fig. 8.6) along the axis of a circular current.

As $d\boldsymbol{l}$ and $\boldsymbol{r}$ are normal to one another, the induction at P caused by the current element $d\boldsymbol{l}$ is

$$d\boldsymbol{B} = \frac{\mu I \, d\boldsymbol{l}}{4\pi r^2} \qquad [8\text{-}23]$$

It will be seen that dB_y, the component of $d\boldsymbol{B}$ that is normal to the axis, is canceled by the normal-to-the-axis component arising from the current element diametrically opposite to $d\boldsymbol{l}$. The total induction P is, therefore, the sum of the individual *axial* components:

$$B = \oint dB_x = \oint dB \cos \alpha = \oint dB \sin \beta \qquad [8\text{-}24]$$

whence,

$$B = \frac{\mu I \sin \beta}{4\pi r^2} \oint dl \qquad [8\text{-}25]$$

After integration and substitution of $r^2 = a^2 + p^2$ and $\sin \beta = a/r$, equation [8-25] becomes, for N turns

$$B = \frac{\mu N a^2 I}{2(a^2 + p^2)^{3/2}} \text{ webers/m}^2 \qquad [8\text{-}26]$$

or, in the electromagnetic system,

$$B = \frac{\mu_{em} 2\pi N a^2 I}{(a^2 + p^2)^{3/2}} \text{ gauss} \qquad [8\text{-}26 \text{ cgsm}]$$

These equations reduce, when $p = 0$, to the expressions given in the preceding section for the magnetic induction at the center of a circular loop.

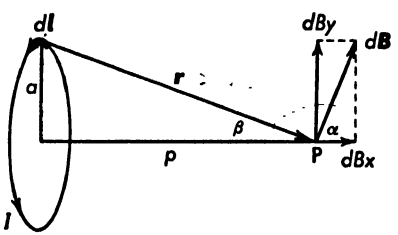


Fig. 8.6

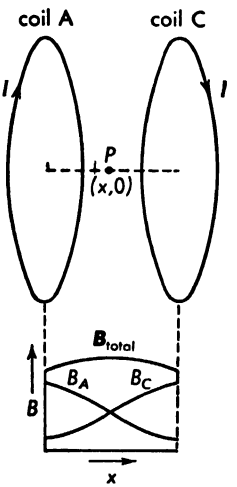


Fig. 8.7

Helmholtz Coils

Two thin coaxial coils, each containing the same number of turns and carrying the same current (Fig. 8.7), can be used to obtain a nearly constant value of induction over a considerable distance along the axis.

Let us first consider the manner in which the induction of a single coil falls off as one moves outward along the axis. At any point P, located a distance

x along the axis from the center of coil A, the induction of coil A, of radius a, is given by the equation [8-26]:

$$B = \frac{\mu N a^2 I}{2(a^2 + x^2)^{3/2}} \qquad [8\text{-}26]$$

The rate of change of B with x is, of course, dB/dx. It is shown in Problem 8.2 that, at the point $x = a/2$, $\frac{d}{dx}\left(\frac{dB}{dx}\right) = 0$. Thus, at the point $x = a/2$, the quantity dB/dx becomes constant, i.e., the B-versus-x curve is a straight line. Actually, at the point $x = a/2$ there is a turning point in the curve, with the result that the curve in the neighborhood of the point approximates a straight line.

If two identical coaxial coils are placed with their centers a distance apart equal to their radii, as in Figure 8.7, not only at the point midway between the two coils, but for some distance on either side of this point, dB_A/dx and dB_C/dx are equal in magnitude and opposite in sign. As a result, as shown in the bottom portion of Figure 8.7, the *total* induction B_{total}, is constant over this range. Two such coils, thus arranged with respect to one another, are known as *Helmholtz coils* and are often employed to secure uniform magnetic fields. Evidently, at the mid-point

$$B_{\text{total}} = \frac{\mu N a^2 I}{\left(a^2 + \frac{a^2}{4}\right)^{3/2}}$$

or

$$B_{\text{total}} = \frac{8\mu N I}{a(5)^{3/2}} \text{ webers/m}^2 \qquad [8\text{-}27]$$

The electromagnetic equivalent is

$$B_{\text{total}} = \frac{2.86\mu_{\text{em}}\pi N I}{a} \text{ gauss} \qquad [8\text{-}27 \text{ cgsm}]$$

8.7 Magnetic Induction Along Axis of a Solenoid

We shall now calculate the value of the induction at a point P on the axis of the solenoid shown in Figure 8.8. The solenoid carries a current of I amperes and contains a total of N turns of wire. As there are $n(= N/l)$ turns per unit length, the number of turns in the element of the solenoid of length dx is $n\,dx$. From equation [8-26] the induction at P arising from this element is

$$dB = \frac{\mu n\,dx\,a^2 I}{2(a^2 + x^2)^{3/2}} \qquad [8\text{-}28]$$

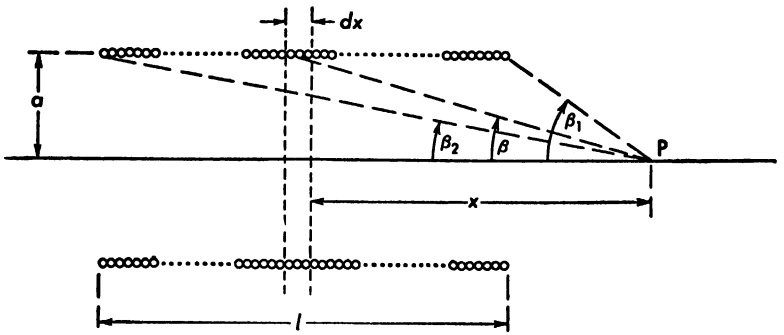


Fig. 8.8

If we now substitute the relationships

$$x = a \operatorname{cotan} \beta \qquad \text{and} \qquad dx = -a \operatorname{cosec}^2 \beta \, d\beta \qquad \text{[8-29]}$$

equation [8-28] may be rewritten

$$dB = -\frac{\mu n I}{2} \sin \beta \, d\beta \qquad \text{[8-30]}$$

The total induction at P of the entire solenoid is

$$B = -\frac{\mu n I}{2} \int_{\beta_1}^{\beta_2} \sin \beta \, d\beta = \frac{\mu n I}{2} (\cos \beta_2 - \cos \beta_1) \text{ webers/m}^2 \qquad \text{[8-31]}$$

The equivalent cgsm expression is

$$B = 2\pi \mu_{em} n I (\cos \beta_2 - \cos \beta_1) \text{ gauss} \qquad \text{[8-31 cgsm]}$$

At the center of a solenoid whose length is large compared to its radius, $\cos \beta_2 \doteq 1$ and $\cos \beta_1 \doteq -1$, whence

$$B \doteq \mu n I \text{ webers/m}^2 \qquad \text{[8-32]}$$

At one end of the coil, $\cos \beta_2 \doteq 1$ and $\cos \beta_1 = 0$, whence

$$B \doteq \frac{\mu n I}{2} \text{ webers/m}^2 \qquad \text{[8-33]}$$

which is one-half of its value at the center.

8.8 Magnetic Induction of a Straight Wire Carrying a Current

The induction at the point P, shown in Figure 8.9, caused by the current element dx, has the magnitude (from Ampère's law—equation [8-14])

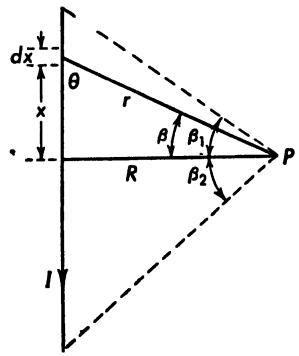


Fig. 8.9

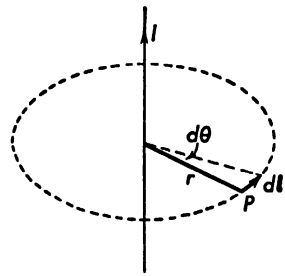


Fig. 8.10

$$dB = \frac{\mu I \sin\theta \, dx}{4\pi r^2} \qquad [8\text{-}34]$$

$$dB = \frac{\mu I \cos\beta \, dx}{4\pi r^2} \qquad [8\text{-}35]$$

and is directed out of the page. If we substitute $dx = R \sec^2\beta \, d\beta$ (from $x = R \tan\beta$) and $r = R \sec\beta$, equation [8-35] becomes

$$dB = \frac{\mu I}{4\pi R} \cos\beta \, d\beta \qquad [8\text{-}36]$$

If the lines joining P to the ends of the wire are inclined at angles β_1 and β_2 to the perpendicular from P to the wire, the total induction at P is

$$B = \frac{\mu I}{4\pi R} \int_{-\beta_2}^{\beta_1} \cos\beta \, d\beta = \frac{\mu I}{4\pi R} (\sin\beta_1 + \sin\beta_2) \text{ webers/m}^2 \qquad [8\text{-}37]$$

When the wire is infinitely long, $\sin\beta_1 = \sin\beta_2 = 1$, in which case

$$B = \frac{\mu I}{2\pi R} \text{ webers/m}^2 \qquad [8\text{-}38]$$

The corresponding cgsm expression is

$$B = \frac{2\mu_{em} I}{R} \text{ gauss} \qquad [8\text{-}38 \text{ cgsm}]$$

Equation [8-38], describing the magnetic field of a long, straight, current-carrying conductor, is known as the *law of Biot and Savart*, named after two of Ampère's contemporaries whose experiments established it. Problems 8.3 to 8.7 are applications of it.

8.9 Line Integral of the Induction: Ampère's Law in Circuital Form

In studying electrostatics we noted that

$$\oint \boldsymbol{E} \cdot d\boldsymbol{l} = 0 \qquad [2\text{-}26]$$

since electrostatic potential is a single-valued function of position. We shall now give thought to the analogous magnetic integral, namely, the integral of $\boldsymbol{B} \cdot d\boldsymbol{l}$ around a closed path.

Let us assume, for simplicity, that the induction under consideration arises from a current in a long wire, as shown in Figure 8.10. The value of induction at the point P is given by equation [8-38] as

$$B = \frac{\mu I}{2\pi r} \qquad [8\text{-}38]$$

We propose to evaluate $\boldsymbol{B} \cdot d\boldsymbol{l}$ around the circular path shown in Figure 8.10. In this case, since $\boldsymbol{B}$ and $d\boldsymbol{l}$ are parallel, the integral can be written

$$\oint \boldsymbol{B} \cdot d\boldsymbol{l} = \frac{\mu I}{2\pi} \oint \frac{r\, d\theta}{r} = \frac{\mu I}{2\pi} \oint d\theta = \mu I \qquad [8\text{-}39]^{6}$$

The sign of the integral depends, of course, upon the direction in which the current is encircled. It is positive if a route parallel to $\boldsymbol{B}$ is followed, negative if antiparallel.

The result is known as *Ampère's law in circuital form.* Although it is derived here with the benefit of certain simplifying assumptions, notably that the current is infinitely long, the same result is got for any shape of current, although the integration is more difficult. In words, the law states that *the line integral of the induction around any closed path is numerically equal to the permeability multiplied by the current through the surface that is bounded by the closed path.* Some examples of the use of this form of Ampère's law are given towards the end of this section.

Equation [8-39] can also be written in the form

$$\oint \boldsymbol{H} \cdot d\boldsymbol{l} = I \qquad [8\text{-}40]$$

since $\boldsymbol{H}$ is defined as $\boldsymbol{B}/\mu$ (equation [8-11]). This version of Ampère's law in circuital form has the advantage that it does not depend upon the nature of the medium.

Although the use of I (rather than i) for the conduction current in equa-

[6] It can be seen from Figure 8.10 that the magnitude of $d\boldsymbol{l}$ is $r\, d\theta$.

tions [8-39] and [8-40] should make this point clear, it is emphasized that these equations only hold in the presence of steady electric fields (see § 8.10).

Magnetic Scalar Potential

The first observation to be made concerning equation [8-39] is that the line integral of the magnetic induction around a closed path does *not* vanish unless it encloses zero current. On the other hand, the line integral of the electric field (equation [2-26]) always vanishes for static cases. It follows, therefore, that if we were to define a *magnetic scalar potential* (corresponding to the electric potential V) from which $\boldsymbol{B}$ could be derived (corresponding to $\boldsymbol{E} = -\nabla V$), this potential would *not* be a single-valued function of position unless the region were current-free. Such a potential has been employed (e.g., § 10.4) but because of this shortcoming it is not of general use. It is not to be confused with the *magnetic vector potential* which is introduced in § 8.11.

Use of Ampère's Circuital Law in Calculating Fields of Current Configurations

Equation [8-39] is useful in computing the magnetic induction arising from certain symmetric current configurations. In this regard it resembles Gauss' theorem, which was seen in § 1.9 to provide a powerful method of computing the electric fields belonging to symmetric charge configurations.

First let us study the *long solenoid* shown in Figure 8.11. It is proposed to

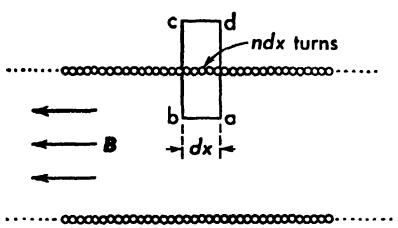


Fig. 8.11

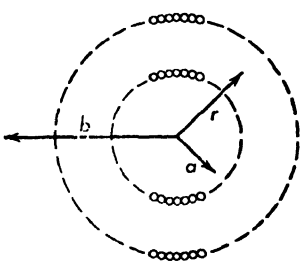


Fig. 8.12

evaluate the line integral of the induction around the closed loop *abcd*, which is located near the middle of the long solenoid.

We know from the symmetry of the situation that the induction inside and far from the ends of the solenoid is parallel to the axis. Outside the solenoid the induction is very weak. It is in this external region that the lines of force which are crowded into the solenoid execute their return paths, and they are so spread out that their density is negligibly small. We can, therefore, conclude

that the integral of $\boldsymbol{B}$ along the paths bc and da is zero (since $\boldsymbol{B}$ is normal to both), and is also zero along the path cd (since $\boldsymbol{B}$ there is negligibly small). Ampère's circuital law,

$$\oint \boldsymbol{B} \cdot d\boldsymbol{l} = \mu \times \text{total current enclosed} \qquad [8\text{-}39]$$

therefore, gives

$$B\,dx \doteq \mu(n\,dx\,I) \qquad [8\text{-}41]$$

Since n is the number of turns per unit length, $n\,dx\,I$ is the total current encircled in this particular case. From equation [8-41] we see that the induction at or near the middle of the long solenoid is

$$B \doteq \mu n I \text{ webers/m}^2 \qquad [8\text{-}32]$$

in agreement with the result previously derived.

Incidentally, we conclude from this, since dx is an indefinite distance from the axis, that the value of the induction inside a solenoid is independent of the distance from the axis. This is a useful piece of information which did not follow from our previous discussion of solenoids in § 8.7.

As another example of the use of equation [8-39], we shall consider the magnetic field within and surrounding a ring or toroidal solenoid (Fig. 8.12). Around the circular path of radius r, lying within the solenoid, the induction is constant and normal to the radius. Equation [8-39] is written as

$$\oint \boldsymbol{B} \cdot d\boldsymbol{l} = B 2\pi r = \mu N I \qquad [8\text{-}42]$$

where N is the total number of turns in the solenoid. Solving equation [8-42] for B, we find:

$$B = \mu I \left(\frac{N}{2\pi r}\right) = \mu n I \qquad [8\text{-}43]$$

everywhere *within* the solenoid. Outside the solenoid, as represented by closed paths of radii a and b, respectively, the situation is quite different. Around the former path, zero current is encircled, and the induction is consequently zero. Around the path of radius b, the total current encircled is $N(I - I) = 0$, since the current has opposite directions in the inner and outer portions of the coils. Hence, in this region also, the induction is zero. In brief, the induction belonging to a toroidal solenoid lies entirely within the solenoid.

In deriving equation [8-43] we have disregarded the fact that r is not constant for all points within the solenoid. Actually, in conformity with equation [8-42], $\boldsymbol{B}$ is larger for paths of smaller r. In practice, however, provided the radius of the mean, interior, circular path is large compared to the radius of the solenoid itself, $\boldsymbol{B}$ is nearly uniform.

The long, straight solenoid discussed earlier can be regarded as a toroidal solenoid with $r = \infty$. Thus, at internal points far from the ends of the straight solenoid, the induction is given by equation [8-43] (which is identical with our earlier result—equation [8-32]) and is uniform across the solenoidal cross section.

8.10 Maxwell's Second Equation

Ampère's law expressed in circuital form is a convenient point of departure from which to obtain Maxwell's second equation. As we have seen, for *steady* electric fields,

$$\oint \boldsymbol{H} \cdot d\boldsymbol{l} = I \qquad [8\text{-}40]$$

or

$$\oint \boldsymbol{H} \cdot d\boldsymbol{l} = \int_s \boldsymbol{J} \cdot d\boldsymbol{s} \qquad [8\text{-}44]$$

where the integral on the right is taken over the surface through which the charge flow corresponding to the current I takes place. In equation [8-44], $\boldsymbol{J}$, the current density, refers to the usual conduction current. We recall, however, that another type of current, the displacement current (§ 4.6), can exist if the electric field is changing. The density of displacement current is seen, from equation [4-30], to be

$$\boldsymbol{J}_D = \frac{\partial \boldsymbol{D}}{\partial t} \qquad [8\text{-}45]$$

We can therefore modify equation [8-44] to cover situations in which the electric field is not constant by adding to the right-hand side of the equation a term representing the displacement current. Thus,

$$\oint \boldsymbol{H} \cdot d\boldsymbol{l} = \int_s (\boldsymbol{J} + \boldsymbol{J}_D) \cdot d\boldsymbol{s} = \int_s \left(\boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} \right) \cdot d\boldsymbol{s} \qquad [8\text{-}46]$$

is a relationship of *general* applicability.

Equation [8-46] can be expressed in a different form by use of Stokes' theorem (see Appendix), which is useful in transforming a line integral into a surface integral, or vice versa. According to it, for any vector, $\boldsymbol{H}$,

$$\oint \boldsymbol{H} \cdot d\boldsymbol{l} = \int_s (\nabla \times \boldsymbol{H}) \cdot d\boldsymbol{s} \qquad [8\text{-}47]$$

The theorem stated in words is the following: *the line integral of any vector around*

a closed path is equal to the surface integral of the curl of the vector over any surface bounded by the closed path.

Equation [8-46] can now be rewritten

$$\int_s (\nabla \times \boldsymbol{H}) \cdot d\boldsymbol{s} = \int_s \left(\boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} \right) \cdot d\boldsymbol{s} \qquad [8\text{-}48]$$

whence

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} \qquad [8\text{-}49]$$

Equation [8-49] is *Maxwell's second equation.* In a conductor, the second term on the right-hand side vanishes, while in a dielectric it is the first term on the right-hand side which vanishes. As mentioned earlier, the justification for including a displacement current term in equation [8-46] *et seq.* is that predictions based on the so-modified equation are borne out by experience.

8.11 Magnetic Vector Potential

Although it is not generally possible to derive $\boldsymbol{B}$ from a scalar potential (§ 8.9), we shall now show that a *vector potential* exists from which $\boldsymbol{B}$ can be derived.

The magnetic induction at a distance r from a current element $I\,dl$ is given by Ampère's law

$$d\boldsymbol{B} = \frac{\mu I}{4\pi} \frac{d\boldsymbol{l} \times \boldsymbol{r}}{r^3} \qquad [8\text{-}14]$$

Noting that

$$\nabla \left(\frac{1}{r} \right) = \left[i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z} \right] (x^2 + y^2 + z^2)^{-\frac{1}{2}} \qquad [8\text{-}50]$$

$$= -\frac{1}{2} \frac{(i2x + j2y + k2z)}{(x^2 + y^2 + z^2)^{\frac{3}{2}}} = \frac{-\boldsymbol{r}}{r^3} \qquad [8\text{-}51]$$

we may rewrite Ampère's law in the form

$$d\boldsymbol{B} = \frac{\mu I}{4\pi} \nabla \left(\frac{1}{r} \right) \times d\boldsymbol{l} \qquad [8\text{-}52]$$

Before tran forming this expression further, we digress briefly to discuss the basic rule for performing vector operations. As was implied in § 2.1, the expression $\boldsymbol{AB}$ is not a sensible one inasmuch as $\boldsymbol{A}$ and $\boldsymbol{B}$ are both vectors and, as such, can only be multiplied together through a scalar (dot) or vector (cross) product. Similarly, because of its vector character, ∇ can only operate upon the vector $\boldsymbol{A}$ to form the "divergence" $(\nabla \cdot \boldsymbol{A})$ or the "curl" $(\nabla \times \boldsymbol{A})$ of the vector. On the other hand, ∇ can operate upon a scalar, V, in one way only—to form the gradient (∇V). The gradient is, of course, a vector. Let us now consider the operation

$$\nabla \times (V\boldsymbol{A})$$

This evidently involves vector differentiation of the product VA. In performing this operation, care must be taken to ensure that ∇ and A are always related through the cross product, and in the same order in which they appear above. The position of V relative to the "cross" is immaterial since V has no vector character. The vector operation therefore takes the form

$$\nabla \times (VA) = (\nabla V) \times A + V(\nabla \times A) \qquad [8\text{-}53]$$

A similar operation upon dl/r yields

$$\nabla \times \left(\frac{dl}{r}\right) = \nabla \left(\frac{1}{r}\right) \times dl + \frac{1}{r}(\nabla \times dl) \qquad [8\text{-}54]$$

whence

$$\nabla \left(\frac{1}{r}\right) \times dl = \nabla \times \left(\frac{dl}{r}\right) - \frac{1}{r}(\nabla \times dl) \qquad [8\text{-}55]$$

In the light of this relationship, Ampère's law may be transformed from the form of equation [8-52] to

$$dB = \frac{\mu I}{4\pi}\left[\nabla \times \left(\frac{dl}{r}\right) - \frac{1}{r}(\nabla \times dl)\right] \qquad [8\text{-}56]$$

Since dl is not a function of the coordinates (x,y,z) of the point at which the induction is being calculated, and since the operation ∇ represents differentiation with respect to these coordinates, the second term on the right-hand side of equation [8-56] is zero. Thus, the total induction produced at the point in question by a closed current is

$$B = \frac{\mu I}{4\pi} \oint \nabla \times \left(\frac{dl}{r}\right) \qquad [8\text{-}57]$$

Further, since the operation $\nabla \times$ is independent of the integration of dl/r around the closed current, equation [8-57] may be rewritten

$$B = \nabla \times \left[\frac{\mu I}{4\pi} \oint \frac{dl}{r}\right] \qquad [8\text{-}58]$$

The quantity within the square brackets is obviously a vector.

From this result we conclude that a vector exists such that, by taking its curl, the induction produced at any point by a closed current may be obtained. This vector, designated as A, is known as the *vector potential*. Hence,

$$B = \nabla \times A \qquad [8\text{-}59]$$

where

$$A = \frac{\mu I}{4\pi} \oint \frac{dl}{r} \qquad [8\text{-}60]$$

We noted earlier (§ 8.9) that, whenever currents are present, B cannot be derived from a *scalar* potential. A vector potential, however, always exists from which B can be derived in accordance with equation [8-59]. Frequently it is simpler to obtain B from the vector potential than it is to compute it directly. The situation is analogous to the electrostatic case, in which the electric field is often most easily obtained from the electric potential.

Use of the Vector Potential

Although the notion of a magnetic vector potential will not be much exploited in this text, a single example[6] will be given of its usefulness in calculating $\boldsymbol{B}$.

Figure 8.13 depicts a square, closed current, of side a, lying in a vacuum in the X-Z plane with its center at the origin of coordinates. We wish to calculate the com-

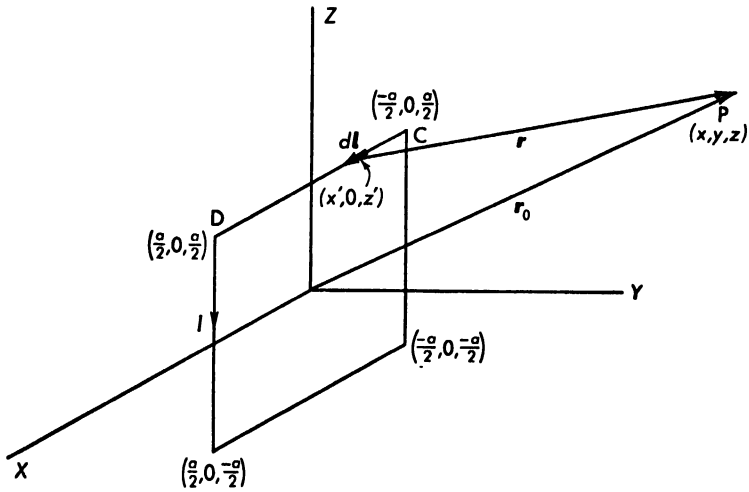


Fig. 8.13

ponents of the magnetic induction at the point $P(x,y,z)$. To do this we first compute the vector potential at P.

As we have seen, the vector potential is given (equation [8-60]) by

$$A = \frac{\mu_0 I}{4\pi} \oint \frac{dl}{r} \qquad [8\text{-}61]$$

This, written out for the square, closed current shown in Figure 8.13, becomes

$$A = \frac{\mu_0 I}{4\pi}\left[\int_{-a/2}^{a/2} \frac{\boldsymbol{i}\,dx'}{r} + \int_{a/2}^{-a/2} \frac{\boldsymbol{k}\,dz'}{r} + \int_{a/2}^{-a/2} \frac{\boldsymbol{i}\,dx'}{r} + \int_{-a/2}^{a/2} \frac{\boldsymbol{k}\,dz'}{r}\right] \qquad [8\text{-}62]$$

where the first integral represents the contribution of the portion of the closed current between C and D, and the three subsequent integrals represent the contributions of the three following sides.[7] The representative current element is located at the point $(x',0,z')$.

We shall confine our attention to cases where $r_0 \gg a$. Under these conditions

---

[6] This example is adapted from one given in Peck, *Electricity and Magnetism*, McGraw-Hill, 1953.

[7] In general, the dl (in equation [8-61]) is given by $\boldsymbol{i}\,dx' + \boldsymbol{j}\,dy' + \boldsymbol{k}\,dz'$. This reduces in the first and third integral to $\boldsymbol{i}\,dx'$ and in the second and fourth to $\boldsymbol{k}\,dz'$.

$r_0 \gg x'$ and $r_0 \gg z'$, and the value of $1/r$ at the point $(x',0,z')$ is given by

$$\frac{1}{r} = \frac{1}{r_0} + \left[\frac{\partial\left(\frac{1}{r}\right)}{\partial x'}\right]_{\substack{x'=0\\z'=0}} x' + \left[\frac{\partial\left(\frac{1}{r}\right)}{\partial z'}\right]_{\substack{x'=0\\z'=0}} z' + \text{higher-order terms} \qquad [8\text{-}63]$$

The subscripts following the partial derivative terms indicate that the partial derivatives are evaluated at the origin.

To evaluate the second term on the right-hand side of equation [8-63] we note that

$$r = \sqrt{(x - x')^2 + y^2 + (z - z')^2} \qquad [8\text{-}64]$$

and

$$\left[\frac{\partial\left(\frac{1}{r}\right)}{\partial x'}\right]_{\substack{x'=0\\z'=0}} = \left[\frac{(x - x')}{[(x - x')^2 + y^2 + (z - z')^2]^{3/2}}\right]_{\substack{x'=0\\z'=0}} \qquad [8\text{-}65]$$

$$= \frac{x}{(x^2 + y^2 + z^2)^{3/2}} = \frac{x}{r_0^3} \qquad [8\text{-}66]$$

Similarly, the partial derivative in the third term on the right-hand side of [8-63] has the value z/r_0^3. Upon substitution of these values, equation [8-63] becomes

$$\frac{1}{r} = \frac{1}{r_0} + \frac{x}{r_0^3} x' + \frac{z}{r_0^3} z' + \text{higher-order terms} \qquad [8\text{-}67]$$

The higher-order terms are negligible by comparison to the ones shown.

When equation [8-67] is substituted in equation [8-62], we have

$$A \doteq \frac{\mu_0 I}{4\pi}\left[i\int_{-a/2}^{a/2}\left[\frac{1}{r_0} + \frac{xx'}{r_0^3} + \frac{z\left(\frac{a}{2}\right)}{r_0^3}\right]dx' + k\int_{a/2}^{-a/2}\left[\frac{1}{r_0} + \frac{x\left(\frac{a}{2}\right)}{r_0^3} + \frac{zz'}{r_0^3}\right]dz'\right.$$

$$\left. + i\int_{a/2}^{-a/2}\left[\frac{1}{r_0} + \frac{xx'}{r_0^3} + \frac{z\left(-\frac{a}{2}\right)}{r_0^3}\right]dx' + k\int_{-a/2}^{a/2}\left[\frac{1}{r_0} + \frac{x\left(-\frac{a}{2}\right)}{r_0^3} + \frac{zz'}{r_0^3}\right]dz'\right] \qquad [8\text{-}68]$$

which, when evaluated, gives

$$A \doteq \frac{\mu_0 I}{4\pi}\frac{a^2}{r_0^3}[iz - kx] \qquad [8\text{-}69]$$

The induction can now be derived from equation [8-69] in accordance with

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} \qquad [8\text{-}59]$$

From the definition of the curl (equation [2-15]) we see that the components of $\boldsymbol{B}$ are given by

$$B_x = \left(\frac{\partial Az}{\partial y} - \frac{\partial Ay}{\partial z}\right), \quad B_y = \left(\frac{\partial Ax}{\partial z} - \frac{\partial Az}{\partial x}\right), \quad B_z = \left(\frac{\partial Ay}{\partial x} - \frac{\partial Ax}{\partial y}\right) \qquad [8\text{-}70]$$

Thus the x-component of B is obtained as follows:

$$B_x \doteq \frac{\mu_0 I a^2}{4\pi}\left[\frac{\partial}{\partial y}\left(\frac{-x}{(x^2+y^2+z^2)^{3/2}}\right) - \frac{\partial}{\partial z}(0)\right] \qquad [8\text{-}71]$$

or

$$B_x \doteq \frac{3\mu_0 I a^2 xy}{4\pi(x^2+y^2+z^2)^{5/2}} = \frac{3\mu_0 I a^2 xy}{4\pi r_0^5} \qquad [8\text{-}72]$$

Similarly, we obtain

$$B_y \doteq \frac{\mu_0 I a^2}{4\pi}\left[\frac{2}{r_0^3} - \frac{3(x^2+z^2)}{r_0^5}\right] \qquad [8\text{-}73]$$

and

$$B_z \doteq \frac{3\mu_0 I a^2 zy}{4\pi r_0^5} \qquad [8\text{-}74]$$

A useful selection of other examples in which B is derived from a previously calculated vector potential can be found in W. T. Scott, *The Physics of Electricity and Magnetism*, Wiley, 1959.

8.12 Further Discussion of the Magnetic Force Exerted upon an Electric Current or a Moving Charge

Having come to an end of the digression (§ 8.4 to § 8.9 plus § 8.11) that was made to discuss means of calculating the magnetic induction B, let us now return to the subject of the magnetic force exerted upon electric currents and charges in motion.

The Direction of the Magnetic Force upon a Current or a Charge in Motion. Two Alternative Rules

The direction of the magnetic force experienced by a current (or charge in motion) is best deduced from equation [8-3]:

$$dF = I\,dl \times B$$

or equation [8-9]:

$$dF = qv \times B$$

which provide at the same time the means of calculating its magnitude. There are, however, a couple of other useful methods of visualizing the direction of this force which should be briefly mentioned.

The first of these is based upon the picture of lines of force, developed so realistically by Faraday. An important property of lines of force, it will be recalled, is that of mutual lateral repulsion. With this in mind, let us look at the situation portrayed in Figure 8.14.

When the conductor (which extends perpendicular to the paper) is carrying no current, as on the left of Figure 8.14, the lines of force of the uniform mag-

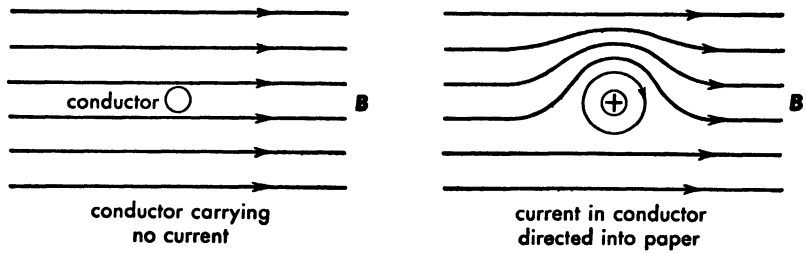


Fig. 8.14

netic field in which the conductor is placed are undisturbed. The situation is altered, however, when a current exists in the conductor, say, into the paper[8] (as on the right of Fig. 8.14), thus establishing its own clockwise system of coaxial lines of force. Now there is an interaction between the magnetic field of the current and the external field, with the result that the lines of force of the external field are distorted in the manner shown. The high density of lines of force above the conductor leads to a strong mutual repulsion, with the effect that the conductor experiences a downward force. A positive charge moving into the paper would similarly experience a downward force in the presence of the depicted magnetic field.

The direction of the force that is exerted upon a moving charge or current by a magnetic field is also given by *Fleming's left-hand rule*, or the *motor rule*. If the thumb, forefinger, and center finger of the left hand are extended so as to be mutually orthogonal, the center finger pointing in the direction of the current (or direction of motion of a positive charge), the forefinger pointing in the direction of the magnetic field ($\boldsymbol{B}$), then the thumb will point in the direction of the thrust experienced by the current or charge. If a negative charge is involved, the center finger should be pointed in the direction *opposite* to that in which the charge is moving. This rule is obviously intended for use in situations in which $\boldsymbol{B}$ and the current direction are at right angles. If the direction of the current is not perpendicular to $\boldsymbol{B}$, the center finger should be pointed in that direction, normal to $\boldsymbol{B}$, in which the current direction has its largest component.

Force on a Current in a Uniform Magnetic Field

In many cases of practical interest, the current-carrying conductor is located in a field of uniform induction, $\boldsymbol{B}$. The force experienced by an element of the conductor is then seen from equation [8-3] to be

$$d\boldsymbol{F} = I\, d\boldsymbol{l} \times \boldsymbol{B} \qquad [8\text{-}75]$$

[8] Current or change motion in a direction perpendicular to the plane of the paper is indicated by the symbols ⊙ and ⊕. The former represents the point of an arrow emerging from the paper, while the latter is a rear view of an arrow entering the paper.

If the wire is straight, the force per unit length is

$$\boldsymbol{F} = I\hat{\boldsymbol{l}} \times \boldsymbol{B} \qquad [8\text{-}76]$$

where $\hat{\boldsymbol{l}}$ is a unit vector pointing in the direction of current flow.

Frequently the conductor is normal to $\boldsymbol{B}$. In this case a conductor of length l experiences a magnetic force of magnitude

$$F = BIl \qquad [8\text{-}77]$$

and direction $\boldsymbol{l} \times \boldsymbol{B}$. The electric motor operates by virtue of such a force. Also, as we shall see in Chapter 11, the d'Arsonval galvanometer and the commonest types of DC ammeter and voltmeter are motivated by a force of this type.

Force Between Parallel Currents

Parallel currents provide an important special case of the force between current-carrying conductors.

The general expression (equation [8-1]) for the force, $d\boldsymbol{F}_{12}$, exerted by one current-carrying element, $d\boldsymbol{l}_1$, upon a second, $d\boldsymbol{l}_2$, takes the following form in the mks system:

$$d\boldsymbol{F}_{12} = \frac{\mu}{4\pi} \frac{I_1 I_2}{r_{12}^3} d\boldsymbol{l}_2 \times (d\boldsymbol{l}_1 \times \boldsymbol{r}_{12}) \qquad [8\text{-}78]$$

The situation described by [8-78] is depicted, for the general case of parallel currents, in Figure 8.15.

Let us apply equation [8-78] to the parallel currents shown in the figure. By doing so one can readily convince himself that the force, $\boldsymbol{F}_{12}$, exerted by $d\boldsymbol{l}_1$ upon $d\boldsymbol{l}_2$, is in the direction shown ($d\boldsymbol{l}_1 \times \boldsymbol{r}_{12}$ gives a vector, say $\boldsymbol{A}$, into the paper; then $d\boldsymbol{l}_2 \times \boldsymbol{A}$ gives a vector having the direction of $\boldsymbol{F}_{12}$). Likewise, the force, $\boldsymbol{F}_{21}$, exerted by $d\boldsymbol{l}_2$ upon $d\boldsymbol{l}_1$, has the direction shown in the figure. Thus, the two current elements—and, for that matter, the two parallel currents—are pulled together. If the currents are antiparallel, the force between them is a repulsive one.

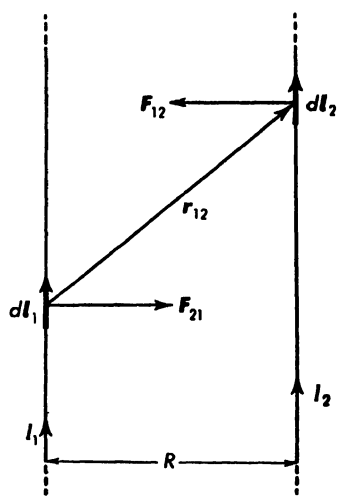


Fig. 8.15

When the first wire is long, the induction produced by it at a distance R from its axis is given by equation [8-38]:

$$B_1 = \frac{\mu I_1}{2\pi R} \qquad [8\text{-}38]$$

We can then picture the second conductor as located in this field of induction which, at the position of the second wire, has a direction into the paper (i.e., normal to dl_2). The force experienced by dl_2, given by equation [8-3], is

$$dF = I_2\, dl_2 \times B_1 \qquad [8\text{-}79]$$

or

$$dF = I_2\, dl_2 \left(\frac{\mu I_1}{2\pi R}\right) \qquad [8\text{-}80]$$

Thus, the magnetic force exerted on unit length of the second wire is

$$F = \frac{\mu I_1 I_2}{2\pi R} \text{ newtons/m} \qquad [8\text{-}81]$$

Meters of the electrodynamometer type (§ 11.6) utilize the force between parallel currents.

Pinch Effect

The mutual attraction of parallel currents exhibits itself in an electric discharge as an attraction between parallel-moving groups of ions. Here, of course, the ions are free to move in response to the attractive force arising from the interaction of their magnetic fields. With sufficiently high current densities this effect leads to a sudden constriction of the discharge, known as the *pinch effect*. Some of the attempts to obtain thermonuclear power have been concerned with provoking a sudden enough pinch to cause the nuclei of the "pinching" ions to fuse.

Forces Between Other Currents

One can calculate the force between any two current-carriers by integrating equation [8-78] over the length of both conductors.

8.13 Definition of the Coulomb and the Ampere

The meter, kilogram, second, and coulomb define the mks system. The meter and kilogram repose near Paris, the second is 1/86,400 of a mean solar day, but the coulomb must be realized experimentally. Since 1 ampere corresponds to a charge flow of 1 coulomb per second, the coulomb is realized through the ampere.

The ampere is defined in terms of the force that exists between current-carrying conductors. For example, we see from equation [8-81] that two long parallel wires, 1 m apart, each carrying a current of 1 ampere, experience in a vacuum attractive (or repulsive) forces of magnitude:

$$\text{Force} = \frac{\mu_0}{2\pi} = \frac{4\pi \times 10^{-7}}{2\pi} = 2 \times 10^{-7} \text{ newtons/m} \qquad [8\text{-}82]$$

In actual standardization work a pair of *coils* is so arranged that the force between them, when they carry a current, can be measured with precision. Such an arrangement is known as a *current balance.*[9] The current whose absolute amperage has thus been ascertained is also present in a standard resistor which is on series with the current balance. The resulting voltage drop is compared, in a potentiometer, with the emf of a standard cell. The standard cell whose emf has been determined in this manner is then used as the standard of reference. The emf of a standard cell is much more permanent than the calibration of any ammeter that might be put in series with the current balance.

One ampere is that current which, flowing in the same direction in each of two identical coaxial loops, produces in a vacuum a mutual attraction of

$$\frac{6\pi^2 a^4}{r^4} \times 10^{-7} \text{ newtons} \qquad [8\text{-}83]$$

where a and r are the radii and axial separation, respectively, of the two coils, and $r \gg a$.

8.14 The Mass Spectroscope

The mass spectroscope, an instrument used for the study of isotopes, provides an example of the force that is exerted upon a moving charge as it passes through a magnetic field. The name of the instrument arises from analogy with the ordinary spectroscope, in which it is possible to separate a beam of light into its various wavelength (or color) components. Similarly, the mass spectroscope separates a beam of charged particles into its several mass components. In the simplest form of the mass spectroscope, shown in Figure 8.16, the mass analysis takes place in a uniform magnetic field.

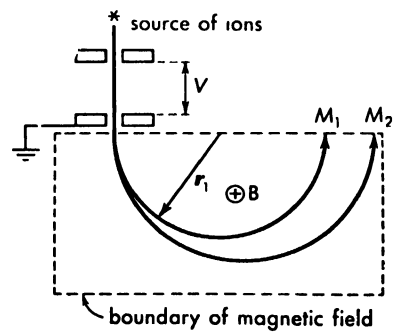


Fig. 8.16

[9] The current balance in use at the National Physical Laboratory in England is described in Vigoureux and Webb, *Principles of Electric and Magnetic Measurements*, Prentice-Hall, 1936.

The atoms to be analyzed are ionized in the source, each acquiring a net positive charge q, after which they fall through the potential difference V, thereby gaining kinetic energy,

$$\tfrac{1}{2}Mv^2 = Vq \qquad [8\text{-}84]$$

or momentum,

$$Mv = \sqrt{2VqM} \qquad [8\text{-}85]$$

Before and after this acceleration, which takes place between two parallel plates, the ions pass through narrow slits, as shown, which collimate the ion beam. Upon emerging from the second slit the ions enter a uniform magnetic field (of induction $\boldsymbol{B}$) which is directed into the paper.

Here each ion experiences a magnetic force, at right angles to its direction of motion, given by equation [8-9]:

$$\boldsymbol{F} = q\boldsymbol{v} \times \boldsymbol{B} \qquad [8\text{-}9]$$

Because this force is at right angles to the direction of motion of the ion, the magnitude of the ion velocity is not affected—only its direction is changed. Also, since $\boldsymbol{v}$ and $\boldsymbol{B}$ are normal to one another, the equation [8-9] reduces to

$$\boldsymbol{F} = -qvB\left(\frac{\boldsymbol{r}}{r}\right) \qquad [8\text{-}86]$$

That is, the force is centripetal (antiparallel to the radius vector $\boldsymbol{r}$) with the constant magnitude

$$F = qvB \qquad [8\text{-}87]$$

Under the action of such a force the ion traverses a circular orbit, experiencing a centripetal acceleration:

$$a = \frac{v^2}{r} \qquad [8\text{-}88]$$

The equation of motion of the ion is given by Newton's second law:

$$F = Ma \qquad [8\text{-}89]$$

which becomes, in this case,

$$qvB = \frac{Mv^2}{r} \qquad [8\text{-}90]$$

from which the momentum of the ion is seen to be

$$Mv = qBr \qquad [8\text{-}91]$$

Now, equating equations [8-85] and [8-91], we obtain

$$\frac{mv^2}{r} = evB = eV \qquad r^2 = \frac{2V}{B^2}\left(\frac{M}{q}\right) \qquad [8\text{-}92]$$

relating the specific charge (q/M) of the ion to the radius of curvature (r) of its path, the accelerating potential (V), and the magnetic induction (B). The charge, q, carried by the ion is usually one elementary charge (1.6×10^{-19} coulombs).

As equation [8-92] predicts, the heavier the ion the larger the radius of curvature of the path it traces in the magnetic field. In this way mass dispersion is achieved.

There are two types of mass spectroscope: (a) the mass spectrograph, in which the ions, after dispersion, strike a photographic plate on which the mass spectrum is recorded, and (b) the mass spectrometer, in which the separated ions are collected and their relative numbers determined by electrical means. These two methods for the detection of the separated ion beams are illustrated in Figure 8.17.

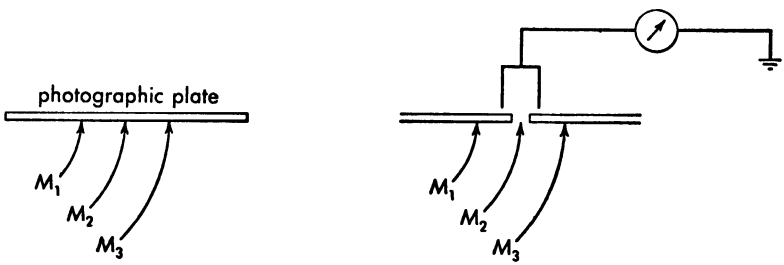


Fig. 8.17

In the mass spectrometer, the various ion groups are brought in turn to the collector by varying either the magnetic field or the accelerating voltage. In this way, the isotopic constitution of an element can be determined. Mass spectroscopes, usually more sophisticated than the type illustrated in Figure 8.16, are also used to determine atomic masses.

8.15 The Cyclotron

The *cyclotron* is a particle accelerator in which the ions are accelerated many times by the same potential. In this way very high particle energies are achieved with modest accelerating voltages. The cyclotron was invented by E. O. Lawrence (1901–1958), American physicist, who completed the first experimental model in 1931.

The ions to be accelerated are introduced into the gap between the two halves of a flat, circular, copper box that has been cut along a diameter. These two halves, known as the "Dees," are connected to opposite terminals of an rf (radio frequency) oscillator (left side of Fig. 8.18), and the entire Dee-

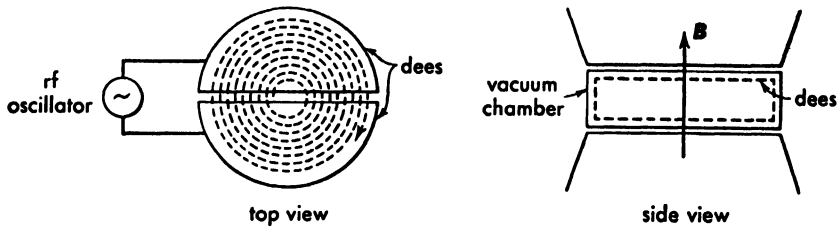


Fig. 8.18

structure is contained in a vacuum chamber which is located between the poles of a constant-field electromagnet [10] (right side of Fig. 8.18).

The acceleration begins when the ion is accelerated, under the action of the rf field, across the gap between the Dees. Once the ion is inside one of the Dees (an electric-field-free region), it follows a semicircular path, in accordance with equation [8-91], of radius

$$r = \frac{Mv}{Bq} \qquad [8\text{-}93]$$

If the time required for the ion to traverse this semicircular path is one-half the period of the rf voltage, the ion, on next reaching the gap, is again accelerated toward the opposite Dee, and within it traverses a semicircular path of larger radius. Thus the ion follows a roughly spiral course, gaining energy twice in each revolution, and eventually achieving a maximum kinetic energy of amount

$$\tfrac{1}{2}Mv_m^2 = \frac{q^2B^2r_m^2}{2M} \qquad [8\text{-}94]$$

where r_m is the largest permitted radius. This expression, which is not relativistic, is got from equation [8-93].

It will be realized from what has been said that the frequency, f, of the rf voltage must be related to the strength of the magnetic field and the specific charge of the ion that is accelerated. It is left to be shown in Problem 8.15 that these quantities are related as follows:

$$f = \frac{B}{2\pi}\left(\frac{q}{M}\right) \qquad [8\text{-}95]$$

Equation [8-95], which is also nonrelativistic, gives the frequency of the rf oscillator and, of course, gives the frequency of the motion of ions in the magnetic field as well. This latter frequency, known as the *cyclotron frequency*, is independent of the radius of curvature; otherwise the operation of the cyclotron would not be possible.

[10] The production of magnetic fields by means of electromagnets is discussed in Chapter 10.

It can be seen, from equation [8-94], that for a given ion the radius of curvature at any stage during the acceleration is characteristic of the ion energy. Thus the most energetic ions are found in the orbits of largest radius of curvature. Indeed, when the ion energy becomes too great, the ions can no longer be accommodated within the Dees ($r > r_m$). These maximum-energy ions may emerge through an aperture in one of the Dees or they may be ejected by some magnetic or electrostatic device.

The energy spread in the ion beam from the cyclotron is much wider than in the case of the Van de Graaff generator discussed in § 2.9. The ion current, however, is two to five times as great and the attainable energies are significantly higher.

The Synchrocyclotron (Frequency Modulated Cyclotron)

Let us rewrite equation [8-95] in terms of angular velocity:

$$\omega = B\frac{q}{M} \qquad [8\text{-}96]$$

The condition for successful cyclotron operation can then be restated as follows: the angular frequency, ω, of the rf voltage must be identical with the angular frequency, $B(q/M)$, of the motion of the ions.

The resonant condition is destroyed when the ion is accelerated to relativistic energies. Because of the relativistic increase in the ion mass (equation [2-51]), the angular velocity of the ion motion is reduced, with the result that the ion arrives at the gap too late to receive the maximum acceleration. It is still further out of step with the accelerating voltage when it crosses the next gap, and eventually traverses the gap in such a phase as to be *de*celerated. Thus, there is a relativistic limit to the energy which any particular ion can acquire in a conventional cyclotron.

Means of overcoming this limitation were suggested independently in 1945 by V. Veksler in Russia and E. M. McMillan in the U.S.A. The effect of the increasing mass of the circulating particle can be compensated for by a regular increase in B, so that B/M (in equation [8-96]) remains constant, or by gradually decreasing the angular frequency, ω, of the rf oscillator. In practice, it is usually the frequency which is altered. In the University of Chicago synchrocyclotron, when protons are being accelerated, the rf frequency is reduced from 28.6 to 18.0 Mc in 2 μsec. During acceleration the protons travel approximately 350 miles, finally traveling in an orbit of radius 76 in. and possessing an energy of 450 MeV. Protons of this energy are very useful for meson production. In the standard cyclotron the energy limit is 10–20 MeV.

Because this accelerator follows a cycle during which the rf frequency is

changed by the necessary amount, the ions are emitted in bursts, usually sixty per second.

8.16 The Hall Effect

It was observed in 1879, by Edwin Hall of Harvard University, that when a current-carrying conductor is placed in a transverse magnetic field, as in Figure 8.19, the top of the conductor becomes positive with respect to the bottom. This effect was for decades of little interest, but has recently become important in the study of solids, where it gives the number and sign of the current-carriers. It is explicable in terms of the principles which we have discussed.

Let us consider the case investigated by Hall—that of a metallic conductor —where the current-carriers are electrons. We shall assume the *average* drift velocity of the conduction electrons to be $\boldsymbol{v} = j v_y$. The induction, as shown, is $\boldsymbol{B} = -iB_x$. The conduction electrons are therefore subject to a force:

$$\boldsymbol{F} = -e\boldsymbol{v} \times \boldsymbol{B} = (-e)(jv_y) \times (-iB_x) \qquad [8\text{-}97]$$

or

$$\boldsymbol{F} = -ev_y B_x \boldsymbol{k} \qquad [8\text{-}98]$$

which forces them to the bottom surface. This nonuniform electron distribution gives rise to an emf between the top and bottom surfaces, known as the *Hall emf*, and written as $\mathcal{E}_H$. This corresponds to an electric field,

$$\boldsymbol{E}_H = -\boldsymbol{k}E_H \left(= -k\frac{\mathcal{E}_H}{d} \right) \qquad [8\text{-}99]$$

When equilibrium has been established, the conduction electrons are subject to zero force. The electrical force arising from the Hall emf is exactly balanced by the magnetic force, that is,

$$\boldsymbol{E}_H(-e) + (-e)\boldsymbol{v} \times \boldsymbol{B} = 0 \qquad [8\text{-}100]$$

or

$$\boldsymbol{E}_H e\boldsymbol{k} - ev_y B_x \boldsymbol{k} = 0$$

or

$$E_H = v_y B_x \qquad [8\text{-}101]$$

We can now make use of the expression derived in § 5.1 for the current density, viz.,

$$\boldsymbol{J} = nq\boldsymbol{v} \qquad [5\text{-}4]$$

In this case,

$$\boldsymbol{J} = iJ_x + jJ_y + kJ_z = n(-e)jv_y \qquad [8\text{-}102]$$

whence,

$$J_y = -nev_y \qquad [8\text{-}103]$$

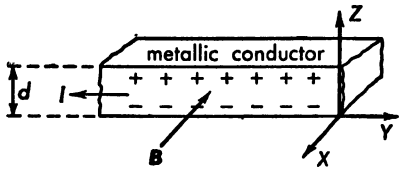


Fig. 8.19

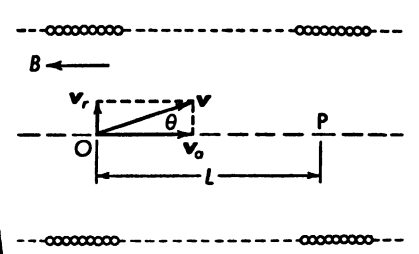


Fig. 8.20

When equation [8-103] is substituted in [8-101] we obtain

$$\frac{E_H}{J_y B_x} = -\frac{1}{ne} \qquad [8\text{-}104]$$

The *Hall coefficient*, R_H, is defined as

$$R_H = \frac{E_H}{J_y B_x} = -\frac{1}{ne} \qquad [8\text{-}105]$$

R_H is numerically equal to the electric field produced across a longitudinal conductor that is located in a transverse magnetic field of unit induction when the current density in the conductor is unity. The Hall coefficient contains only quantities that can be determined experimentally. From it, as can be seen in equation [8-105], the number of conduction electrons per unit volume can be computed. In this way it has been shown, for example, that in silver there are ~1.2 conduction electrons per atom.

When the current consists of conduction electrons the Hall coefficient is, of course, always negative. This is the *normal* Hall effect. In some metals, however, as well as in semiconductors, the Hall coefficient is positive. This *anomalous* Hall effect is an indication that the current is effectively carried by positive charges. These are the *holes* mentioned in § 6.3. The Hall effect can thus be used to ascertain both the number and sign of the current-carriers.

8.17 Focusing of Charged Particles by Magnetic Lenses

Long Solenoid

The simplest type of magnetic lens is a long solenoid, shown in Figure 8.20. Let us consider a charged particle, say, an electron, emerging from the point O on the axis of the solenoid with velocity v, which makes an angle θ with the axis. Thus, the axial and radial components of v are $v_a = v \cos\theta$ and $v_r = v \sin\theta$, respectively.

Let us write the total velocity of the electron as the vector sum of these two components:

$$\boldsymbol{v} = \boldsymbol{v}_a + \boldsymbol{v}_r \qquad [8\text{-}106]$$

Since $\boldsymbol{B}$ of the solenoid is also axial, it has no effect upon v_a. v_r, however, is normal to $\boldsymbol{B}$, with the result that the electron experiences a constant force

$$\boldsymbol{F} = -e\boldsymbol{v}_r \times \boldsymbol{B} \qquad [8\text{-}107]$$

of magnitude

$$F = ev_rB \qquad [8\text{-}108]$$

which is perpendicular to both $\boldsymbol{v}_r$ and $\boldsymbol{B}$. This force, at the instant shown in Figure 8.20, is directed into the paper.[11] Under the action of this force the electron executes a circular motion while traveling down the solenoid with axial velocity v_a.

The frequency of this circular motion is simply the *cyclotron frequency* which we have already referred to in equation [8-95]. The period of the motion is

$$T = \frac{1}{f} = \frac{2\pi m}{Be} \qquad [8\text{-}109]$$

The axial distance traveled during one cycle of this cyclotron motion is

$$L = Tv_a = \left(\frac{2\pi m}{Be}\right) v \cos\theta \qquad [8\text{-}110]$$

Thus, after traveling an axial distance L from the point O, the electron again returns to the axis. In short, the electron follows a helical path in its passage down the solenoid, crossing the axis at distances L, $2L$, etc., from its point of origin.

Any other electron that emerges from O with the same axial velocity, $v_a = v\cos\theta$, also returns to the axis after the same distance L, because the cyclotron period [8-109] is independent of the magnitude of the radial component of the velocity. In this way an "image," or at least a *recombination* occurs at the point P. If the electrons under consideration all possess the same total velocity, and θ is small enough to make $\cos\theta \doteq 1$, a very good focus occurs at P.

The Short Magnetic Lens

A short magnetic lens differs in its focusing action from a solenoidal field in two fundamental respects: (a) both object and image are normally outside

[11] The radial velocity, v_r, does not have a constant direction. Rather, under the action of the force described by equation [8-107], its direction undergoes a steady rotation about an axis parallel to $\boldsymbol{B}$.

the field, and (b) the deflection of the electrons is primarily due to the radial rather than to the axial component of the magnetic field. The approximate shape of the magnetic field for a short solenoid is shown in Figure 8.21.

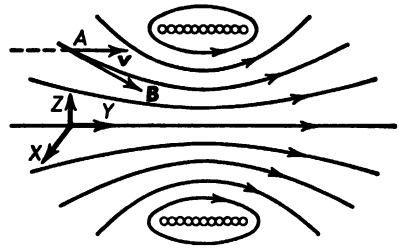


Fig. 8.21

Let us consider the force at point A on an electron traveling in the Y-Z plane with a velocity $\boldsymbol{v} = j v_y$. The value of $\boldsymbol{B}$ at point A is $\boldsymbol{B} = jB_y + kB_z$; B_z is evidently negative. The force on the electron is given, as usual, by

$$\boldsymbol{F} = -e\boldsymbol{v} \times \boldsymbol{B}$$

$$= -e(jv_y) \times (jB_y + kB_z) \qquad [8\text{-}111]$$

$$= -iev_yB_z \qquad [8\text{-}112]$$

Since B_z is a negative quantity, this force is in the positive x-direction, or out from the paper, giving rise to a motion about the axis. This motion, at the point in question, corresponds to a component of velocity in the x-direction, say, iv_x, which is likewise acted upon by the magnetic flux. The force so exerted is

$$\boldsymbol{F}' = -e(iv_x) \times (jB_y + kB_z) \qquad [8\text{-}113]$$

$$= -kev_xB_y + jev_xB_z \qquad [8\text{-}114]$$

the first term of which is directed toward the axis. In this way electrons approaching the lens in a direction parallel to the axis but displaced radially from it are deflected toward the axis. This focusing action of a short current-carrying coil was known prior to 1900 but was not understood until the problem was analyzed by Busch in 1926.

To summarize: the *radial* component of the induction is responsible for a lateral component of velocity which, under the action of the *axial* component of the induction, gives rise to a component of velocity toward the axis. Magnetic lenses embodying this principle are integral components of most electron microscopes and many cathode ray tubes. They are also used to study the energetic electrons which are emitted (as β-rays) from certain radioactive materials.

8.18 Problems

E (8.1) Knowing that the erg is the unit of energy in both the cgsm and cgse systems of units, deduce the relationship between (a) the statvolt and the abvolt and (b) the statohm and the abohm.

E (8.2) Find the point along the axis of a current-carrying circular loop at which the rate of change of field is constant. [This result forms the basis for the Helmholtz coils arrangement (§ 8.6).]

E (8.3) What is the magnetic induction at the center of an equilateral triangle 100 cm on a side formed by a wire that is carrying a current of 10 amperes?

M (8.4) A current of 100 amperes is carried by a long wire of radius 0.01 m. Calculate the total flux ($\Phi_m = \int_s \boldsymbol{B}\cdot d\boldsymbol{s}$), per unit length of wire, that is enclosed between the surface of the wire and a coaxial cylinder of radius 0.02718 m (recall e = 2.718). [Ans.: $\Phi_m = 50\mu/\pi$ webers.]

E (8.5) A circuit in the form of a regular polygon of n sides is inscribed in a circle of radius a. Calculate the value of the magnetic induction at the center of the circle if the current in the circuit is I amperes.

[Ans.: $B = \dfrac{\mu n I}{2\pi a} \tan(\pi/n)$ webers/m^2.]

M (8.6) Show that, as n is indefinitely increased, the result of the previous problem approaches the value of B at the center of a circular current.

M (8.7) A circuit, carrying a current I amperes, is in the form of a square of side length l meters. Find the value of B at a point on the axis of the square at a distance p from its center.

$$\left[\text{Ans.: } B = \frac{\mu I l^2}{2\pi\left(p^2 + \dfrac{l^2}{4}\right)\sqrt{p^2 + \dfrac{l^2}{2}}} \text{ webers/m}^2.\right]$$

E (8.8) An electric and a magnetic field are so arranged with respect to one another that ions with a particular velocity v_0 are not deflected in passing through them. How is this done? (This combination of fields is a well-known one and is called a *Wien velocity filter*.)

M (8.9) A certain region is pervaded by a uniform electric field of value $\boldsymbol{E} = 2\boldsymbol{i} - 3\boldsymbol{j}$ volts/m and by a uniform magnetic induction of value $\boldsymbol{B} = 4\boldsymbol{i} + 3\boldsymbol{j} + 3\boldsymbol{k}$ webers/m^2. What is the total force exerted upon a moving charge q coulombs at the instant that it possesses the velocity $\boldsymbol{v} = \boldsymbol{i} + 2\boldsymbol{j}$ m/sec? [Ans.: $q(8\boldsymbol{i} - 6\boldsymbol{j} - 5\boldsymbol{k})$ newtons.]

E (8.10) Electrons traveling with a velocity of 10^5 m/sec pass through a grid and enter an electric-field-free region in which the magnetic induction is just sufficient to prevent the electrons from reaching a

plate parallel to the grid and 1 cm from it. What is the direction and magnitude of the magnetic induction?

H (8.11) Electrons emitted from a surface A (Fig. 8.22) are accelerated

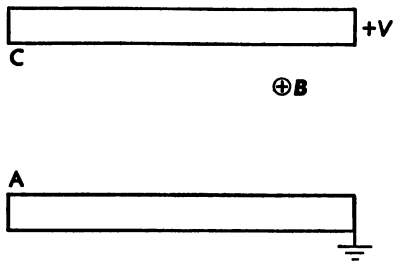


Fig. 8.22

toward surface C which is maintained at a potential of $+V$ volts. At what value of the flux density B will current cease to flow to plate C? The separation of the plates is d m.

[Ans.: $B = \sqrt{2mV/ed^2}$.]

M (8.12) An electron is traveling at such a speed that its mass is 1.25 times its rest mass. What will be the radius of curvature of its path in a perpendicular magnetic field with $B = 0.1$ webers/m^2?

[Ans.: 0.013 m.]

E (8.13) Two straight wires carrying equal DC currents in opposite directions are so arranged that one lies on top of the other. If the wires each have diameters of 5 mm and linear densities of 1 gm/cm, what current will cause the top wire to be just lifted from the bottom one?

[Ans.: 157 amperes.]

M (8.14) Suppose a magnet with semicircular pole pieces, 1.22 m in diameter, is used for the mass spectroscopic separation of U^{235} from U^{238}. If an accelerating voltage of 30,000 volts is employed, what magnetic flux density is required? [These are the conditions under which the isotope separator (*calutron*) at Oak Ridge operates.]

[Ans.: 0.63 webers/m^2.]

E (8.15) Prove that the frequency of the rf voltage that is applied to the Dees of a cyclotron is given by

$$f = \frac{qB}{2\pi M} \text{ cycles/sec}$$

where q/M is the specific charge of the ion being accelerated.

M (8.16) An oscillator applies a peak potential of 10,000 volts to the Dees of a cyclotron. The flux density provided by the cyclotron magnet is 2 webers/m^2. Calculate the radius of the orbit that a proton will describe after it has completed 100 complete revolutions.

[Ans.: 0.1 m.]

M (8.17) What is the focal length of a long solenoid?

Chapter 9

Electromagnetic Induction

Oersted's discovery of the magnetic field associated with an electric current suggested to several investigators the possibility that magnetism could in some way give rise to electricity. The method by which this could be done was discovered by Michael Faraday in 1831. The discovery was made independently, and perhaps earlier, by the American physicist, Joseph Henry (1797–1878), who later became first director of the Smithsonian Institution. Faraday's priority, which was admitted even by Henry, is based on his prior publication of his results.

9.1 Lenz's Law

The first experiment performed by Faraday, in which a current was induced by a magnetic field, employed the apparatus shown in Figure 9.1. The helices, H_1 and H_2, were wound side by side on a piece of wood. One of the helices was connected to a galvanometer and the other through a switch to a battery. It was found, on making and breaking the battery circuit, that the galvanometer indicated brief currents, in opposite directions in the two cases. In later experiments Faraday discovered that the mere approach or removal of a current-carrying conductor induced a current in a neighbouring closed circuit, and that the motion of magnets produced similar effects.

In 1834 Heinrich Friedrich Emil Lenz (1804–1865), a German physicist, made an important series of experiments which, among other results, led him

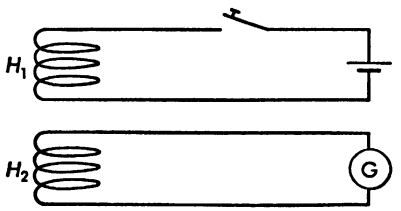


Fig. 9.1

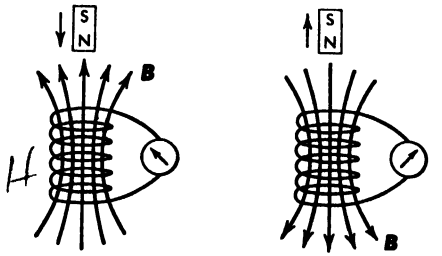


Fig. 9.2

to his celebrated rule for predicting the direction of an induced current. The rule is as follows: the *direction of the induced current is always such that its electromagnetic action tends to oppose the change which produces it.* Faraday had also shown how to predict the direction of an induced current but had not stated the principle as succinctly as did Lenz.

Lenz's law is illustrated in Figure 9.2. As the bar magnet approaches the top of the coil, as on the left, the current induced in the helix gives rise to an induction that opposes the approach of the magnet.[1] Likewise, as the magnet is withdrawn, as on the right in Figure 9.2, its motion is again opposed by the magnetic field arising from the current induced in the coil.

Although we have spoken of an induced "current," it will be realized that the primary effect is the inducing of an *emf*. This gives rise to a current, provided the circuit is complete.

9.2 Magnitude of the Induced emf: Faraday's Law of Electromagnetic Induction

Magnetic Flux

The surface integral of $\boldsymbol{B}$ is known as *magnetic flux*, and is designated by the symbol Φ_m. Thus

$$d\Phi_m = \boldsymbol{B}\cdot d\boldsymbol{s} \qquad [9\text{-}1]$$

and the total flux through any surface s is given by the integral of [9-1] over the surface:

$$\Phi_m = \int_s \boldsymbol{B}\cdot d\boldsymbol{s} \qquad [9\text{-}2]$$

Clearly, magnetic flux is a scalar.

---

[1] We shall see in the next chapter that the north pole of a magnet, when placed in an external magnetic field, experiences a force in the direction of the induction.

Faraday's Law of Electromagnetic Induction

Faraday's experiments led him to conclude that the magnitude of the emf induced in a closed loop is proportional to the time rate of change of magnetic flux through the loop.

The total flux, Φ_m, through the closed loop shown in Figure 9.3, is evidently

$$\Phi_m = \int_s \boldsymbol{B} \cdot d\boldsymbol{s} \qquad [9\text{-}2]$$

where s is any surface having the loop as boundary. Then, according to Faraday,

$$\varepsilon = -C\frac{d\Phi_m}{dt} \qquad [9\text{-}3]$$

where ε is the instantaneous value[9] of the induced emf, and C is a constant of proportionality. The negative sign in equation [9-3] is the basis for a sign convention governing ε. This is discussed below.

The mks unit of flux is the *weber* (named in honor of the German physicist, Wilhelm Eduard Weber [1804–1891]. When the flux is expressed in webers and the emf in volts, the constant of proportionality in equation [9-3] is equal to unity. Thus,

$$\varepsilon = -\frac{d\Phi_m}{dt} \qquad [9\text{-}4]$$

This emf is induced in a closed loop whether a conductor comprises the loop or not. It is present, for example, around any closed loop in *space*, through which the flux is changing.

If the flux passes through N identical turns of wire connected in series, instead of the single loop shown in Figure 9.3, the total induced emf is the sum of the emf's generated in the separate turns:

$$\varepsilon = -N\frac{d\Phi_m}{dt} \qquad [9\text{-}5]$$

This is Faraday's *law of electromagnetic induction.*

Sign Convention Governing ε

Let us note the convention of signs that is implied in Faraday's law. If the flux Φ_m is decreasing, the induced emf must, by Lenz's law, give rise to a current whose induction (given by the right-hand rule) is in the same direction

[9] We shall frequently deal, from now on, with emf's and currents that are fluctuating. In such cases (e.g., in AC circuits) it is necessary to consider instantaneous values. These will be designated by the symbols e and i, in contrast to the steady-state values, $\mathcal{E}$ and I.

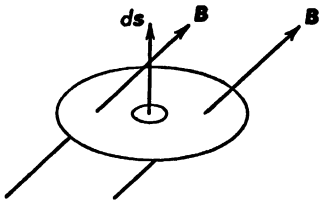


Fig. 9.3

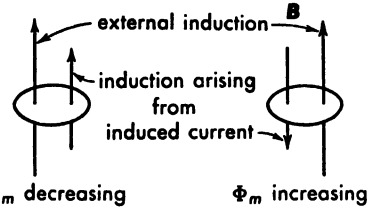


Fig. 9.4

as the external induction. This is shown on the left in Figure 9.4. Similarly, when $d\Phi_m/dt$ is positive, as on the right, Lenz's law requires the induced current to flow so as to give rise to an "induced" induction that is opposite in direction to the external induction. In the former case (from equation [9-4]), ε is positive; in the latter case it is negative. A positive emf is, therefore, one in which the resulting current gives rise to an induction that is either parallel to the external induction or has a component that is parallel to it. For a negative emf, the induced induction, or a component of it, is antiparallel to the external induction.

9.3 Faraday's Law Derived from Ampère's Law

Although Faraday arrived at the law of electromagnetic induction independently and empirically, it is a logical consequence of Ampère's law and the *conservation of energy*, as we shall now demonstrate.

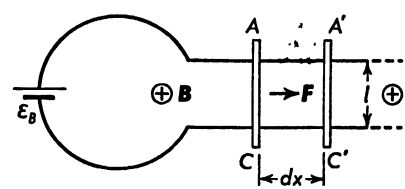


Fig. 9.5

Figure 9.5 represents a circuit which is completed at the right end by a shorting bar *AC*. This bar can be slid on parallel rails. The circuit is located in a region of uniform magnetic induction. For simplicity, the induction is assumed to be perpendicular to the paper and pointing into it.

The current in the circuit is I amperes, causing an energy expenditure during the time dt of

$$\mathcal{E}_B I\,dt \qquad [9\text{-}6]^9$$

When *AC* is in a fixed position, this energy is all spent in overcoming the resistance in the circuit, viz.,

$$\mathcal{E}_B I\,dt = I^2 R\,dt \qquad [9\text{-}7]$$

[9] The emf of the battery, $\mathcal{E}_B$, is the work done on unit charge by the battery (§ 5.8). In time dt the charge transported through the battery is $(dq/dt)\,dt = I\,dt$. Thus, the total work done by the emf (and, hence, the electrical energy expended in the circuit) in time dt is given by equation [9-6].

Meanwhile, the shorting bar is subject to a force F, in the direction shown in Figure 9.5, of magnitude:

$$F = BIl \qquad [8\text{-}77]$$

Let us now suppose that the shorting bar slides at a steady rate during time dt to a new position $A'C'$, a distance dx from its original location. During this motion the system does work equal to

$$dW = Bil\,dx \qquad [9\text{-}8]^4$$

The energy expended by the battery in time dt is now employed not only to overcome the resistance in the circuit but also to do this work, dW. Thus

$$\mathcal{E}_B i\,dt = i^2 R\,dt + Bil\,dx \qquad [9\text{-}9]$$

or

$$i = \frac{\mathcal{E}_B - \dfrac{Bl\,dx}{dt}}{R} \qquad [9\text{-}10]$$

Evidently, during this motion, there is an emf in the circuit in addition to the battery emf, of value

$$\mathcal{E} = -\frac{Bl\,dx}{dt} \qquad [9\text{-}11]$$

Meanwhile, it is seen from Figure 9.5 that $l\,dx$ is the amount by which the area enclosed by the circuit has been increased. Hence,

$$Bl\,dx = d\Phi_m \qquad [9\text{-}12]$$

is the change in the total flux passing through the circuit. In short, the emf arising from the change of flux through the circuit is given by

$$\mathcal{E} = -\frac{d\Phi_m}{dt} \qquad [9\text{-}4]$$

in agreement with Faraday's law. If the shorting bar were moved, by some external agent, from $A'C'$ to AC, the work done by the system would be negative, and the induced emf would have the same direction as the battery emf.

In deriving Faraday's law from Ampère's law, we have considered the magnetic field to be stationary and the conductor to be in motion. The combination of fixed conductor and varying magnetic field, discussed in § 9.2, is entirely equivalent.

[4] The value of the current, i, with the shorting bar in motion differs, as is seen below, from the current value when the shorting bar is at rest.

9.4 Alternative Deduction of the Direction of an Induced Current in a Conductor Moving in a Magnetic Field

We have already acknowledged that Lenz's law provides a general principle for deducing the direction of an induced current. Another useful expression of the same idea can be deduced from Ampère's law.[5]

Let us begin with the experiment depicted in Figure 9.6, which shows a conductor moving to the left with velocity v, in a region of uniform normal induction. This conductor consists of nuclei of atoms, electrons that are bound to the nuclei, and free electrons. The free electrons alone are able to move under the action of an electric field. Essentially, then, Figure 9.6 represents a group of electrons moving to the left with velocity v.

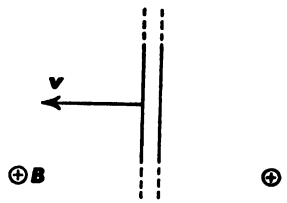


Fig. 9.6

The force exerted upon each of these electrons by the magnetic field is given by equation [8-9]:

$$\boldsymbol{F} = -e\boldsymbol{v} \times \boldsymbol{B} \qquad [8\text{-}9]$$

By the definition of electric field,

$$\boldsymbol{F} = -e\boldsymbol{E} \qquad [9\text{-}13]$$

where $\boldsymbol{E}$ (force per unit charge) is the electric field associated with the induced emf. This field is, evidently,

$$\boldsymbol{E} = \boldsymbol{v} \times \boldsymbol{B} \qquad [9\text{-}14]$$

This field gives rise to a current provided, of course, that the conductor forms part of a complete circuit. The direction of current, in the orthogonal configuration of Figure 9.6, is *down*. If the situation is not an orthogonal one, the current in the conductor flows in the direction in which $\boldsymbol{E}$ has a component. The total induced emf is the integral of $\boldsymbol{E} \cdot d\boldsymbol{l}$ over the length of the conductor.

Equation [9-14], giving the direction of the induced field (and, therefore, the direction of the induced current), is a much more convenient rule to keep in mind than the rather cumbersome generator rule described in footnote 5.

---

[5] The direction of an induced current can also be predicted from *Fleming's generator rule.* This rule corresponds to the *left-hand* or *motor rule* referred to in § 8.12. If the thumb and first two fingers of the *right hand* are extended so as to point in mutually perpendicular directions, the thumb pointing in the direction of motion of the conductor and the forefinger pointing in the direction of the magnetic field (B), then the center finger points in the direction of the induced current. This rule, like the left-hand rule, is obviously intended for situations in which the conductor extends in a direction that is normal to the induction, and the motion to which it is subjected is in a direction that is normal to both.

9.5 Maxwell's Third Equation

We can derive from Faraday's law of electromagnetic induction the third of Maxwell's field equations.

The emf induced in a closed loop is given by Faraday's law:

$$\mathcal{E} = -\frac{d\Phi_m}{dt} = -\int_s \frac{\partial \boldsymbol{B}}{\partial t} \cdot d\boldsymbol{s} \qquad [9\text{-}2]\ [9\text{-}4]$$

This emf can also be found by computing the work done in carrying a unit charge completely around the loop. Thus

$$\mathcal{E} = \oint \boldsymbol{E} \cdot d\boldsymbol{l} \qquad [9\text{-}15]$$

where $\boldsymbol{E}$ is the intensity of the field associated with the induced emf. Attention is drawn to the fact that the line integral of this electric field around a closed path does not vanish as it does for a *static* field (equation [2-26]).

From the two above equations we obtain

$$\oint \boldsymbol{E} \cdot d\boldsymbol{l} = -\int_s \frac{\partial \boldsymbol{B}}{\partial t} \cdot d\boldsymbol{s} \qquad [9\text{-}16]$$

the third of Maxwell's equations (in integral form).

By means of Stokes' theorem (§ 8.10 and Appendix), equation [9-16] can be rewritten as

$$\int_s (\nabla \times \boldsymbol{E}) \cdot d\boldsymbol{s} = -\int_s \frac{\partial \boldsymbol{B}}{\partial t} \cdot d\boldsymbol{s} \qquad [9\text{-}17]$$

from which the differential form of Maxwell's second equation is obtained:

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \qquad [9\text{-}18]$$

This equation, relating the rate of change of magnetic induction to the resulting electric field, will be further discussed in Chapter 14 when Maxwell's four equations will be considered as a group.

9.6 Inductively Coupled Circuits: Mutual Inductance and Self-inductance

Consider the two coils shown in Figure 9.7, which are so located with respect to one another that the passage of current in one of them (called the

primary coil) gives rise to a magnetic flux which extends, at least in part, to the region occupied by the other (called the *secondary* coil). When the current changes in the primary coil, emf's are induced in both coils. We shall discuss that which is induced in the secondary coil first.

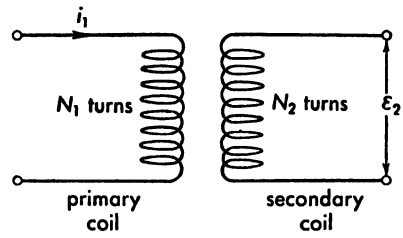


Fig. 9.7

Mutual Inductance

Let us suppose that the number of turns of wire in the secondary coils is N_2 and that the flux from the primary coil which passes through it at any time is Φ_{m_2} webers. Then the emf induced in the secondary coil is given by Faraday's law:

$$\mathcal{E}_2 = -N_2 \frac{d\Phi_{m_2}}{dt} \qquad [9\text{-}5]$$

Since Φ_{m_2} is produced by the current i_1 flowing in the primary coil, we may write:

$$N_2\Phi_{m_2} = M_{21}i_1 \qquad [9\text{-}19]$$

Substituting [9-19] in [9-5], we obtain

$$\mathcal{E}_2 = -M_{21} \frac{di_1}{dt} \qquad [9\text{-}20]$$

The constant M_{21} is known as the *mutual inductance.*

Likewise, if the current changes in the second coil, an emf

$$\mathcal{E}_1 = -M_{12} \frac{di_2}{dt} \qquad [9\text{-}21]$$

is induced in the first coil. Although it is not self-evident in the general case, it will be proved in § 9.10 that

$$M_{12} = M_{21} \qquad [9\text{-}22]^6$$

That is, the mutual inductance is a property of the two circuits and has the same value regardless of which is regarded as the "primary" and which the "secondary." We shall therefore drop the subscripts, and simply write M. Thus equations [9-20] and [9-21] become

$$\mathcal{E}_2 = -M \frac{di_1}{dt} \quad \text{and} \quad \mathcal{E}_1 = -M \frac{di_2}{dt} \qquad [9\text{-}23]$$

---

[6] Equation [9-22] is proved in Problem 9.10 for the very special case in which all the flux passing through one of two coils passes through the other as well.

In the mks system, mutual inductance is expressed in *henrys*. From the above equation we see that the *mutual inductance of a pair of circuits is 1 henry if a current changing in one of the circuits at the rate of 1 ampere per second induces in the other circuit an emf of 1 volt.*

The Definition of M

Equation [9-19] assumes that the primary flux passing through each turn of the secondary coil is the same, namely, Φ_{m_2}. If so, the emf induced in the secondary coil is correctly given by equation [9-5]. This assumption is valid for certain special combinations of coils (e.g., two coaxial toroidal solenoids), but it is far from valid for the pair of loosely wound coils shown in Figure 9.7. Equation [9-19], therefore, should not be regarded as the definition of M; M is defined by equation [9-20] or [9-23]. On the other hand, in circumstances in which Φ_{m_2} is calculable and common to all turns in the secondary coil, equation [9-19] may be used to compute M (cf. § 9.10).

Self-inductance

The emf induced in the primary coil as a result of a change in its own flux is given by an expression similar to equation [9-23]:

$$\varepsilon_1 = -L\frac{di_1}{dt} \qquad [9\text{-}24]$$

where the constant L is called the *self-inductance*, or, simply, the *inductance* of the primary coil. Self-inductance is also expressed in henrys. *A coil has an inductance of 1 henry if a current changing in it at the rate of 1 ampere per second gives rise to an induced emf of 1 volt.*

We should note that when the current in the coil is increasing (di_1/dt positive), the induced emf is negative, thereby opposing the change by which it is produced. For this reason the emf that is induced in a coil as a result of current changing in it is known as a *back emf* or a *counter emf*. Lenz's law is similarly obeyed when di_1/dt is negative. Because of this back emf, the rise or decay of current in a coil does not take place instantaneously (cf. § 12.5).

For certain special coil geometries, the flux through each turn is the same, say, Φ_{m_1}. Then the back emf, given by equation [9-5], is

$$\varepsilon_1 = -\frac{d}{dt}(N_1\Phi_{m_1}) \qquad [9\text{-}25]$$

When this is compared with equation [9-24], written for the occasion as

$$\varepsilon_1 = -\frac{d}{dt}(Li_1) \qquad [9\text{-}26]$$

we see that

$$N_1\Phi_{m_1} = Li_1 \qquad [9\text{-}27]$$

This relation is useful in computing the inductances of coils for which the assumption that the flux passing through each turn is the same is valid (cf. § 9.10).

9.7 Inductances in Series

We shall now derive the expressions for the effective inductance and effective resistance of a series combination of two inductors. The arrangement is illustrated in Figure 9.8, where M is the mutual inductance of the two coils whose self-inductances are L_1 and L_2. The coil resistances are R_1 and R_2. At every instant the current is changing at the same rate, di/dt, at every point in the circuit. The back emf across the first coil is

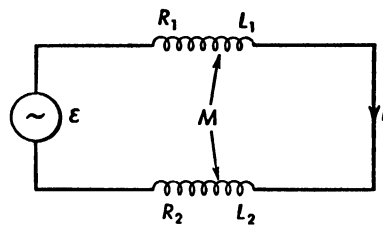


Fig. 9.8

$$\mathcal{E}_1 = -L_1\frac{di}{dt} \mp M\frac{di}{dt} \qquad [9\text{-}28]$$

where the second term on the right-hand side results from the changing current in the second coil. The sign of this term depends upon the direction in which the second coil is wound relative to the first. The back emf across the second coil is given by a similar expression:

$$\mathcal{E}_2 = -L_2\frac{di}{dt} \mp M\frac{di}{dt} \qquad [9\text{-}29]$$

We may write Kirchhoff's second law for the entire circuit:

$$\mathcal{E} + \mathcal{E}_1 + \mathcal{E}_2 = i(R_1 + R_2) \qquad [9\text{-}30]$$

or

$$\mathcal{E} - \frac{di}{dt}[L_1 + L_2 \pm 2M] = i(R_1 + R_2) \qquad [9\text{-}31]$$

where $\mathcal{E}$, the instantaneous emf of the generator, is the external emf that is applied to the circuit. In general, when an emf $\mathcal{E}$ is impressed across a series circuit containing effective inductance L and effective resistance R, Kirchhoff's second law gives

$$\mathcal{E} - L\frac{di}{dt} = iR \qquad [9\text{-}32]$$

Comparing equations [9-31] and [9-32] we see that the effective inductance of

the two coils in series is

$$L = L_1 + L_2 \pm 2M \qquad [9\text{-}33]$$

and their effective resistance is

$$R = R_1 + R_2 \qquad [9\text{-}34]$$

The last term in [9-33] is negligible if the two coils are far apart.

Equation [9-33] serves as the basis for a widely used type of variable inductance. Two flat, circular coils are mounted as shown in Figure 9.9. Coil *A* is fixed in position while coil *C* is free to rotate about their common diameter. The mutual inductance of the two coils can thus be continuously varied from maximum positive to maximum negative coupling. It is zero when the planes of the coils are mutually perpendicular.

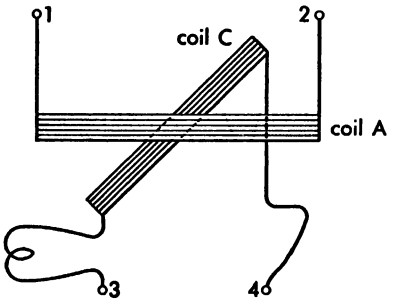


Fig. 9.9

When this combination is employed as a variable mutual inductance, the primary terminals are *1* and *2*, while the secondary terminals are *3* and *4*. When the unit is used as a self-inductance, terminals *2* and *3* are joined, usually by a shorting bar provided for the purpose, and terminals *1* and *4* are the binding posts for the so-formed variable inductor or *inductometer*. Under these conditions L can be varied between the limits given by equation [9-33].

The calculation of the effective inductance and resistance of inductors in parallel is postponed until Chapter 13, where methods to be developed for the handling of alternating currents will simplify what would otherwise be a laborious chore.

9.8 Energy Stored in a Coil Through Which a Current Is Flowing

We have seen that work is done, in the amount of $\frac{1}{2}CV^2$ joules, in establishing a potential difference across a capacitor. Similarly, work is done in establishing current in an inductor. In this case the rise of current in the coil is opposed by the back emf

$$\varepsilon_B = -L\frac{di}{dt} \qquad [9\text{-}35]$$

Consequently, in establishing the current, the impressed emf does work in dt seconds, of amount

$$dW = -\varepsilon_R i\,dt = Li\,di \qquad [9\text{-}36]$$

The total work done during the time required for the current to rise to the value of i amperes is

$$W = L\int_0^i i\,di = \tfrac{1}{2}Li^2 \qquad [9\text{-}37]^7$$

This energy is possessed by the current-carrying coil on account of its self-inductance, and can be regarded as *magnetic energy*. As in the electrical case, it can be considered as occupying the space pervaded by the magnetic field of the coil (cf. § 9.9).

The energy possessed by a coil is relinquished by it, when the current is shut off, in the following way. When the flux emanating from the coil collapses, it induces in the coil an emf which does work on the conduction electrons. The total energy thus acquired by the conduction electrons is $\frac{1}{2}Li^2$ joules.

9.9 Energy in a Magnetic Field

It was shown in §§ 1.11 and 4.5 that the volume density of energy in a region pervaded by an electric field is

$$\frac{DE}{2} \qquad [4\text{-}27]$$

The analogous magnetic expression is

$$\frac{BH}{2} \qquad [9\text{-}39]$$

It will now be shown that the energy density of space pervaded by a magnetic field is in fact given by expression [9-39].

We note from the previous section that the energy of a coil carrying a steady current I is

$$\text{Energy} = \tfrac{1}{2}LI^2 \qquad [9\text{-}37]$$

This energy may be regarded as residing in the region pervaded by the magnetic field of the coil. In order to calculate the volume density of magnetic energy, let us consider the case of a toroidal solenoid, and make use of the conclusions relating thereto that we have already drawn in § 8.9. First, we recall that

---

[7] Similarly the work done in overcoming the back emf arising from the mutual inductance of two coils is

$$W = Mi_1i_2 \qquad [9\text{-}38]$$

the magnetic field of such a coil is entirely confined to the region enclosed by the coil. Second, when the cross-sectional radius of the coil is small by comparison to the toroidal radius, the value of induction within the coil is everywhere given by

$$B = \mu nI \qquad [8\text{-}43]$$

The inductance of such a coil is given by equation [9-27]:

$$L = \frac{N\Phi_m}{I} \qquad [9\text{-}27]$$

which may be rewritten

$$L = \frac{NBA}{I} \qquad [9\text{-}40]$$

where A is the interior cross-sectional area of the coil. Equation [9-37] giving the energy of the coil can then be expressed as

$$\text{Energy} = \tfrac{1}{2}LI^2 = \frac{NBAI}{2}$$

or

$$\text{Energy} = \frac{(nl)BAI}{2} = \frac{B(nI)(Al)}{2} \qquad [9\text{-}41]$$

Here the quantity Al (where l is the "length" of the solenoid) represents the volume to which the magnetic energy is confined. Substituting (from [8-11]) $H = B/\mu = nI$ in equation [9-41], we conclude that the magnetic energy density of the space occupied by the magnetic field of the toroidal solenoid is given by:

$$\text{Magnetic energy density} = \frac{BH(Al)}{2(Al)} = \frac{BH}{2} \qquad [9\text{-}42]$$

as predicted above. This result, it will be noted, does not involve any quantity that is peculiarly toroidal. Although it has been derived in the context of a specific type of coil, the specific characteristics of the coil do not appear in the final result. Equation [9-42] is of general validity.

We may conclude, on the basis of equations [4-27] and [9-42], that the *total* energy density in a region pervaded by both electric and magnetic fields is

$$\text{Electromagnetic energy density} = \frac{(DE + BH)}{2} \qquad [9\text{-}43]$$

This, incidentally, is the instantaneous energy density in any space traversed by electromagnetic radiation.

9.10 Calculation of Self-inductance and Mutual Inductance

The results of the preceding two sections can be used to calculate the inductances of conductor configurations whose magnetic fields, when the conductors are carrying current, are amenable to calculation. On the one hand we know that the magnetic energy is given by

$$\text{Magnetic energy} = \tfrac{1}{2}LI^2 \qquad [9\text{-}37]$$

On the other hand, we have the alternative expression

$$\text{Magnetic energy} = \int_v (\text{magnetic energy density})\, dv \qquad [9\text{-}44]$$

$$= \int_v \frac{BH}{2}\, dv \qquad [9\text{-}45]$$

where the integration is carried out over the volume occupied by the magnetic field of the conductor configuration under study. Thus, by equating [9-37] and [9-45], we see that

$$L = \frac{1}{I^2}\int_v BH\, dv \qquad [9\text{-}46]$$

gives the inductance value that is sought.

Self-inductance of a Long Solenoid

We have shown in § 8.7 that the induction along the axis of a long solenoid is given by

$$B \doteq \mu nI \qquad [8\text{-}32]$$

In addition, it follows from the argument presented in § 8.9 that the values of B at nonaxial, interior points are identical. Assuming the contribution to the energy integral of the region outside the solenoid is negligible (that is, the energy is confined to the volume Al) and neglecting end effects, we have from [9-46]

$$L \doteq \frac{1}{I^2}(\mu nI)(nI)(Al) \doteq \frac{\mu N^2 A}{l} \qquad [9\text{-}47]$$

where A is the cross-sectional area of the solenoid, and N and l are its total turns and length, respectively. For an air-filled solenoid, $\mu = \mu_0$; the value of μ for an iron core will be discussed in the next chapter.

Equation [9-47] also applies to a toroidal solenoid whose toroidal radius is large compared to its cross-sectional radius.

Self-inductance of a Coaxial Cable

The two commonest types of "line" for transmitting electromagnetic energy are the parallel-wire line and the coaxial cable. The inductance of the parallel-wire line is discussed in Problem 9.7: that of the coaxial cable will now be derived as an additional example of the use of equation [9-46].

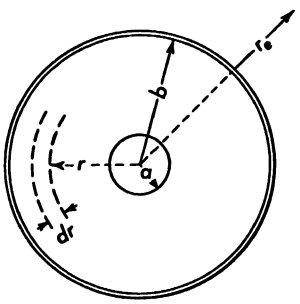


Fig. 9.10

The coaxial cable is shown in cross-section in Figure 9.10. The current has one direction in the central conductor and the opposite direction in the outer conductor (often called the "return" path). The wall of the outer conductor is assumed to be negligibly thin. As in the preceding example, we shall proceed to compute the total magnetic energy (per unit length) that is associated with charge flow in this pair of conductors. We know, of course, from symmetry, that the field lines are circles about the axis. It is assumed that in the space between the inner and outer conductor $\mu = \mu_0$.

First, we notice that the entire flux is confined to the region between the conductors. This we can see from the fact that the value of

$$\oint \boldsymbol{B} \cdot d\boldsymbol{l}$$

integrated around a circular path (in the plane of the paper) of radius r_e is zero, inasmuch as the two currents that are circumnavigated are equal and opposite (cf. § 8.9). Further, the outer cylinder contributes zero flux to the internal region. This is seen by imagining the inner conductor to be removed and evaluating the integral

$$\oint \boldsymbol{B} \cdot d\boldsymbol{l}$$

around a circular path of radius r. The integral is zero since zero current is enclosed by such a closed loop. Hence, all flux is confined to the region within the outer conductor and is generated by the current in the inner conductor.

For unit length of conductor the value of $BH\,dv$ for the radial range r to $r + dr$ is given by

$$BH\,dv = \left(\frac{\mu_0 I}{2\pi r}\right)\left(\frac{I}{2\pi r}\right) 2\pi r\,dr = \frac{\mu_0 I^2}{2\pi}\frac{dr}{r} \qquad [9\text{-}48]$$

where B and H are given by equation [8-38] for a long straight wire. When

equation [9-46] is evaluated for the between-conductor region, we have

$$L = \frac{1}{I^2}\int_a^b BH\,dv = \frac{1}{I^2}\left(\frac{\mu_0 I^2}{2\pi}\right)\int_a^b \frac{dr}{r} = \frac{\mu_0}{2\pi}\ln\frac{b}{a} \qquad [9\text{-}49]$$

There is another $BH\,dv$ contribution which we have not taken into account, namely,

$$\int_0^a BH\,dv$$

or that contained *within* the inner conductor. This can be shown (Problem 9.7) to increase the value of L by an amount $\mu/8\pi$ when the current is uniformly distributed across the cross-section of the inner conductor (as at low frequencies). Thus, the total inductance for the coaxial cable per unit length is

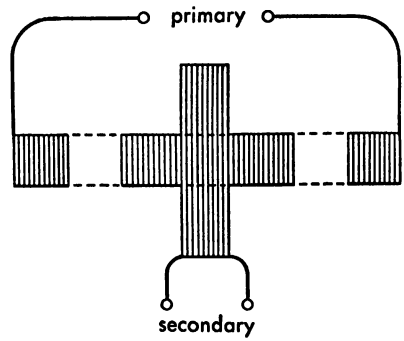


Fig. 9.11

$$L = \frac{1}{8\pi}\left[4\mu_0 \ln\frac{b}{a} + \mu\right] \qquad [9\text{-}50]$$

For a nonmagnetic central conductor, and $b \gg a$, the second term is negligible.

Calculation of Self-inductance by Flux Linkage

When dealing with a conductor configuration for which the flux passing through each turn is the same, we can readily calculate L from equation [9-27]:

$$N\Phi_m = LI \qquad [9\text{-}27]$$

Thus, for the toroidal solenoid whose axial radius is large in comparison to its cross-section radius,

$$\Phi_m = BA = \mu nIA \qquad [9\text{-}51]$$

and

$$L = \frac{N\Phi_m}{I} = \frac{N}{I}\left(\mu\frac{N}{l}IA\right) = \frac{\mu N^2 A}{l} \qquad [9\text{-}52]$$

in agreement with equation [9-47].

Calculation of Mutual Inductance by Flux Linkage

Calculations corresponding to the preceding can also be made for mutual inductance. Let us consider the case shown in Figure 9.11. The "primary"

coil is a long solenoid while the "secondary" coil is short and has a total of N_2 turns.

From equation [9-19], which holds in this case, we have

$$M = \frac{N_2 \Phi_{m_2}}{I_1} \qquad [9\text{-}19]$$

where Φ_{m_2}, the flux through each turn of the secondary coil, is the same as the flux, Φ_{m_1}, in the interior of the primary coil, since the flux density outside the primary coil is negligibly small. Thus

$$M = \frac{N_2 \Phi_{m_1}}{I_1} = \frac{N_2 (B_1 A_1)}{I_1} = \frac{N_2 (\mu n_1 I_1 A_1)}{I_1} = \frac{\mu N_1 N_2 A_1}{l_1} \qquad [9\text{-}53]$$

This expression is independent of the diameter of the secondary coil, provided it exceeds that of the primary.

■ *Neumann's Formula for the Calculation of Mutual Inductance*

Consider the two circuits shown in Figure 9.12. When a current i_1 flows in circuit

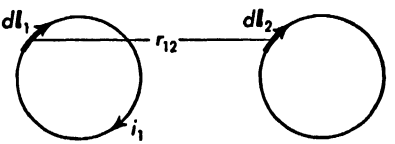


Fig. 9.12

1 it generates a field of induction B_1 which extends in part to the region of circuit *2*. When i_1 changes, an emf ε_2 is induced in circuit *2* of amount

$$\varepsilon_2 = -\int_{s_2} \frac{\partial}{\partial t} (\boldsymbol{B}_1 \cdot d\boldsymbol{s}_2) = -\frac{\partial}{\partial t} \int_{s_2} \boldsymbol{B}_1 \cdot d\boldsymbol{s}_2 \qquad [9\text{-}54]$$

$\boldsymbol{B}_1$, as we have seen in § 8.11, can be derived from a vector potential:

$$\boldsymbol{B}_1 = \nabla \times \boldsymbol{A}_1 \qquad [8\text{-}59]$$

Equation [9-54] may then be rewritten

$$\varepsilon_2 = -\frac{\partial}{\partial t} \int_{s_2} (\nabla \times \boldsymbol{A}_1) \cdot d\boldsymbol{s}_2 \qquad [9\text{-}55]$$

and transformed, using Stokes' theorem, to

$$\varepsilon_2 = -\frac{\partial}{\partial t} \oint \boldsymbol{A}_1 \cdot d\boldsymbol{l}_2 \qquad [9\text{-}56]$$

Finally, if we substitute the value of A_1 from equation [8-60], we obtain

$$\mathcal{E}_2 = -\frac{\partial}{\partial t}\oint\left[\frac{\mu i_1}{4\pi}\oint\frac{dl_1}{r_{12}}\right]\cdot dl_2 \qquad [9\text{-}57]$$

or

$$\mathcal{E}_2 = -\left[\frac{\mu}{4\pi}\oint\oint\frac{dl_1\cdot dl_2}{r_{12}}\right]\frac{\partial i_1}{\partial t} \qquad [9\text{-}58]$$

where dl_1 is to be integrated around the boundary of circuit *1* and dl_2 around the boundary of circuit *2*.

When we compare this expression with equation [9-23], we see that the mutual inductance between circuit *1* and circuit *2* is

$$M = \frac{\mu}{4\pi}\oint\oint\frac{dl_1\cdot dl_2}{r_{12}} \qquad [9\text{-}59]$$

which depends only upon the geometry of the two circuits and the permeability of the intervening medium. Equation [9-59], known as *Neumann's formula*, substantiates the assertion made earlier that mutual inductance is a joint property of two inductively coupled circuits and does not depend upon which circuit is the originator of the induction process.

Neumann's formula provides a direct and useful method of calculating mutual inductance, although we shall not give any examples of its use.

In principle a formula similar to [9-59] should hold also for the calculation of the self-inductance of a circuit. In this formula both elements of length would lie in the same circuit. The following difficulty, however, is encountered: Neumann's formula assumes that the currents in the conductors may be regarded as thin current filaments. This is usually justified in calculations of M_1, but when a double integral of the form of [9-59] is taken around a single filamentary path, as would be required to calculate L, each coincidence of the two elements of length gives a value of infinity (since $r = 0$). This difficulty can be circumvented by taking into account the finite cross section of the conducting paths and the current distribution across them.

Standard Inductances

The examples given above of self- and mutual-inductance calculations were deliberately chosen for their simplicity. There are, however, many other coil configurations for which the self- or mutual-inductance values can be computed exactly. It is thus possible, by constructing coils to certain fixed specifications, to obtain inductors whose absolute values of inductance are known. The interested reader should consult F. W. Grover, *Inductance Calculations*, Van Nostrand, 1946.

Unknown inductances are usually determined by comparing them with standard inductances by means of an AC bridge (§ 13.19).

Coefficient of Coupling

The efficiency with which one circuit is inductively coupled to another is described by the term *coefficient of coupling*, K, defined by the equation

$$K = \frac{M}{\sqrt{L_1 L_2}} \qquad [9\text{-}60]$$

It can be shown (Problem 9.10) that, for two circuits in which the entire flux of the one passes through the other (e.g., two coaxial, toroidal solenoids of the same dimensions),

$$M = \sqrt{L_1 L_2} \qquad [9\text{-}61]$$

a relation which underlines the fact that mutual inductance is a property of the two coils. It is immaterial which of the two is regarded as the primary and which the secondary.

In practice there is usually flux leakage between any two coupled circuits: the flux from the one does not pass completely through the other, or vice versa. Equation [9-60] describes this situation. For *close* coupling, $K \rightarrow 1$; for *loose* coupling, $K \rightarrow 0$.

9.11 Eddy Currents

Eddy currents arise in a conductor through which the magnetic flux is changing. The heat produced by these currents usually represents an unintentional energy loss, although in one case to be discussed the loss is deliberate. These currents are, of course, in accord with Faraday's law of electromagnetic induction. They flow in closed loops in planes that are normal to the direction of the magnetic induction, $\boldsymbol{B}$.

Transformer

The action of a transformer will be studied in Chapter 13, but the effect of eddy currents upon transformer design will be briefly discussed here. Figure 9.13 shows the essential elements of a transformer, the iron serving the dual purpose of enhancing the flux (cf. Chapter 10) and of leading it from the primary to the secondary coil.

The transformer is essentially an AC device, the changing current in the primary coil giving rise to a changing flux in the iron which in turn gives rise to an emf across the secondary coil. Typical eddy currents arising from this $d\Phi_m/dt$ are shown in the upper portion of Figure 9.14, which is a cross-sectional

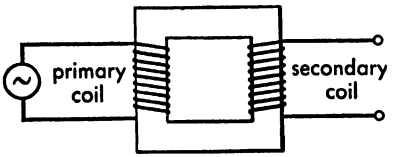


Fig. 9.13

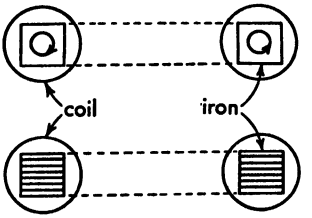


Fig. 9.14

view of the transformer. In practice these currents are frustrated, and the energy loss associated with them avoided, by constructing the transformer core of sheets or *laminations* of iron, each of which is insulated from its neighbors, as shown in the lower half of Figure 9.14. The interlaminar insulation interrupts the eddy currents that would otherwise flow.

Induction Heating

This is a case in which the eddy current loss is deliberately fostered. A metal to be heated to a high temperature is placed inside a coil that is carrying a radio-frequency current (Figure 9.15). The temperature to which the metal is heated depends upon the strength of the rf current, its frequency, and the resistance which the metal offers to the resulting eddy currents (Problem 9.11).

Since glass offers no significant impediment to the passage of magnetic flux, the metal to be heated in the induction heater can be housed, if desired, in a glass envelope that is being evacuated. In this way electrodes in vacuum tubes, for example, can be *outgassed* before the tube is sealed off.

Electromagnetic Shielding

Let us suppose that a conducting sheet lies in a plane normal to a magnetic field, as indicated in Figure 9.16. If $\boldsymbol{B}$ suddenly changes, eddy currents are set up in the sheet, whose directions, by Lenz's law, are such as to compensate for the change in $\boldsymbol{B}$. As a result, at point A, the full change in $\boldsymbol{B}$ is not immedi-

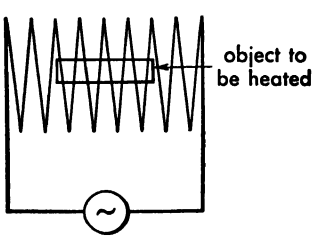


Fig. 9.15

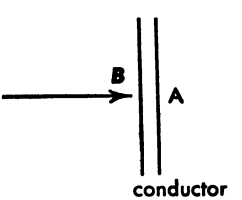


Fig. 9.16

ately felt. Indeed, if the resistance of the sheet were zero, it would never be felt. In actual conductors the eddy currents quickly die out, because of the resistance that is present, and $\boldsymbol{B}$ soon assumes, in region A, its new value.

In the meantime, however, region A has been partially shielded from the change in flux. Furthermore, if the conductor be one of low resistance, such as copper, and the changes in $\boldsymbol{B}$ are taking place periodically at a high frequency, this shielding can be practically complete. For this reason, copper is in common use as a shield against the magnetic flux from rf coils. It is of no use as a shield from static magnetic fluxes: for this purpose magnetic materials are needed (cf. Chapter 10).

9.12 The Betatron—an Induction Accelerator

The *betatron* is an electron accelerator in which the accelerating field is produced by a changing magnetic flux. The electrons travel in circular orbits in an evacuated *doughnut* which is located between the poles of an electromagnet,

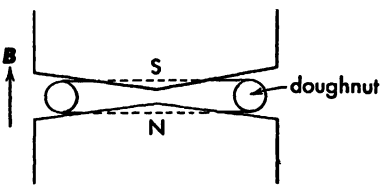


Fig. 9.17

as shown in Figure 9.17. The radius of each orbit is given (cf. equation [8-91]) by

$$r = \frac{mv}{Be} \qquad [8\text{-}91]$$

which may be written as

$$p = Ber \qquad [9\text{-}62]$$

where p $(= mv)$ is the momentum of the electron.

When the magnetic flux, Φ_m, enclosed by the orbit is increased, the resulting emf, ε, given by Faraday's law, represents a tangential electric field of magnitude

$$E = \frac{\varepsilon}{2\pi r} = \frac{1}{2\pi r}\frac{d\Phi_m}{dt} \qquad [9\text{-}63]$$

The electrons gain momentum under the action of this field in accordance with Newton's second law of motion

$$\frac{dp}{dt} = \text{Force} = Ee = \frac{e}{2\pi r}\frac{d\Phi_m}{dt} \qquad [9\text{-}64]$$

If the electron starts from rest when the magnetic field is zero, then its momentum at the time the flux has attained the value Φ_m is

$$p = \frac{e}{2\pi r}\Phi_m \qquad [9\text{-}65]$$

Eliminating p between equations [9-62] and [9-65], we obtain

$$B = \frac{\Phi_m}{2\pi r^2} = \frac{\Phi_m}{2A} \qquad [9\text{-}66]$$

where A is the area enclosed by the orbit. Therefore, if the magnetic field can be shaped in such a way that the value of the induction at an orbit of radius r is always *half* its average value (Φ_m/A) within this orbit, the electron, during acceleration, will follow an orbit of this constant radius. The pole pieces of the magnet, as shown in Figure 9.17, are designed to provide a strong central flux and a weaker one at the orbit in order to fulfil requirement [9-66].

The first usable beam of electrons to be accelerated in this manner was obtained in 1940 by D. W. Kerst at the University of Illinois. The idea had been suggested earlier by R. Wideröe (Germany, 1928) and E. T. S. Walton (England, 1929). In Kerst's first betatron the electrons acquired 2.3 MeV energy. Most present-day betatrons are in the 25 MeV range, although a small number accelerate electrons to energies $\geq$100 MeV. In a 100 MeV instrument the electron makes 250,000 turns, traveling a distance of 900 miles and gaining 400 eV per turn. When the magnetic flux reaches its peak value, corresponding to the maximum electron energy, current is sent through an auxiliary winding on the magnet. This disturbs the orbit to the extent that the beam strikes a target for the production of x-rays or emerges as an external electron beam.

The ultimate energy achievable by this method is limited by the radiation loss of the electron. Whenever a charged particle is accelerated, it radiates electromagnetic waves. In this case, the traversing of the circular orbit constitutes the principal acceleration. Eventually, therefore, the electron loses energy by radiation at the same rate as it gains it by induction acceleration. There is no relativistic limitation as in the cyclotron. The electron output, like that of the synchrocyclotron, is pulsed at the frequency of the magnet excitation.

9.13 Problems

E (9.1) A wire running perpendicular to the page is moved in a uniform magnetic field, as indicated in Figure 9.18. The wire is part of a complete circuit. Use Lenz's law and the type of argument employed in Figure 8.14 *et seq.* to find the direction of the induced current.

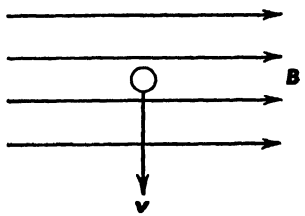


Fig. 9.18

E (9.2) A wire of length 2 m is perpendicular to the X-Y plane. It is moved with velocity $\boldsymbol{v} = 2\boldsymbol{i} + 3\boldsymbol{j} + \boldsymbol{k}$ m/sec through a region of uniform induction $\boldsymbol{B} = \boldsymbol{i} + 2\boldsymbol{j}$ webers/m^2. By the method of § 9.4 compute the potential difference between the ends of the wire. [Ans.: 2 volts.]

E (9.3) A metal airplane propeller rotates at 20,000 rpm. What potential difference exists between the hub and the tip of the propeller when the airplane is flying horizontally in a direction due magnetic north? The propeller is 3 m in total length. Take the horizontal component of the earth's magnetic field to be 10^{-5} webers/m^2.

[Ans.: 2.4×10^{-2} volts.]

E (9.4) A U-shaped conductor lies in the Y-Z plane, as shown in Figure 9.19, in a region of uniform induction $B = \boldsymbol{i}B$. What is the potential difference between the terminals A and C when the loop is moved with velocity (a) $\boldsymbol{v} = \boldsymbol{i}v$, (b) $\boldsymbol{v} = \boldsymbol{j}v$, (c) $\boldsymbol{v} = \boldsymbol{k}v$? Throughout each motion the orientation of the conductor remains constant.

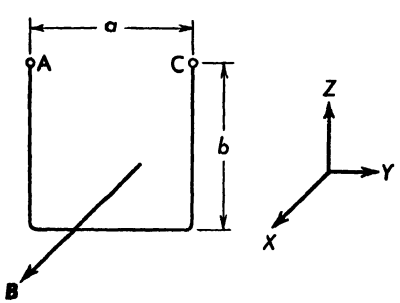


Fig. 9.19

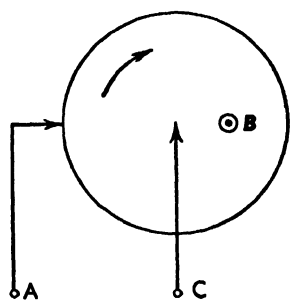


Fig. 9.20

E (9.5) Figure 9.20 shows the *Faraday disc generator*. A copper disc of radius a is mounted with its axis parallel to a uniform field of induction B. If the disc rotates at 120 rpm, find the emf appearing across the terminals A and C, which are connected to brushes, as shown, one making contact with the circumference of the disc and the other with its axis. [Ans.: $2\pi a^2 B$.]

M (9.6) This problem deals with the basic AC generator. A rectangular coil of N turns and area A is rotated with angular velocity ω in a field of uniform induction B. The direction of B is at right angles to the axis of rotation of the coil. Show that the emf induced in the coil is given by $\varepsilon = BAN\omega \sin \omega t$. Assume that at $t = 0$ the plane of the coil is normal to B.

H (9.7) A long current-carrying wire has uniform current density. Show that the magnetic energy per unit length *within* the wire is equal to $\mu I^2/16\pi$, where I is the total current carried by the wire.

M (9.8) Calculate the self-inductance per meter of two parallel telephone wires, each of radius a meters and spaced D meters apart ($D \gg a$). Since the two wires represent the "lead" and "return" wires of a single circuit, they carry equal currents in opposite directions.

M (9.9) Calculate the mutual inductance of the two coils shown in Figure 9.21, where $a \gg b$. The large coil contains N_1 turns and the small coil N_2 turns.

$$\left[\text{Ans.: } \frac{\mu N_1 N_2 \pi a^2 b^2}{2(a^2 + p^2)^{3/2}}.\right]$$

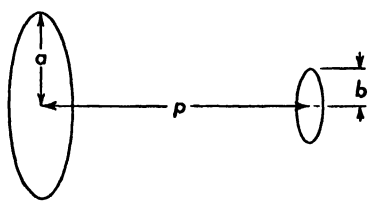


Fig. 9.21

E (9.10) Show that, for two coaxial toroidal solenoids of the same dimensions, $M = \sqrt{L_1 L_2}$.

E (9.11) Other factors being equal, will a piece of copper or a piece of nichrome (see Table 3, § 6.4) experience the greater heating effect in an induction heater?

Chapter 10

The Magnetic Properties of Matter

10.1 History of Magnetism

As with electrostatics, the early history of magnetism is shrouded in antiquity. The word *magnetism* is derived from the Greek *magnes*, applied by Euripides to an iron ore (*magnetite*) that displayed a selective attraction for iron. It appears, from the writings of Lucretius, that the Greeks and Romans were aware that magnetic iron ore could endue ordinary iron in contact with it with its own peculiar property. The existence of magnetic *poles* was apparently not known, although many other interesting facts concerning magnetism had been catalogued. We are told, for example, that a magnet is effective in the cure of disease; that it affects the brain, causing melancholy; that it acts as a love philtre; that it loses its power when rubbed with garlic, but recovers it when treated with goat's blood; and much else that has not yet been substantiated by modern experiments.

The science of magnetism made little progress until the development of the mariner's compass, possibly invented in China in the latter part of the eleventh century. William Gilbert, whose name has already been mentioned in connection with the development of electrostatics, collected in his *De Magnete* (1600) all that was then known concerning magnetism, including many observations of his own. He confirmed the fact that a freely suspended magnet not only points north and south but also dips towards the earth. In the northern hemisphere it is the north pole of the magnet that points downward, the *angle of dip* depending upon the latitude. From these experiments Gilbert inferred

that the earth itself acts as a huge magnet with poles almost, but not quite, coincident with its geographic poles.

The observation that like poles of a magnet repel each other had been made towards the end of the sixteenth century. In view of the widespread use of compasses prior to this time, it is surprising that the discovery had not been made earlier. Undoubtedly, a cause that contributed to the delay is the fact that a magnetized piece of iron, even if it is initially repelled when brought near a strong magnet, may be induced to reverse its polarity and so experience an attraction.

10.2 Force Between Magnetic Poles: Coulomb's Law for Magnets

The conclusion that there were two types of magnetism, analogous to the two types of static electricity, was a natural consequence of the discovery that the poles of a magnet either attract or repel one another. The north (i.e., north-seeking) pole was said to be *positive*, whereas the south pole was *negative*.

Attempts were made to separate the positive and negative poles of a magnet by breaking the magnet into two pieces, but it was always found that negative and positive poles, respectively, appeared on the two sides of the break. It was, therefore, reluctantly accepted that the privilege of independent existence, enjoyed by electric charges, is not a prerogative of magnetic poles. Whenever a positive pole is discovered in nature, it is inevitably associated with a negative pole of equal strength. (This fact is, of course, explicable on the modern view that the magnetic properties of a magnet are due to the elementary closed currents that it contains.)

The problem of obtaining "free" poles, in order to determine the force between them, was overcome by using long, thin magnets. Here, as shown in Figure 10.1, the positive pole was sufficiently far removed from its associated negative pole as to be practically isolated from it. The law of force between poles was investigated by Coulomb, using the same torsion balance with which he established the law of force between electric charges, and was found to be similar in form to that for charges. Thus, the magnitude of the force between two poles of strength q_{m_i} and q_{m_j}, located a distance r_{ij} apart, is written as

$$F_{ij} = C\frac{q_{m_i}q_{m_j}}{r_{ij}^2} \qquad [10\text{-}1]$$

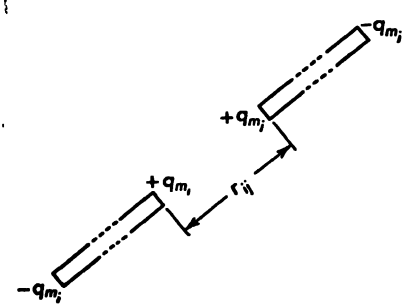


Fig. 10.1

where the value of C, the constant of proportionality, depends upon the system of units employed and the magnetic properties of the intervening medium. The vector form of Coulomb's law is similar to equation [1-2] for electric charges, inasmuch as the force between poles acts along the line joining them.

Equation [10-1] is the fundamental equation for magnetostatics. From it a unit pole can be defined. Then, by analogy to electrostatics, the ideas of magnetic field, magnetic potential, and so on, can be developed. The primary entities, in this magnetostatic picture, are small bar magnets possessing positive and negative poles, q_m and $-q_m$, which are separated physically from one another. This method of representing elementary magnets, although artificial, is useful for handling many aspects of magnetism. For example, it provides a convenient method of calculating the external effect of a large assembly of elementary closed currents, as in a compass or ordinary magnet. In addition, the small bar magnet is a valid replacement for a single closed current when one is calculating magnetic effects at distances which are large by comparison to the dimensions of the current loop.

Coulomb's Law in the mks System of Units

By definition, the mks unit pole, when placed at a point in space where the magnetic induction is 1 weber/m^2, experiences a force of 1 newton. Thus, the force experienced by a pole of strength q_{m_j} mks units is given by

$$\boldsymbol{F}_j = \boldsymbol{B}q_{m_j} \qquad [10\text{-}2]$$

where B is the value of the magnetic induction at the position of the pole. By comparing equations [10-1] and [10-2], we see that the magnitude of the magnetic induction at a distance r from a magnetic pole q_m is given by

$$B = \frac{Cq_m}{r^2} \qquad [10\text{-}3]$$

Let us now deduce the value, for the mks system of units, of the constant C in Coulomb's law for magnetic poles (equation [10-1]). This we can do by studying the interaction between a circular current and a magnetic pole located at its center (Fig. 10.2). We shall calculate, and then equate, the force exerted by a current upon the pole and the force exerted by the pole upon the current.

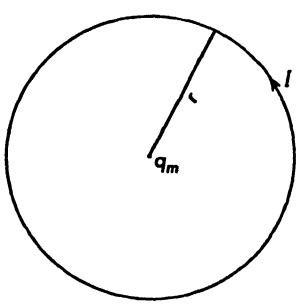


Fig. 10.2

The circular current in a vacuum produces an induction at its center, given by equation [8-17], of magnitude

$$B = \frac{\mu_0 I}{2r} \text{ webers/m}^2 \qquad [8\text{-}17]$$

Hence, the force experienced by the pole q_m is

$$F = Bq_m = \frac{\mu_0 I q_m}{2r} \text{ newtons} \qquad [10\text{-}4]$$

Meanwhile, the pole produces an induction given by equation [10-3]. Since this induction is everywhere normal to the direction of the current, the current experiences a force (equation [8-77])

$$F = BIl = \left(\frac{Cq_m}{r^2}\right)(I)(2\pi r) \text{ newtons} \qquad [10\text{-}5]$$

From Newton's third law, these two forces are equal. Equating them, we find

$$\frac{2\pi C q_m I}{r} = \frac{\mu_0 I q_m}{2r} \qquad [10\text{-}6]$$

or

$$C = \frac{\mu_0}{4\pi} \qquad [10\text{-}7]$$

Substituting this value in equation [10-1], we have Coulomb's law for magnetic poles in the mks system:

$$F_{ij} = \frac{\mu_0}{4\pi}\frac{q_{m_i} q_{m_j}}{r_{ij}^2} \qquad [10\text{-}8]$$

The dimensions of the mks pole are obtained as follows. Since $\boldsymbol{B}$ (from $\boldsymbol{F} = q\boldsymbol{v} \times \boldsymbol{B}$) has dimensions of F/qv, the dimensions of q_m are seen from equation [10-2] to be

$$[q_m] = \left[\frac{F}{\frac{F}{qv}}\right] = [qv] = Qlt^{-1} = (Q/t)l \qquad [10\text{-}9]$$

Evidently, the units of the mks pole are *ampere-meters*.

We shall utilize the foregoing results after briefly discussing the manner in which electromagnetic units are derived from Coulomb's law.

10.3 The Electromagnetic System of Units

Until fairly recently all magnetic effects were discussed in terms of electromagnetic units. Indeed, as was mentioned earlier, frequent reference is still made to the *gauss* and the *oersted*, the cgsm (centimeter-gram-second-magnetic) units of $\boldsymbol{B}$ and $\boldsymbol{H}$, respectively. Since the electromagnetic system has its starting point in Coulomb's law for magnetic poles, we shall now briefly discuss it and, in so doing, supplement the remarks concerning it that were made in § 8.5.

In the cgsm system the reciprocal of the constant in equation [10-1] is the *permeability*, μ_{em}, a dimensionless quantity having the value unity in free space. Hence, the cgsm expression of Coulomb's law is

$$F_{ij} = \frac{q_{m_i} q_{m_j}}{\mu_{em} r_{ij}^2} \text{ dynes} \qquad [10\text{-}10]$$

Thus, the *cgsm unit pole, when placed 1 centimeter away from a similar pole in a vacuum, is repelled by it with a force of 1 dyne.*

If a small magnetic pole, q_m, is brought into a magnetic field and there experiences a force of $\boldsymbol{F}$ dynes, the *cgsm magnetic field intensity*, $\boldsymbol{H}$, is defined as

$$\boldsymbol{H} = \lim_{q_m \to 0} \frac{\boldsymbol{F}}{q_m} \qquad [10\text{-}11]$$

corresponding to the definition (equation [1-11]) of the electric field intensity, $\boldsymbol{E}$. *An intensity of 1 dyne per unit pole is called 1 oersted.* Thus, in a magnetic field of intensity $\boldsymbol{H}$ oersteds a pole of q_{m_j} cgsm units experiences a force

$$\boldsymbol{F}_j = \boldsymbol{H} q_{m_j} \text{ dynes} \qquad [10\text{-}12]$$

Magnetic induction, or *flux density*, is related to the intensity through the permeability:

$$\boldsymbol{B} = \mu_{em} \boldsymbol{H} \qquad [10\text{-}13]$$

The cgsm unit of $\boldsymbol{B}$ is the *gauss*. Because of the dimensionless character of μ_{em}, $\boldsymbol{B}$ and $\boldsymbol{H}$ have the same dimensions in the electromagnetic system.

Relation Between Unit cgsm Pole and Unit cgsm Current

The *abampere*, the cgsm unit of current, was introduced in § 8.5. We shall now relate it to the cgsm unit pole defined above.

From § 8.5 we know the cgsm expression for the magnetic induction at the center of a circular current to be

$$B = \mu_{em} \frac{2\pi I}{r} \text{ gauss} \qquad [8\text{-}20]$$

whence

$$H = \frac{2\pi I}{r} \text{ oersteds} \qquad [10\text{-}14]$$

Thus, the *abampere*, when flowing in a circular loop of 1 centimeter radius, gives rise to a magnetic intensity of 2π oersteds at the center of the loop. In other words, *it is that current which, when flowing in a circular loop of 1 centimeter radius, exerts a force of 2π dynes upon a unit cgsm pole located at the center.*

Incorrect Emphasis upon H Rather than upon B

The perceptive reader will have noticed the fundamental difference between equations [10-2] and [10-12]. On the one hand, in the mks system, the force experienced by a magnetic pole is determined by $\boldsymbol{B}$, while on the other hand, in the cgsm system, the force on the pole is determined by $\boldsymbol{H}$. In both cases, of course, the pole picture is a fiction, providing a convenient means of representing the behavior at a distance of a closed current or group of closed currents. The use of the fiction does not alter the fact that all magnetic forces are exerted upon electric currents, not poles, and their magnitude is determined by $\boldsymbol{B}$ rather than by $\boldsymbol{H}$. Thus, the mks relationship (equation [10-2]) is correct to the extent that the pole representation is valid, while the cgsm relationship (equation [10-12]) is misleading. Since $\boldsymbol{B}$ and $\boldsymbol{H}$ have the same free-space value in the cgsm system, the consequences of the above error do not manifest themselves unless the medium involved has a permeability that differs from that of free space.

It is this error underlying the electromagnetic system that is responsible for the illogical situation described in § 8.3. Fortunately, the mks system provides an opportunity to make a fresh start, in which the joint-primacy of $\boldsymbol{E}$ and $\boldsymbol{B}$ is recognized and an unnecessary source of confusion is removed.

10.4 The Field of a Magnetic Dipole and the Torque Experienced by a Dipole in a Magnetic Field

The small isolated bar magnet shown in Figure 10.3 gives rise to a magnetic induction that is analogous to the electric field produced by an electric dipole. Since the electric case has been worked out in detail in § 4.2, the corresponding results for the magnetic case will simply be put on record.

First, we may define a *magnetic scalar potential*, V_m, which represents the work done in bringing a unit positive pole from infinity to the point P. By a line of reasoning similar to that leading to equation [4-7], we obtain in vacuo

$$V_m(P) \doteq \frac{\mu_0 q_m d \cos\theta}{4\pi r^2} = \frac{\mu_0 m \cos\theta}{4\pi r^2} = \frac{\mu_0 \boldsymbol{m}\cdot\boldsymbol{r}}{4\pi r^3} \qquad [10\text{-}15]$$

where $\boldsymbol{m} = q_m \boldsymbol{d}$ is the *magnetic moment* of the bar magnet.

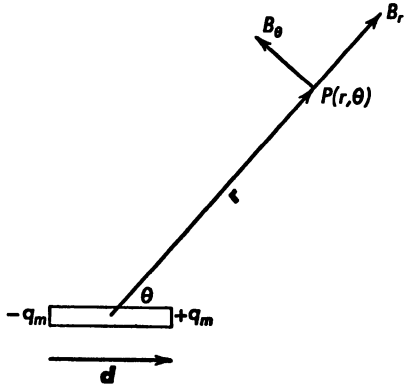


Fig. 10.3

The magnetic induction may be derived by differentiating the potential.[1] Thus,

$$B_r = -\frac{\partial V_m}{\partial r} \doteq \frac{\mu_0 m \cos\theta}{2\pi r^3} \text{ webers/m}^2 \qquad [10\text{-}16]$$

and

$$B_\theta = -\frac{\partial V_m}{r\partial\theta} \doteq \frac{\mu_0 m \sin\theta}{4\pi r^3} \text{ webers/m}^2 \qquad [10\text{-}17]$$

It is emphasized that equations [10-15] to [10-17] presuppose that $r \gg d$.

Torque Experienced by a Magnetic Dipole in a Field of Uniform Induction

The situation to be studied is presented in Figure 10.4. A dipole of length d and pole strength q_m is inclined at an angle θ to a field of uniform induction $\boldsymbol{B}$. The moment of the dipole has the value

$$\boldsymbol{m} = q_m \boldsymbol{d}$$

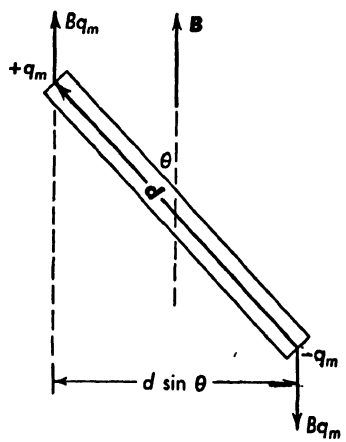


Fig. 10.4

The forces acting upon the two poles of the dipole are equal and opposite to one another, as shown, with the result that the magnet undergoes no translational motion. Since, however, the two forces do not act in the same straight line, the dipole is subject to a *turning couple*, or *torque*, of magnitude:

$$T = Bq_m d \sin\theta = Bm \sin\theta \qquad [10\text{-}18]$$

The vector nature of this torque may be expressed (see § 2.1) as follows:

$$\boldsymbol{T} = \boldsymbol{m} \times \boldsymbol{B} \qquad [10\text{-}19]$$

We shall discover in the next section that a closed current in a magnetic field experiences a similar sort of torque.

---

[1] It will be recalled from § 8.9 that, if currents are present, the magnetic scalar potential is not a single-valued function. If, in bringing the unit positive pole from infinity to the point P, we encircle a current n times, we must add to expression [10-15] an additional term $\pm n\mu I$ (cf. equation [8-39]). However, this additional term disappears in the differentiation of V_m, so that the resulting value of B is single-valued.

10.5 Magnetic Dipole Equivalence of a Current Loop

We have demonstrated above that a magnetic dipole located in a field of uniform induction suffers a torque,

$$T = m \times B \qquad [10\text{-}19]$$

which tends to rotate it to a position parallel to the field. We shall now examine the torque experienced by a current-carrying loop when it is placed in a similar magnetic field.

Consider the rectangular current loop shown in Figure 10.5. The axis of rotation of the loop coincides with the X-axis, while the entire region is pervaded by a uniform magnetic flux, $B = jB$, which is parallel to the Y-axis. The area of the loop is represented by the vector A, drawn, as usual, normal to the plane of the loop. The force on any element dl of the current loop is given by

$$dF = Idl \times B \qquad [8\text{-}3]$$

The ends of the loop lie in the Y-Z plane. It will be seen that the magnetic forces on the two ends of the loop are not only equal in magnitude, but act in opposite directions in the same straight line. The resultant force and torque are both zero.

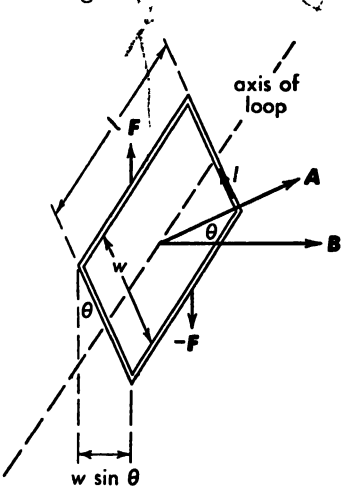


Fig. 10.5

On the other hand, the forces on the two sides of the loop, although equal in magnitude and opposite in direction, do not act in the same straight line. The loop thus experiences a torque of magnitude:

$$T = Fw \sin \theta \qquad [10\text{-}20]$$

The magnitude of F is seen from equation [8-3] to be

$$F = \int_0^l I|dl \times B| = IlB \qquad [10\text{-}21]$$

enabling us to rewrite equation [10-20] in the form

$$T = IlwB \sin \theta$$

or

$$T = IAB \sin \theta$$

or, in vector form,

$$T = IA \times B \qquad [10\text{-}22]$$

When equation [10-22] is compared with equation [10-19], it is seen that the rectangular current loop is equivalent to a magnetic dipole of moment

$$m = IA \qquad [10\text{-}23]$$

This equivalence holds for any shape of closed current, A being in every case the area, regular or irregular, which the current encloses. The magnetic moment of the current loop is normal to the plane of the loop and points in the direction a right-handed screw would advance when rotated in the direction of the current. Hereafter, we shall usually represent the magnetic character of a current loop by its magnetic moment vector.

The dipole nature of a closed current was first recognized by Ampère and led him to suspect that all magnetic phenomena could be attributed to elementary closed currents—*Ampèrian currents*, as they came to be called. The correctness of this suspicion has been completely borne out.

A dipole and a closed current can be considered equivalent only at distances which are large compared to the dimensions of the current loop. When the area enclosed by the current is large, the current loop is equivalent, at a distance, to a *family* of dipoles (§ 10.6).

10.6 The Magnetic Shell and Its Closed Current Equivalent

A *magnetic shell* is a thin sheet of material that is everywhere magnetized at right angles to its surface. It can be visualized as consisting of a large number of magnetic dipoles whose positive poles lie in one surface of the sheet, whose negative poles lie in the other, and whose moments are parallel to the vectors representing the elements of surface. Although actual magnetic shells are rarely encountered in nature, the concept is useful for the reason that an easily derived equivalence exists between magnetic shells and closed currents. This equivalence frequently simplifies the calculation of the magnetic effect of electric currents.

In Figure 10.6 the current I, which encloses the total area A, is shown resolved into a large number of Ampèrian currents, each of magnitude I and each enclosing a surface element. At all interior boundaries of these surface elements the currents cancel one another, so that the array of Ampèrian currents is in fact equivalent to the single large loop. Thus, we are justified in replacing the large closed loop by many small dipoles (one for each Ampèrian current) which collectively constitute a magnetic shell. We therefore conclude that extensive closed currents can be pictured as magnetic shells, or vice versa.

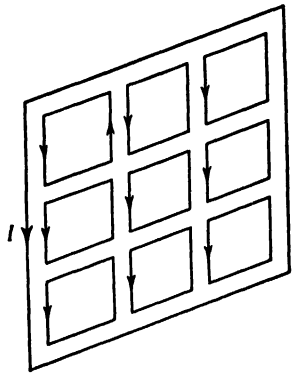


Fig. 10.6

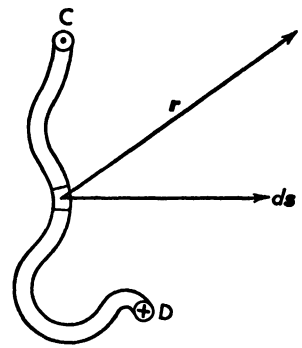


Fig. 10.7

The *strength of a magnetic shell*, ψ, is defined as the magnetic moment per unit area of its surface:

$$\psi = \frac{dm}{ds} \qquad [10\text{-}24]$$

This, from equation [10-23], can be expressed in terms of the current encircling the shell:

$$\psi = \frac{IA}{A} = I \qquad [10\text{-}25]$$

At a distant point, therefore, the magnetic field produced by the closed current I is identical with that produced by a magnetic shell whose area is that enclosed by the current and whose strength is equal to the encircling current.

■ *Indifference of the Shape of a Magnetic Shell*

Let us consider in detail the magnetic shell shown in cross-section in Figure 10.7. It is bounded by a peripheral current of I amperes with direction out of the paper at C and into the paper at D.

The dipole moment, dm, of the element of the shell of area ds is given by $\psi ds = Ids$. This gives rise, in accordance with equation [10-15], to a magnetic potential at P of magnitude

$$dV_m(P) = \frac{\mu Ids \cdot r}{4\pi r^3} = \frac{\mu I}{4\pi}\left(\frac{ds \cdot r}{r^3}\right) \qquad [10\text{-}26]$$

or

$$dV_m(P) = \frac{\mu I}{4\pi} d\omega \qquad [10\text{-}27]$$

The reader will recognize that the quantity $(ds \cdot r)/r^3$ is, as indicated, an element of solid angle (see § 1.8).

The total magnetic potential at P, arising from the entire magnetic shell, is given by

$$V_m(P) = \frac{\mu I}{4\pi}\int_0^{\omega} d\omega = \frac{\mu I}{4\pi}\omega \qquad [10\text{-}28]$$

or

$$V_m(P) = \frac{\mu}{4\pi}\psi\omega \qquad [10\text{-}29]$$

where ω is the solid angle subtended at P by the entire magnetic shell or, alternatively, the solid angle subtended at P by the closed current.

We thus see, from equation [10-29], that *all* magnetic shells with the same strength and boundary are equivalent (provided, of course, that μ is constant). If, therefore, in the course of solving a problem, we replace a closed current by its equivalent magnetic shell, we are free to choose whatever shape of shell is best suited to the specific problem at hand (see Problem 10.11).

Equation [10-28] provides a direct means of computing the magnetic scalar potential of a closed current at any point. V_m is simply the product of $(\mu I/4\pi)$ and the solid angle subtended at the point by the current.

10.7 Magnetic Moment of an Orbital Electron: The Bohr Magneton

The expression derived above (equation [10-23]) for the magnetic moment of a current loop, is immediately applicable to the case of an electron following a circular orbit within an atom. In this case the circular current has the magnitude

$$I = \text{charge/time} = \frac{e}{T} \qquad [10\text{-}30]$$

where e is the elementary charge and T is the period of the motion.

As mentioned in a footnote to § 1.2, the Bohr condition for a permitted orbit is that the angular momentum associated with the electron's orbital motion be an integral multiple of the quantity $h/2\pi$:

$$m_e v r = \frac{nh}{2\pi} \qquad [10\text{-}31]^{\ddagger}$$

where m_e and v represent the electron's mass and velocity, respectively, r is the radius of its orbit, and h is Planck's constant. We may substitute in [10-31] the value of the electron's velocity,

$$v = \frac{2\pi r}{T} \qquad [10\text{-}32]$$

to obtain the period of the electron's motion

$$T = \frac{4\pi^2 m_e r^2}{nh} \qquad [10\text{-}33]$$

Using equations [10-30] and [10-33], and noting that the area of a circular

‡ For this section only, in order to distinguish between magnetic moment and electronic mass, we shall designate the latter by m_e.

orbit is πr^2, we may now write the expression for the magnetic moment of the orbital motion of the electron:

$$m = IA = e\left[\frac{nh}{4\pi^2 m_e r^2}\right]\pi r^2 = \frac{neh}{4\pi m_e} \qquad [10\text{-}34]$$

Thus, for example, in the lowest-energy or *ground* state ($n = 1$) of the hydrogen atom, the magnetic moment of the electron's orbital motion is

$$m = \frac{eh}{4\pi m_e} \qquad [10\text{-}35]$$

This quantity is known as the *Bohr magneton*. It has the value 9.27×10^{-24} ampere-m^2, and is the unit in which all atomic magnetic moments are expressed.

It is pertinent to note that, by means of equation [10-31], equation [10-34] may be rewritten in the form

$$m = \frac{e}{2m_e}(m_e v r) \qquad [10\text{-}36]$$

indicating that the magnetic moment of the electron's orbital motion is proportional to its angular momentum. Designating angular momentum by the vector, $\vec{G}$, we may rewrite [10-36]:

$$\boldsymbol{m} = -\frac{e}{2m_e}\boldsymbol{G} \qquad [10\text{-}37]$$

where the negative sign expresses the fact that the current in an electron orbit is opposite in direction to the electron's motion. The magnetic moment and angular momentum vectors are thus oppositely directed.

■ *Elliptic Electron Orbits*

Equation [10-37] has been derived on the simplifying assumption that the electron's orbit is circular. It can easily be demonstrated that the result is equally true for elliptic orbits. First, we note with reference to Figure 10.8, which represents an electron at one position in its orbit, that the instantaneous value of the angular momentum is

$$\boldsymbol{G} = m\boldsymbol{r} \times \boldsymbol{v} \qquad [10\text{-}38]$$

or

$$G = mvr \sin\theta \qquad [10\text{-}39]$$

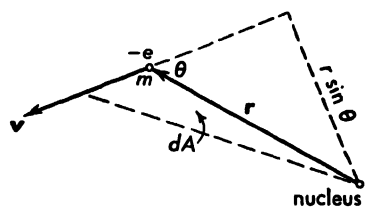


Fig. 10.8

For motions of this type, in which the particle moves in a plane under the action of a force that is always directed toward the origin (the nucleus)[3], the angular momentum is a constant. In general the rate of change of angular momentum about any fixed point is equal to

[3] Such a force is called a *central* force.

the moment of the force about the point. For a central force, the moment of the force is zero and, hence, the angular momentum is constant.

Next we calculate the area of the orbit, beginning with the element of area, dA, shown in Figure 10.8. This is the area swept out in time dt and has the value

$$dA = \tfrac{1}{2}(v\,dt)(r\sin\theta) \qquad [10\text{-}40]$$

or

$$dA = \frac{\mathbf{r}\times\mathbf{v}}{2}\,dt = \frac{G}{2m_e}\,dt \qquad [10\text{-}41]$$

From equation [10-41], recalling that the angular momentum is constant, the area of the orbit is found to be

$$A = \int_S dA = \frac{G}{2m_e}\int_0^T dt = \frac{GT}{2m_e} \qquad [10\text{-}42]$$

Now, using [10-30], which, of course, holds for any shape of orbit, the magnetic moment of the orbital motion is seen to be

$$m = IA = -\left(\frac{e}{T}\right)\left(G\frac{T}{2m_e}\right) = -\frac{e}{2m_e}G \qquad [10\text{-}43]$$

in agreement with equation [10-37].

Magnetic Moment Associated with Particle Spin

Reference has been made in § 6.2 to the intrinsic angular momentum of the electron. Inasmuch as the electron is a charged entity, its intrinsic angular momentum should represent a circulation of charge. Such is the case: the electron possesses a spin magnetic moment, although it is approximately twice as great as one would expect on the basis of relation [10-37]. Apparently electron spin angular momentum is about twice (actually 2.00232) as effective as electron orbital angular momentum in creating magnetic moments.

The proton and the neutron also possess spin angular momenta equal to that of the electron. As one would expect, the spin magnetic moment of the proton is parallel to its angular momentum vector. Because of the difference between the electron and proton masses, one would further expect (from equation [10-37]) the magnetic moment of the proton to be $\sim 1/1840$ that of the electron. This expectation is roughly, but by no means exactly, borne out. The magnetic moment of the neutron is of the same order of magnitude as that of the proton, but has the opposite sign. The fact that its magnetic moment is not zero presumably indicates that the neutron, although electrically neutral, possesses some sort of structure.

10.8 Potential Energy Associated with the Orientation of a Magnetic Dipole or a Closed Current in a Uniform Magnetic Field

The potential energy of an electric dipole in a uniform electric field has already been discussed in Problem 4.6. In a similar manner the potential energy of a magnetic dipole, say, a bar magnet, in a uniform magnetic field depends upon its orientation with respect to the field. As in Problem 4.6, the bar magnet is assumed to possess zero potential energy when its moment is normal to the field (left side of Fig. 10.9). The potential energy associated with any other direction of the moment is equal to the work done in turning it to that position.

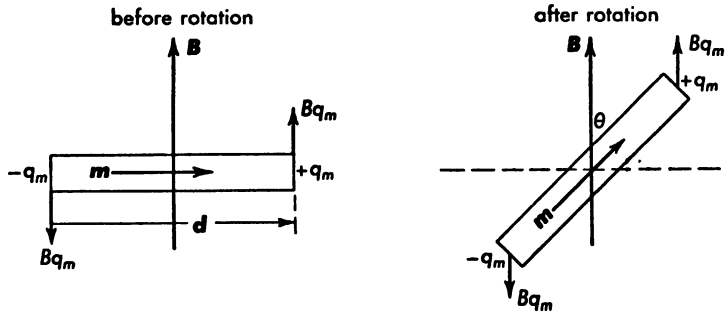


Fig. 10.9

Let us suppose that the bar magnet is rotated from the zero position until its moment makes an angle θ with the field, as shown on the right side of Figure 10.9. During this particular motion the external work required is negative, inasmuch as the motion takes place under the action of the field. Since the force on each pole is the product of the induction B and the pole strength q_m, the external work done per pole during this rotation is

$$\text{Work per pole} = -Bq_m \frac{d}{2} \cos\theta \qquad [10\text{-}44]$$

The potential energy associated with the new orientation is then

$$\text{P.E.} = 2\left[-Bq_m \frac{d}{2} \cos\theta\right]$$

or

$$\text{P.E.} = -\boldsymbol{m}\cdot\boldsymbol{B} \qquad [10\text{-}45]$$

The potential energy of a current loop in a field of uniform induction is, of course, given by equation [10-45] as well.

10.9 Zeeman Effect

The dependence of potential energy upon the orientation of a current loop is responsible for the splitting of spectral lines when the radiating atoms are located in a magnetic field. This effect was first observed in 1896 by the Dutch physicist, Pieter Zeeman (1865–1943), and has played an important part in the explanation of atomic spectra. It will now be explained.

Most atoms, by virtue of the motion of their orbital electrons,[4] including electron spin, possess net magnetic moments. Such atoms, when placed in a magnetic field, are not allowed to orient themselves arbitrarily with respect to the field. Instead, as predicted by the quantum theory, their magnetic moments can point only in certain permitted directions, a restriction known as *space quantization*. Actually, it is the projection of the magnetic moment in the direction of the field that is quantized. The larger the value of the atom's magnetic moment, the larger the number of orientations permitted to it. Associated with each permitted orientation there is, of course, a value of potential energy given by equation [10-45]. Thus, to whatever internal energy the atom otherwise possesses, the amount $-\boldsymbol{m}\cdot\boldsymbol{B}$ joules must be added when it is located in a magnetic field.

The emission of an atomic spectral line represents a transition in the atom from a higher to a lower energy state. Since different electronic orbits are involved in the two cases, the magnetic moments of the two states are not necessarily the same. For example, the *ground* state of a zinc atom has a net magnetic moment that is zero. The thirty planetary electrons belonging to the zinc atom are grouped into fifteen pairs, the members of each pair circulating in similar orbits but in opposite directions. The orbital magnetic moments and the spin magnetic moments of the thirty electrons neutralize one another. In *excited* states of a zinc atom, however, this complete neutralization seldom occurs. The magnetic moment of certain of the excited states has such a value, say, $\boldsymbol{m}$, that it is permitted three orientations in a magnetic field, making angles of $\pi/4$, $\pi/2$, and $3\pi/4$ radians, respectively, with it, as shown in Figure 10.10.

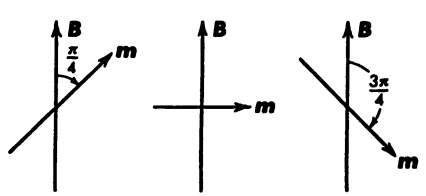


Fig. 10.10

[4] Actually, the particles in the nucleus also contribute to the total magnetic moment of the atom, but this contribution is smaller and will be disregarded here.

Let us consider transitions to the ground state of zinc from one of the excited states that has this value of magnetic moment. These transitions, in the absence of a magnetic field, take place between states of energy U_e and U_g, respectively, as shown on the *energy-level diagram* on the left of Figure 10.11. Each transition

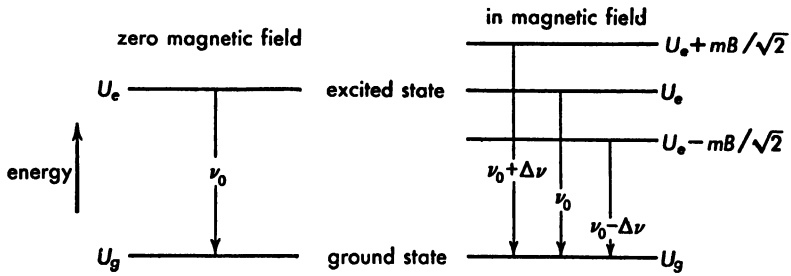


Fig. 10.11

results in the emission of a photon of light of frequency, ν_0, given by:

$$h\nu_0 = U_e - U_g \qquad [10\text{-}46]$$

a relation known as *Bohr's second postulate* (see footnote to § 1.2 for Bohr's first postulate).[5]

When the zinc atoms are placed in a magnetic field, the excited energy level splits into three levels, corresponding to the orientations shown in Figure 10.10 and having energies $U_e - (mB/\sqrt{2})$, U_e, and $U_e + (mB/\sqrt{2})$, respectively, as shown on the right side of Figure 10.11. The ground state, since it has zero magnetic moment, is not altered by the presence of the magnetic field. The radiation emitted in the transition now consists of three components —the original single spectral line has become a triplet. The middle component has the same frequency, ν_0, as the original, but the frequencies of the other two are higher and lower, respectively, than the original and by the same amount,

$$\Delta\nu = \frac{\boldsymbol{m}\cdot\boldsymbol{B}}{h} = \frac{mB}{\sqrt{2}h} \qquad [10\text{-}47]$$

This splitting of a spectral line into *three* components when the source of light is placed in a magnetic field is known as the *normal* Zeeman effect. The term *anomalous* Zeeman effect is applied when the number of components exceeds three. Despite its name, this occurrence is more common than the "normal" behavior.

[5] Niels Bohr based his momentous explanation of the hydrogen spectrum on these two postulates.

10.10 Force on Dipole in a Nonuniform Magnetic Field: The Stern-Gerlach Experiment

As we have seen, a dipole that is located in a uniform field experiences a torque of magnitude $\boldsymbol{m} \times \boldsymbol{B}$, which tends to line it up with the field. In a nonuniform field it experiences, in addition, a translatory force. We shall now compute the magnitude of the translatory force by reference to Figure 10.12.

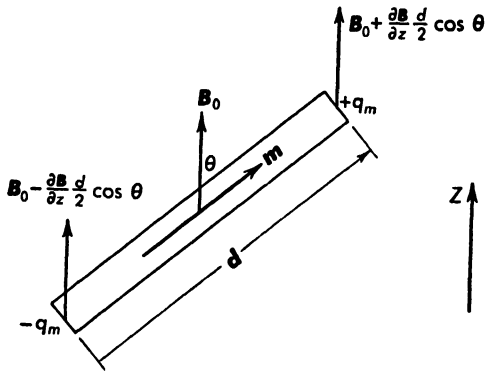


Fig. 10.12

Here the magnet of moment $\boldsymbol{m} = q_m\boldsymbol{d}$ is inclined at an angle θ to a field of induction $\boldsymbol{B}$ that is directed in the Z-direction and increases with increasing z at the rate $\partial \boldsymbol{B}/\partial z$ webers/m^3. The value of $\boldsymbol{B}$ at the center of the magnet is $\boldsymbol{B}_0$ webers/m^2. Hence, its values at the $+q_m$ and $-q_m$ poles are, respectively,

$$\boldsymbol{B}_0 + \left(\frac{\partial \boldsymbol{B}}{\partial z}\right)\frac{d}{2}\cos\theta \qquad \text{and} \qquad \boldsymbol{B}_0 - \left(\frac{\partial \boldsymbol{B}}{\partial z}\right)\frac{d}{2}\cos\theta \qquad [10\text{-}48]$$

The net force experienced by the magnet is given by:

$$\boldsymbol{F} = \boldsymbol{k}\left[B_0 + \left(\frac{\partial B}{\partial z}\right)\frac{d}{2}\cos\theta\right]q_m - \boldsymbol{k}\left[B_0 - \left(\frac{\partial B}{\partial z}\right)\frac{d}{2}\cos\theta\right]q_m \qquad [10\text{-}49]$$

or

$$\boldsymbol{F} = \boldsymbol{k}\left(\frac{\partial B}{\partial z}\right)q_m d\cos\theta \qquad [10\text{-}50]$$

or

$$\boldsymbol{F} = \boldsymbol{k}\left(\frac{\partial B}{\partial z}\right)m\cos\theta \qquad [10\text{-}51]$$

where $\boldsymbol{k}$, as usual, is the unit vector in the Z-direction. The magnet thus experiences a force toward the stronger portion of the field. This force is proportional to the gradient of the induction and to the component of the magnetic moment in the direction of the field.

The Stern-Gerlach Experiment

It follows from equation [10-51] that an atom possessing a magnetic moment $\boldsymbol{m}$ will experience a translatory force of magnitude

$$\left(\frac{\partial B}{\partial z}\right) m \cos\theta \qquad [10\text{-}52]$$

if placed in a nonuniform magnetic field. This type of force was utilized by Stern and Gerlach (Hamburg, 1922) to provide direct evidence of the existence of space quantization (that is, that the magnetic moment of an atom can only assume certain orientations with respect to an external magnetic field).

The inhomogeneous magnetic field needed for this experiment was obtained by employing a magnet, shown in cross-section on the left in Figure 10.13, with

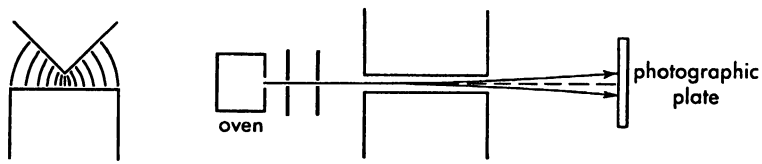


Fig. 10.13

strikingly dissimilar pole shapes. This combination of sharp and flat poles gave rise, as the sketched lines of force indicate, to a field with a large gradient.

Silver atoms, evaporated from an oven and collimated thereafter by slits, were allowed to traverse the length of the inhomogeneous field, as shown on the right in Figure 10.13. It was found that the beam of atoms was split into two components, one of which had suffered a positive displacement, while the other had undergone a negative one of equal magnitude.

The result indicates that for the magnetic moment of the silver atom only two orientations are permitted, making angles of θ and $\pi - \theta$, respectively, with the direction of the applied magnetic field. This follows from equation [10-51], since the displacements of the two components were equal and opposite. In this way not only was the reality of space quantization directly established, but the magnetic moment of the ground state of the silver atom was determined as well.

10.11 Gauss' Theorem for Magnetism

We have seen that the integral of the electric displacement over any closed surface is numerically equal to the free charge enclosed within the surface, a result known as Gauss' theorem:

$$\int_s \boldsymbol{D}\cdot d\boldsymbol{s} = \int_v \rho\, dv \qquad [4\text{-}25]$$

When the net charge within the surface is positive the value of the integral $\int_s \boldsymbol{D}\cdot d\boldsymbol{s}$ is positive, indicating that the displacement vectors at the surface are pointing predominantly outward. This can be visualized by imagining "lines" of displacement that originate upon positive charges and terminate upon negative charges. Thus, if the net charge is positive, more lines of displacement emerge through the surface than enter it.[6] It is seen by this qualitative argument that it is the ability of charge of one sign to exist independently of charge of the other sign that is responsible for the non-vanishing of the integral $\int_s \boldsymbol{D}\cdot d\boldsymbol{s}$.

With magnetism the situation is entirely different. The elementary magnets are small closed currents. Lines of induction emerge from one face of each current and converge toward the other: that is, the lines of induction associated with each elementary current are continuous and closed. Consequently, it is not possible to construct a surface through which more lines of induction emerge than enter. In short, the integral of the induction over any closed surface is zero. Gauss' theorem for magnetic media is then

$$\int_s \boldsymbol{B}\cdot d\boldsymbol{s} = 0 \qquad [10\text{-}53]$$

Since $\boldsymbol{B}\cdot d\boldsymbol{s} = d\Phi_m$, Gauss' theorem may be expressed in words by the statement that the total flux of magnetic induction through any closed surface is zero.

Maxwell's Fourth Equation

By means of the divergence theorem (equation [2-17]) equation [10-53] may be written as

$$\int_v \nabla\cdot\boldsymbol{B}\, dv = 0 \qquad [10\text{-}54]$$

[6] If desired, this picture could be developed further. For example, we could define the displacement quantitatively in terms of the density of such lines, in which case Gauss' theorem would give the total number of displacement lines emerging from the closed surface in question.

which leads to the differential form of Maxwell's fourth equation:

$$\nabla \cdot \boldsymbol{B} = 0 \qquad [10\text{-}55]$$

■ This result can be arrived at analytically by use of the magnetic vector potential. Since

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} \qquad [8\text{-}59]$$

we have

$$\nabla \cdot \boldsymbol{B} = \nabla \cdot (\nabla \times \boldsymbol{A}) \qquad [10\text{-}56]$$

By making use of equations [2-14] and [2-15] it is easily demonstrated (Problem 10.13) that the right-hand side of equation [10-56] vanishes. Thus Maxwell's fourth equation is a direct consequence of the fact that the induction is the curl of a vector.

10.12 Refraction of B and H in Passing from One Magnetic Medium to Another

We are now in a position to predict the relative directions of $\boldsymbol{B}$ and $\boldsymbol{H}$ at the interface between two magnetically different media. The method employed is similar to that used in studying the behavior of $\boldsymbol{D}$ and $\boldsymbol{E}$ at an interface (§§ 4.7 and 4.8). In Figure 10.14 the line AB lies in the interface between

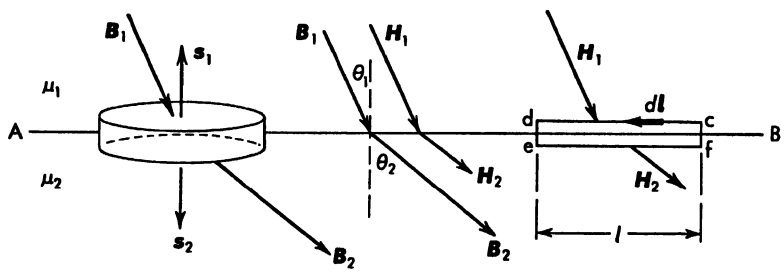


Fig. 10.14

magnetically isotropic materials of permeability μ_1 and μ_2, respectively. $\boldsymbol{B}_1$ (and $\boldsymbol{H}_1$) and $\boldsymbol{B}_2$ (and $\boldsymbol{H}_2$) make angles θ_1 and θ_2, respectively, with the normal to the interface.

We shall study the behavior of $\boldsymbol{B}$ at the boundary by applying Gauss' theorem (equation [10-53]) to the pillbox shown on the left of Figure 10.14. Since the area of the sides of the pillbox is negligibly small, we have

$$0 = \int_s \boldsymbol{B} \cdot d\boldsymbol{s} = \boldsymbol{B}_1 \cdot \boldsymbol{s}_1 + \boldsymbol{B}_2 \cdot \boldsymbol{s}_2 \qquad [10\text{-}57]$$

or

$$0 = (-B_1 \cos \theta_1 + B_2 \cos \theta_2)s \qquad [10\text{-}58]$$

whence,

$$B_1 \cos \theta_1 = B_2 \cos \theta_2 \qquad [10\text{-}59]$$

Hence, *the normal component of the induction is continuous across the boundary.*

The behavior of $\boldsymbol{H}$ is studied with reference to the rectangular path shown on the right in Figure 10.14. Since the closed path encloses zero current, the line integral of $\boldsymbol{H}$ around the path is zero:

$$\oint \boldsymbol{H} \cdot d\boldsymbol{l} = 0 \qquad [8\text{-}40]$$

When this expression is written out, the contributions associated with the ends of the rectangle can be disregarded. Thus,

$$0 = \int_c^d \boldsymbol{H}_1 \cdot d\boldsymbol{l} + \int_e^f \boldsymbol{H}_2 \cdot d\boldsymbol{l} \qquad [10\text{-}60]$$

or

$$0 = (-H_1 \sin \theta_1 + H_2 \sin \theta_2) l \qquad [10\text{-}61]$$

or

$$H_1 \sin \theta_1 = H_2 \sin \theta_2 \qquad [10\text{-}62]$$

Hence, *the tangential component of* $\boldsymbol{H}$ *is continuous across the boundary.*

If now equation [10-62] is divided by equation [10-59], and the result is rearranged, we obtain

$$\frac{\tan \theta_1}{\tan \theta_2} = \frac{B_1 H_2}{B_2 H_1} = \frac{\mu_1}{\mu_2} \qquad [10\text{-}63]$$

This expression describes, in terms of their directions, the behavior of $\boldsymbol{B}$ and $\boldsymbol{H}$ in passing from one medium to another.

The concept of magnetic lines of force is undoubtedly a familiar one to the reader. He has probably at some time been party to experiments in which lines of force have been visualized by means of iron filings. The density of lines of force at a given point in space is proportional to the magnitude of $\boldsymbol{B}$, while the tangent to a line of force provides the direction of $\boldsymbol{B}$. Equation [10-63] may, therefore, be interpreted as describing the refraction of lines of force in passing from one medium to another.

Because of the large values of permeability displayed by ferromagnetic materials (see § 10.18), magnetic lines of force may experience a very large refraction upon entering them. For example, if we consider the two media air ($\mu_1 = \mu_0$) and cast iron ($\mu_2 = 240\mu_0$), we find that the value of θ_1 corresponding to $\theta_2 = \pi/4$ is approximately 0.2°. Thus, except for the values of θ_2 that approach $\pi/2$, the lines of force enter the iron practically normally, as shown in Figure 10.15.

It is evident from Figure 10.15 that there is a strong concentration of lines of force in the iron, with a consequent weakening of the magnetic field in the adjacent regions *A* and *C*. These regions are to a large extent "shielded" from the magnetic field. This principle of magnetic shielding is in common use in the construction of instruments that require for their proper performance the

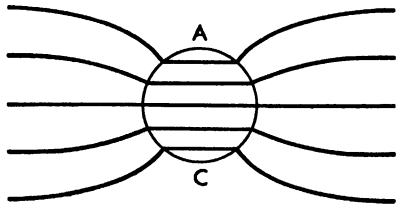


Fig. 10.15

absence of a magnetic field. Certain types of watches are housed in a soft iron case and many pieces of scientific apparatus that demand an environment that is free of a magnetic field are completely encased in iron. Most of the lines of force normally in the region are concentrated in the walls of the iron housing, leaving the interior of the housing practically field-free.

Equations [10-59] and [10-62], governing the behavior of $\boldsymbol{B}$ and $\boldsymbol{H}$ at an interface, correspond exactly to equations [4-33] and [4-35], governing the interface behavior of $\boldsymbol{D}$ and $\boldsymbol{E}$. For this reason, the operational definitions given for $\boldsymbol{D}$ and $\boldsymbol{E}$ in § 4.7 have their $\boldsymbol{B}$- and $\boldsymbol{H}$-counterparts. That is to say, in a disc-shaped cavity (axis parallel to $\boldsymbol{B}$) in a material, $\boldsymbol{B}$ is identical with its value within the material. Also, in a needle-shaped cavity (long axis parallel to $\boldsymbol{B}$), $\boldsymbol{H}$ is identical with its value within the material.

10.13 Magnetic Polarization and Magnetic Susceptibility

So far in our discussion of magnetic effects we have taken little account of the magnetic properties of magnetic materials, except to write μ rather than μ_0 in a number of relationships. Thus, for example, the induction inside a toroidal solenoid is given approximately by

$$B = \mu n I \qquad [8\text{-}43]$$

in the general case, or by

$$B_0 = \mu_0 n I$$

if the interior of the solenoid is a vacuum. If the interior of the coil is filled with a material medium, B may differ from B_0. On the other hand H has, for both cases, the same value:

$$H = \frac{B}{\mu} = \frac{B_0}{\mu_0} = nI \qquad [10\text{-}64]$$

Since H depends only upon the nature of the coil and the current which it carries, and not at all upon the core material, it has come to be regarded as

a *magnetizing force.* In response to this applied force the elementary magnets in the core material respond, either singly or collectively. This response is described by the term *magnetic polarization*, or *magnetization.* For different core materials the magnetic polarization differs greatly in degree, as we shall see, and even in sign. For certain materials the induction [7] within the core is very much larger than B_0. Such is the case in soft iron, for example, where the response of the elementary magnets to the magnetizing force is vigorous, as they align themselves in the direction of H. In this case the magnetic polarization is said to be positive.

The qualitative picture, then, which we shall presently sharpen up with proper quantitative definitions, is the following: when a magnetizing force $\boldsymbol{H}$ (which is to be identified with some type of current flow) is applied to a volume of space, it gives rise to an induction $\boldsymbol{B}_0 = \mu_0 \boldsymbol{H}$. If, however, the volume is occupied by some material medium, the resulting induction, $\boldsymbol{B}$, may differ from $\boldsymbol{B}_0$ as a result of the polarization of the material. Thus,

$$\boldsymbol{B} = \boldsymbol{B}_0 + \text{contribution from magnetic polarization} \qquad [10\text{-}65]$$

or

$$\boldsymbol{B} = \mu_0 \boldsymbol{H} + \text{contribution from magnetic polarization} \qquad [10\text{-}66]$$

The exact form of the "contribution from magnetic polarization" depends upon the manner in which magnetic polarization is defined.

The magnetic polarization, $\boldsymbol{M}$, which is the magnetic analogue of the electric polarization, $\boldsymbol{P}$, is defined for any material as the magnetic moment per unit volume. Thus,

$$\boldsymbol{M} = \frac{d\boldsymbol{m}}{dv} \qquad [10\text{-}67]$$

By definition $\boldsymbol{M}$ is related to $\boldsymbol{H}$ through the *magnetic susceptibility*, χ_m, as follows:

$$M = \chi_m H \qquad [10\text{-}68]$$

The magnetic susceptibility, as was the electric susceptibility, is a dimensionless quantity.[8]

[7] B within the solenoid can be ascertained by winding a secondary coil on the solenoid and observing the emf that is induced in it when the primary circuit is closed. This emf is given by equation [9-5]:

$$\mathcal{E}_2 = -N_2 \frac{d\Phi_{m_2}}{dt} = -N_2 A_1 \frac{dB_1}{dt}$$

since, for this coil configuration, $\Phi_{m_2} = \Phi_{m_1} = B_1 A_1$. This subject is continued in § 10.20.

[8] Equation [10-68] is consistent with the point of view presented above: that the magnetic polarization results from an applied force, $\boldsymbol{H}$. This interpretation of $\boldsymbol{H}$ as a magnetizing force, in response to which magnetic polarization takes place, is an integral part of the modern view of magnetism, and will be employed exclusively hereafter. Although we have concentrated on the calculation of $\boldsymbol{B}$ rather than of $\boldsymbol{H}$, the magnetizing force for any current configuration, such as any one of those studied in Chapter 8, is got simply from $\boldsymbol{H} = \boldsymbol{B}/\mu$.

Clearly, since $\boldsymbol{H}$ and $\boldsymbol{M}$ are dimensionally alike, the proper expression for equation [10-66] is

$$\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M}) \qquad [10\text{-}69]$$

This equation corresponds to equation [4-20] in electrostatics, but differs slightly from it owing to the fact that $\boldsymbol{M}$ has the dimensions of $\boldsymbol{H}$, whereas $\boldsymbol{P}$ was defined to have the dimensions of $\boldsymbol{D}$, not of $\boldsymbol{E}$. Equation [10-69] can also be written as

$$\boldsymbol{B} = \mu_0(\boldsymbol{H} + \chi_m\boldsymbol{H}) \qquad [10\text{-}70]$$

from which we see, by comparing it to

$$\boldsymbol{B} = \mu\boldsymbol{H} \qquad [8\text{-}11]$$

that

$$\mu = \mu_0(1 + \chi_m) \qquad [10\text{-}71]$$

corresponding to the electrostatic counterpart, equation [4-23].

In isotropic media, $\boldsymbol{B}$, $\boldsymbol{H}$, and $\boldsymbol{M}$ have the same directions: consequently χ_m and μ are both scalar quantities. In nonisotropic media, such as crystalline materials, $\boldsymbol{B}$, $\boldsymbol{H}$, and $\boldsymbol{M}$ are rarely parallel. In such media, which will not be considered in this text, χ_m and μ are quantities with nine components (*tensors*). Crystalline materials which consist of large numbers of randomly oriented small crystals behave, macroscopically, as isotropic media.

Some actual relationships that exist between $\boldsymbol{B}$, $\boldsymbol{H}$, and $\boldsymbol{M}$ will be discussed in the sections that follow.

10.14 Diamagnetism, Paramagnetism, and Ferromagnetism

Materials are classified, with respect to their bulk magnetic properties, as *diamagnetic*, *paramagnetic*, or *ferromagnetic*.

Diamagnetic materials have negative values of magnetic susceptibility and consequently possess permeabilities (cf. equation [10-71]) that are less than that of free space. When placed in a magnetic field they display magnetic moments that are antiparallel to the field. In a nonuniform field they experience a force toward the weaker portions of the field.[9]

Paramagnetic materials have small positive values of magnetic susceptibility. When placed in a magnetic field they display magnetic moments that

[9] Any object in a magnetic field moves so as to reduce the total magnetic energy. Since the magnetic energy density is $BH/2 = B^2/2\mu$, materials with $\mu < \mu_0$ move from strong to weak portions of the field, while those with $\mu > \mu_0$ move from weak to strong portions.

are parallel to the applied field. In a nonuniform field they are urged towards the stronger portions of the field.

Ferromagnetic materials have *large* positive values of magnetic susceptibility. They retain a portion of their magnetization when the aligning field is removed. When heated to sufficiently high temperatures they become paramagnetic.

Most of the simple insulators and approximately half the simple metals are diamagnetic. The balance of the insulators and metals, except for the ferromagnetic metals, are paramagnetic. Iron, cobalt, nickel, and certain alloys constitute the ferromagnetic group.

10.15 Diamagnetism

The atoms of a diamagnetic material do not possess permanent magnetic moments. As in the case of zinc, discussed in § 10.9, the individual magnetic moments of the planetary electrons neutralize each other. However, when placed in a magnetic field the atoms become "magnetized" in a direction antiparallel to the field.

This effect is an example of electromagnetic induction. The electronic orbits of the atoms constituting the material in question can be regarded as closed conducting loops in which emf's are induced when the atoms are thrust into a magnetic field. In accordance with Lenz's law, the direction of these induced emf's is such as to give rise to currents whose magnetic fields oppose the applied one. Since the electronic orbits offer zero electrical resistance, these opposing fields persist until the applied field is removed.

The actual mechanism by which this effect is produced is the following: The force that is exerted by the applied magnetic field upon an orbital electron is perpendicular to both the field and the direction of motion of the electron (that is, $\boldsymbol{F} = -e\boldsymbol{v} \times \boldsymbol{B}$). This causes the electronic orbit to precess about the direction of the field, in much the same manner as a spinning top that is tilted with respect to the vertical precesses about the vertical. This so-called *Larmor precession* represents a circulation of charge (in effect, either an increase or decrease of the angular velocity, depending upon their direction of motion) and the dipole moment that is associated with it is always antiparallel to the applied field. Thus the magnetic polarization (or magnetization) is always negative.

This diamagnetic character is displayed by all materials, although it is masked by more prominent effects in paramagnetic and ferromagnetic substances.

10.16 Paramagnetism and Its Variation with Temperature

The atoms or molecules of a paramagnetic material possess *permanent* magnetic moments. These moments tend to align themselves in the direction of an applied magnetic field but are being continually de-aligned as a result of thermal agitation. As a result, there is a net, but by no means predominant, magnetic polarization parallel to the external field. We have discussed at some length, in § 4.11, the electrical analogue of this state of affairs, and will now make use of the results therein derived.

The Curie Law and the Curie Temperature

The average component of the dipole moment in the direction of the applied field (for weak fields or high temperatures) is given by an expression similar to equation [4-74] (and derived in the same way):

$$\overline{m \cos \theta} = \frac{m^2 \mu_0 H_i}{3kT} \qquad [10\text{-}72]$$

where H_i, the "inner field," is the effective magnetizing force acting upon any individual dipole within the material. As one would expect, the degree of with-the-field alignment increases with increasing H_i. Also, because of the increased thermal agitation, the alignment deteriorates with increasing temperature.

If there were no interaction between the permanent magnetic dipoles, the inner field would be identical with the external field, and equation [10-72] would take the form

$$\overline{m \cos \theta} = \frac{m^2 \mu_0 H}{3kT} \qquad [10\text{-}73]$$

This was the result obtained by the French physicist, Paul Langevin (1872–1946), early in this century. Let us carry his line of reasoning a step further.

The value of $m \cos \theta$ is usually written in terms of the magnetic susceptibility, $\chi_m = M/H$ (equation [10-68]). For a substance with n elementary dipoles per unit volume the magnetic polarization (magnetic moment per unit volume) is given by

$$M = \overline{nm \cos \theta} = \frac{nm^2 \mu_0 H}{3kT} \qquad [10\text{-}74]$$

and the susceptibility by

$$\chi_m = \frac{M}{H} = \frac{nm^2\mu_0}{3kT} = \frac{M_0^2\mu_0/3nk}{T} = \frac{C}{T} \qquad [10\text{-}75]^{10}$$

This result, that the magnetic susceptibility varies inversely as the absolute temperature, is known as the *Curie law*. It was discovered empirically by Pierre Curie in 1895 while studying the magnetic susceptibilities of solutions of paramagnetic salts. For such solutions, inasmuch as the elementary dipoles are widely spaced with the effect that their interaction is small, Langevin's analysis is applicable.

For *solid* paramagnetic materials the susceptibility does not obey the Curie law but rather the relationship

$$\chi_m = \frac{C}{T - T_C} \qquad [10\text{-}76]$$

where T_C is a constant. Equation [10-76] can be rationalized, as was first pointed out by Weiss (1907), by postulating that the actual "field" acting upon a dipole is the sum of the applied field, H, and that arising from the presence of neighboring dipoles. His extension of Langevin's work thus began with the use of equation [10-72] instead of equation [10-73]. The effective magnetizing force was then written as

$$H_i = H + \alpha M \qquad [10\text{-}77]$$

explicitly assuming that the contribution of the nearby dipoles to the effective field is proportional to the magnetic polarization. This form of dependence is not self-evident but was chosen by Weiss because, as we shall see below, it leads to equation [10-76]. The factor of proportionality, α, in equation [10-77] is known as the *molecular field coefficient*.

When equation [10-77] is substituted in equation [10-72], and the susceptibility deduced therefrom, we find

$$M = \overline{nm\cos\theta} = \frac{nm^2\mu_0(H + \alpha M)}{3kT} = \frac{n\mu_0 M_0^2(H + \alpha M)}{3n^2kT} \qquad [10\text{-}78]$$

and

$$\chi_m = \frac{M}{H} = \frac{\mu_0 n M_0^2(H + \alpha M)}{3n^2kTH} = \frac{\mu_0 M_0^2}{3nkT} + \frac{\alpha\mu_0 M_0^2\chi_m}{3nkT} \qquad [10\text{-}79]$$

whence

$$\chi_m = \frac{\dfrac{\mu_0 M_0^2}{3nk}}{T - \dfrac{\alpha\mu_0 M_0^2}{3nk}} = \frac{C}{T - \alpha C} = \frac{C}{T - T_C} \qquad [10\text{-}80]$$

[10] $M_0 = nm$ represents the maximum possible value of the magnetic polarization, when all the elementary dipoles are aligned with the field.

Equation [10-80] is known as the *Curie-Weiss law.* The constant T_C $(= \alpha C)$ appearing in the denominator is known as the *Curie temperature.* One might predict from the Curie-Weiss law that, below the Curie temperature, the susceptibility of an otherwise paramagnetic material would be negative: that is, a paramagnetic material would become diamagnetic as it is cooled through the Curie temperature. This prediction overlooks the fact that the analysis leading up to equation [10-72] is predicated upon the existence of high temperatures. For most paramagnetic materials the Curie temperature is so low that even the Curie law (equation [10-75]) is approximately correct at room temperature.

Effect of Space Quantization

In what has been said above, no account has been taken of the fact that the elementary atomic dipoles can only point in certain directions with respect to the applied field. Equation [10-72] has been derived on the assumption that all orientations are possible. This error can be corrected, as has been done by Brillouin, by replacing the integration which led to equation [10-72] by a summation over all permitted orientations. This introduces a significant, but not large, correction. As one would expect, the magnitude of the correction decreases as the number of permitted orientations increases.

10.17 Ferromagnetism

Ferromagnetic materials are characterized by high susceptibility and residual magnetism. The large values of magnetic polarization that these materials display is not to be attributed to atoms possessing exceptionally large magnetic moments. Rather, the effect is caused by the strong tendency in ferromagnetic materials of both atomic and subatomic magnetic moments to align themselves parallel to one another. The existence of an interaction between elementary dipoles that would lead to their parallel alignment was first postulated by Weiss, but the cause of the interaction was not understood until the problem had been studied by quantum-mechanical methods (Heisenberg, 1928).

Weiss' Theory of Ferromagnetism

Weiss' theory consisted of two distinct parts. In the first of these the general theory of paramagnetism outlined in the preceding section was applied to the temperature region below the Curie temperature. Since the results so deduced did not agree with the known behavior of bulk ferromagnetic material, it was assumed by Weiss that they

applied only to small *domains* within the bulk material. The second part of the theory related the properties of the domains to the observed bulk properties.

Equation [10-72], which formed the basis for our discussion of paramagnetism, was written down by analogy to the corresponding electrical relationship (equation [4-74]). Both are approximations which are valid for weak fields or high temperatures. It is necessary in discussing ferromagnetism to write down the exact expression for the average component of the dipole moment in the direction of the applied field. By analogy to equation [4-73] we therefore write

$$\overline{m\cos\theta} = \frac{m\displaystyle\int_0^{\pi} e^{\mu_0 H_i m\cos\theta/kT}\cos\theta\sin\theta\,d\theta}{\displaystyle\int_0^{\pi} e^{\mu_0 H_i m\cos\theta/kT}\sin\theta\,d\theta} \qquad [10\text{-}81]$$

The integrations may be carried out by writing $x = \cos\theta$. Also, for convenience, let us write $a = \mu_0 H_i m/kT$.

Thus,

$$\frac{M}{M_0} = \frac{\overline{nm\cos\theta}}{nm} = \frac{\displaystyle\int_1^{-1} e^{ax}x\,dx}{\displaystyle\int_1^{-1} e^{ax}\,dx} \qquad [10\text{-}82]$$

or

$$\frac{M}{M_0} = \frac{\dfrac{e^{ax}x}{a}\Big]_1^{-1} - \displaystyle\int_1^{-1}\frac{e^{ax}\,dx}{a}}{\dfrac{e^{ax}}{a}\Big]_1^{-1}} = \frac{\dfrac{e^{ax}x}{a} - \dfrac{e^{ax}}{a^2}\Big]_1^{-1}}{\dfrac{e^{ax}}{a}\Big]_1} \qquad [10\text{-}83]$$

or

$$\frac{M}{M_0} = \frac{(-e^{-a} - e^{a}) - \dfrac{e^{-a} - e^{a}}{a}}{(e^{-a} - e^{a})} = \frac{e^{a} + e^{-a}}{e^{a} - e^{-a}} - \frac{1}{a} \qquad [10\text{-}84]$$

Recalling that $\sinh a = \tfrac{1}{2}(e^{a} - e^{-a})$ and $\cosh a = \tfrac{1}{2}(e^{a} + e^{-a})$, we may rewrite [10-84] in the form:

$$\frac{M}{M_0} = \coth a - \frac{1}{a} \qquad [10\text{-}85]$$

The quantity on the right-hand side of this equation is known as the *Langevin function*. For small values of a, as we have already noted, it has the value $\mu_0 m H_i/3kT$ or $a/3$. In this discussion we are interested in the large-a values of the Langevin function as well.

Weiss postulated, as indicated in equation [10-77], that the effective field acting upon any individual dipole is partially determined by an interaction between the dipole and its neighbors. This contribution to the field was assumed to be proportional to the magnetic polarization existing within the sample: thus,

$$a = \frac{\mu_0 m H_i}{kT} = \frac{\mu_0 m}{kT}(H + \alpha M) \qquad [10\text{-}86]$$

Since the dipole-aligning force depends partially upon the degree of alignment (magnetic polarization) already prevailing, any increase in the magnetic polarization will

result in an increase in the aligning force, which, in turn, will result in an increase in the magnetic polarization, etc. In short, a cumulative effect is to be expected. In this way magnetization, once triggered, could occur without the application of an external field. Let us now investigate the conditions under which such *spontaneous magnetization* could take place.

When the applied field is zero, equation [10-86] reduces to

$$a = \frac{\alpha\mu_0 m M}{kT} = \frac{\alpha\mu_0 M M_0}{nkT} \qquad [10\text{-}87]$$

which can be expressed in the form

$$\frac{M}{M_0} = \frac{ankT}{\alpha\mu_0 M_0^2} \qquad [10\text{-}88]$$

Equations [10-85] and [10-88] constitute a pair of equations which may be solved simultaneously for the two variables M/M_0 and a. The solutions are most easily obtained by graphical methods, as shown in Figure 10.16. At the origin, as mentioned above, equation [10-85] has the value

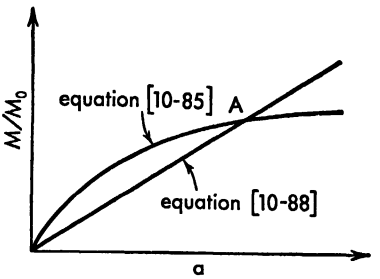


Fig. 10.16

$$\frac{M}{M_0} = \frac{a}{3} \qquad [10\text{-}89]$$

Hence, the slope at the origin of the tangent to the Langevin curve is ⅓. Provided the slope of equation [10-88] is less than this value, a nonzero point of intersection of the two curves exists, as at A. This provision is written symbolically as

$$\frac{nkT}{\alpha\mu_0 M_0^2} < \frac{1}{3} \qquad [10\text{-}90]$$

or

$$T < \frac{\alpha\mu_0 M_0^2}{3nk} \qquad [10\text{-}91]$$

or, recalling from equation [10-80] the value of T_C,

$$T < T_C \qquad [10\text{-}92]$$

Thus, provided the temperature is below the Curie temperature, a state of magnetization is possible without an applied field. This state of spontaneous magnetization is represented by the point of intersection, A, of the two curves in Figure 10.16. Furthermore, as the temperature is reduced, thereby reducing the slope of the straight line belonging to equation [10-88], the spontaneous magnetization becomes an increasingly larger fraction of its maximum possible value, M_0. Incidentally, the Curie temperatures for the ferromagnetic elements Fe, Co, and Ni are 770° C, 1131° C, and 358° C, respectively.

In order to reconcile the prediction of spontaneous magnetization with the fact that ferromagnetic substances are usually found in nature in an *un*magnetized state, Weiss found it necessary to postulate the existence of *domains* whose dimensions are much

smaller than those of the usual ferromagnetic sample. Within a domain the magnetic polarization is practically complete, but since the direction of spontaneous magnetization differs from domain to domain, the net magnetization of the sample is zero. When this type of material is subjected to a magnetizing field the individual domains align themselves with the field, with the result that large values of magnetic polarization can arise for even small values of applied field. Also, when the applied field is removed, all domains do not revert to their original orientation. In this way all the observed features of ferromagnetism, including that of residual magnetism, can be accounted for.

This concludes our account of the conjectures made by Weiss in his attempt to explain ferromagnetism. So astute was his intuition that subsequent work has not substantially altered his proposed mechanism, but has simply provided a theoretical basis for it.

Physical Basis for Spontaneous Magnetization

Let us first briefly summarize Weiss' theory of ferromagnetism. In order to rationalize the observed susceptibility versus temperature behavior of paramagnetic solids, he postulated an interaction between the elementary dipoles of the material. The cause of this interaction was not known. The same notion was then applied to materials with high Curie temperatures, such as iron, cobalt, and nickel, and led to the prediction that such materials should display spontaneous magnetization. This spontaneous magnetization was confined to small, randomly oriented *domains*, each of which was completely magnetized and could be aligned under the action of an applied field. Thus, large values of magnetization could be achieved in macroscopic samples for small magnetizing fields. Also, the sample could retain a portion of its magnetization after removal of the applied field.

The explanation of the interaction between neighboring dipoles was provided indirectly by the Pauli exclusion principle. The reader is probably familiar, at least vaguely, with the application of this principle to the electronic levels in an atom. According to it, two electrons in an atom cannot possess the same energy unless their spins are antiparallel. Thus, each filled energy level in an atom contains a *pair* of electrons whose spins—and, consequently, whose magnetic moments—are antiparallel. For an atom whose occupied levels are all filled, the spin magnetic moments completely cancel one another. The exclusion principle also provides a description of the permitted energy levels, including the electron capacity of the various shells and sub-shells.

The electrons in an energy band in a solid are also governed by the exclusion principle in that, for each permitted value of kinetic energy, two spin orientations are possible. Normally, therefore, there are two electrons for each value of kinetic energy. A second factor, not mentioned so far, occasionally alters this state of affairs. This is the electrostatic repulsion between electrons,

an interaction which obviously raises the energy of the electrons involved and is greater when they are closer together. Since, according to the exclusion principle, two electrons with antiparallel spins can occupy the same energy level, whereas those with parallel spins cannot, it immediately follows that the mean electrostatic energy between electrons with parallel spins is less than it is between those whose spins are antiparallel. If an electron moves from a filled to an unfilled level, where its spin can now assume a direction parallel to that of its former partner, it must acquire additional kinetic energy, but by so doing it may lose electron-interaction energy of greater amount, with the result that its *total* energy is actually lowered. In this event the parallel-spin arrangement will be energetically favored over the antiparallel one. This condition actually exists in one of the electronic bands (not the valence band) of iron, cobalt, and nickel. The density of levels within the band is so high that the negligible gain in kinetic energy required for an electron to move from a filled level to an unfilled one is more than offset by the lowering of the electrostatic energy which accompanies the shift from an antiparallel to a parallel orientation. This type of spin-dependent electrostatic force [11] is responsible for the alignment of electron-spin magnetic moments that constitutes the spontaneous magnetization existing within the domains of ferromagnetic materials.

Properties of Domains

There is ample evidence for the existence of regions in a magnetic material within which the elementary dipoles lie parallel to one another. The boundaries of these domains can be visualized by allowing a colloidal suspension of iron particles to settle upon the smooth surface of a ferromagnetic specimen. At the domain boundaries there is sufficient flux leakage to attract the suspended particles which, after deposition, provide a striking mosaic of the domain structure.

From these and other studies it is known that the magnetization within a domain is parallel to a direction of easy magnetization. In iron, for example, there are six such directions (corresponding to the three crystal axes), and normal unmagnetized iron has one-sixth of its domains oriented in each of these directions.

The domain boundaries are approximately 1000 atomic diameters thick while the domains themselves have dimensions of the order of a tenth of a millimeter.

[11] Spin-dependent electrostatic forces are generically referred to as *exchange forces*. Such a force exists between two atoms that share a pair of electrons. A case in point is the hydrogen molecule. Here the two electrons can be pictured as belonging to both of the constituent atoms and "exchanging" periodically between them. Another example is the pair of shared electrons in the HCl molecule, considered in § 4.1. In these two instances the energetically favored spin state is the antiparallel, rather than the parallel, configuration.

When a weak magnetizing field is applied, the magnetic polarization of the sample changes as a result of *motion of domain walls.* In this process domains whose axes are nearest to the direction of the applied field grow at the expense of other domains, as shown in Figure 10.17a–c. When this process is complete,

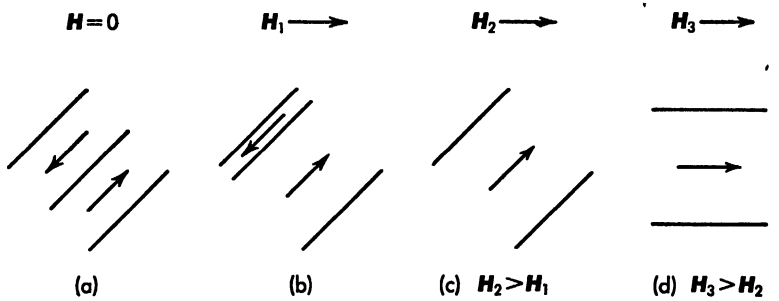


Fig. 10.17

as in (c), the application of a stronger field causes *domain rotation,* during which the magnetic polarization assumes a direction parallel to the applied field, as in (d). Motion of the domain wall accounts for the portion of the magnetization curve below the knee of the curve (0 to A in Fig. 10.18), while domain rotation accounts for a portion above the knee A to C in Fig. 10.18.

Domain motion is a random and discontinuous affair. This can be demonstrated by wrapping the sample being magnetized with a coil of wire in which emf's can be induced whenever $\boldsymbol{B}$ changes within the sample. When these emf's are amplified they can be made audible by use of a microphone. In this way the sporadic nature of part, at least, of the magnetization process can be convincingly shown. This is known as the *Barkhausen effect* (1919).

Magnetostriction

Magnetization of a ferromagnetic material is accompanied by a small change in its length, an effect known as *magnetostriction.* As a result of the motion of domains, the crystal lattice becomes strained, either positively or negatively, in the direction of magnetization. For cobalt the change in length is negative, for iron it is positive at first and then negative for higher values of $\boldsymbol{H}$.

If the magnetostriction is positive, then application of tension to the material will cause the magnetization to prefer the direction of tension, and the reverse for materials with negative magnetostriction. This *structure-sensitive* aspect of magnetism has its analogue in ferroelectricity, where the corresponding macroscopic phenomenon is piezoelectricity.

10.18 The Magnetization Curve

The behavior of a ferromagnetic material during magnetization is conventionally presented as a ***B-H*** curve, similar to the ***D-E*** curve that was briefly touched upon in § 4.13. As we have seen, the ***B-H*** relationship can be expressed in either of two ways:

$$\boldsymbol{B} = \mu \boldsymbol{H} \qquad [8\text{-}11]$$

or

$$\boldsymbol{B} = \mu_0 \boldsymbol{H} + \mu_0 \boldsymbol{M} \qquad [10\text{-}69]$$

Since the magnetizing force is the cause, and the induction is the effect, $\boldsymbol{H}$ is the independent variable in a ***B-H*** curve.

The origin of coordinates in Figure 10.18 represents the unmagnetized condition of a ferromagnetic sample. When a magnetizing force is applied to the sample the resulting magnetic polarization does not increase proportionately, but tends to a saturation value. This process is shown in Figure 10.18 by the solid line, which is known as the *magnetization curve.*

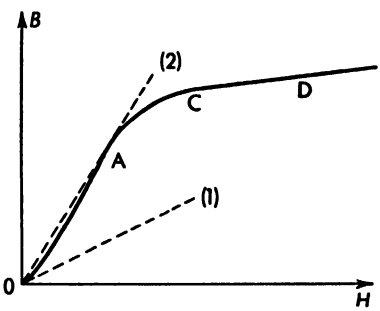


Fig. 10.18

In keeping with the facts presented in the preceding section, the magnetization curve typically consists of three portions:

OA—during this stage of the magnetization, as a result of the motion of domain boundaries, the elementary dipoles assume the direction of easy magnetization which is closest to the direction of the applied field.

AC—during this second stage the applied field is sufficiently large to cause the domain magnetization to rotate from the direction of the preferred crystal axis to that of $\boldsymbol{H}$. At *C* the elementary dipoles are all parallel to $\boldsymbol{H}$. At this point the term $\mu_0 \boldsymbol{M}$ in equation [10-69] has reached its saturation value $\mu_0 \boldsymbol{M}_0$.

CD—as can be seen from equation [10-69], once saturation has occurred, the only gain in $\boldsymbol{B}$ as $\boldsymbol{H}$ is further increased is that due to the increase in $\boldsymbol{H}$ alone. Thus, the curve *CD* is a straight line of slope

$$\frac{\Delta B}{\Delta H} = \mu_0 \qquad [10\text{-}93]$$

From point *C* on is what might be termed the "brute force" region for the achievement of magnetic induction. Every increase in $\boldsymbol{B}$ is gained by increasing

$\boldsymbol{H}$—beyond C the ferromagnetic material itself makes zero additional contribution to the cause.

Shown in Table 6 are illustrative data for a typical sample of Armco iron. This is a very pure iron which is widely used for electromagnets. The saturation of $\boldsymbol{M}$ with increasing $\boldsymbol{H}$ can be seen in the final column.

TABLE 6

B-H Data for Armco Iron

H (ampere-turns/m)	B (webers/m^2)	$k_m = \mu/\mu_0$	M (ampere-turns/m) $\times 10^{-5}$
80	0.415	4150	3.30
120	0.760	5070	6.05
192	1.095	4560	8.73
288	1.285	3570	10.20
800	1.520	1520	12.10
8,000	1.810	181	14.40
80,000	2.250	22.5	17.10
160,000	2.350	11.8	17.10

Initial and Maximum Permeabilities

Two important properties of a ferromagnetic substance are its *initial permeability* and its *maximum permeability*. The former is given in Figure 10.18 by the slope of line (*1*), and the latter by the slope of line (*2*).

Materials with high initial permeability are useful in transformers that transmit the small signals that are encountered in radio communication. In this type of application the maximum induction is often much less than 0.01 webers/m^2. *Permalloy*, an iron-nickel alloy, with initial relative permeability $\sim 10^4$, is a well-known example of this type of material. With certain other recently developed alloys (e.g., *supermalloy*), the initial relative permeability can exceed 10^5. As one might predict, extraordinarily high values of this property have been determined for *single crystals* of iron by employing a magnetizing field that is parallel to the crystal axis.

On the other hand, the materials comprising transformers that are used for power transmission should have large values of relative permeability for *non*vanishing values of H. For these materials the value of maximum relative permeability is the relevant datum: it usually lies in the range 5000–9000.

10.19 Hysteresis and the Hysteresis Loop

Our interest now shifts to the behavior of a ferromagnetic material that has already been magnetized. This behavior is shown in Figure 10.19, which also includes, for reference, the presaturation portion of the magnetization curve.

Three features characterize this behavior: (a) When the magnetizing force is reduced, the magnetic induction does not decrease along the magnetization curve, but displays higher values. When $\boldsymbol{H}$ has been reduced to zero, $\boldsymbol{B}$ still has a positive value, $\boldsymbol{B}_r$, known as the *remanence*. This phenomenon is known as hysteresis. (b) The magnetizing force must be made negative, to a value $\boldsymbol{H}_c$, known as the *coercive force*, in order to bring $\boldsymbol{B}$ back to zero. (c) When $\boldsymbol{H}$ is varied periodically about the origin, the $\boldsymbol{B}$-$\boldsymbol{H}$ curve follows a closed contour, known as the *hysteresis loop*. This description corresponds exactly to that given in § 4.13 for ferroelectric substances.

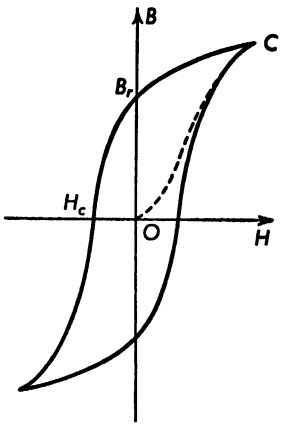


Fig. 10.19

Significance of the Area of the Hysteresis Curve

We shall now see that an expenditure of energy is called for whenever a sample of material is taken through the cycle of operations represented by the hysteresis loop. This energy represents the work done by the current (in the surrounding coil) in overcoming the back emf that is induced in the coil whenever B changes within the sample.

Let us assume that the sample under study completely fills the region within a toroidal solenoid. The flux within the coil is given at any instant by

$$\Phi_m = BA \text{ webers} \qquad [10\text{-}94]$$

where A is the cross-sectional area of the toroid. The instantaneous back emf is given by

$$\varepsilon_B = -N\frac{d\Phi_m}{dt} = -NA\frac{dB}{dt} \qquad [10\text{-}95]$$

where N is the total number of turns on the toroidal coil.

For current to exist, the circuit must contain an applied emf, ε, which is equal and opposite to this back emf: that is, $\varepsilon = -\varepsilon_B$. As pointed out in the

footnote to equation [9-6], the applied emf does work in time dt of amount

$$dW' = \varepsilon i\, dt = -\varepsilon_B i\, dt \qquad [10\text{-}96]$$

Upon substitution of equation [10-95] and $nl = N$, equation [10-96] becomes

$$dW' = NA\frac{dB}{dt} i\, dt = ni\, dB\,(Al) = H\, dB\,(Al) \qquad [10\text{-}97]$$

where Al, the interior volume of the coil, is the volume of the sample. Hence, per unit volume of material, the work done against the back emf when the induction changes from B to $B + dB$ is given by

$$dW/\text{unit volume} = H\, dB \qquad [10\text{-}98]$$

If the unit volume of the sample is carried completely around the hysteresis loop, the work done (and, hence, the energy dissipated) in the process is

$$W/\text{unit volume} = \oint H\, dB \qquad [10\text{-}99]$$

The integral on the right-hand side of equation [10-99] is simply the area of the hysteresis loop. The hysteresis loop thus has peculiar significance. The fact that there is a loop is evidence that work is done in taking the sample through the hysteresis cycle. The energy expended in the process appears as heat in the sample, providing evidence that the magnetization process is not a reversible one but that internal friction is involved. Further, the actual area of the loop provides a value for the amount of energy that is dissipated when unit volume of the material is carried once around the loop.

If the cycle is repeated f times per second, as happens in a transformer, the energy lost per second is

$$\text{Energy lost/sec-m}^3 = f \times (\text{area of hysteresis loop}) \qquad [10\text{-}100]$$

It is therefore essential that transformers be constructed of materials with narrow hysteresis loops. Most commercial power transformers are made of an iron-silicon alloy which has several advantages over ordinary iron: (a) the hysteresis loss is less by reason of its narrower hysteresis loop, (b) the permeability is greater, thereby necessitating smaller exciting currents, and (c) the resistivity is greater, with the result that eddy-current losses are smaller.

Ferrites

Although ferromagnetism is an uncommon phenomenon among elements, it is quite common among compounds, particularly among complex oxides of the so-called *transition group* of elements. Elements in this group are scandium

($Z = 21$), titanium ($Z = 22$), vanadium ($Z = 23$), chromium ($Z = 24$), manganese ($Z = 25$), iron ($Z = 26$), cobalt ($Z = 27$), nickel ($Z = 28$), copper ($Z = 29$) and zinc ($Z = 30$). Typical of such complex oxides are $CrFe_2O_4$, $MnCr_2O_4$ and $MnFeCrO_4$. The outstanding advantage which these materials offer, in addition to their ferromagnetism, is the fact that they are insulators, or sometimes semiconductors. Consequently, when they are used as transformer cores, there is no eddy-current loss.

Permanent Magnets

In a discussion of permanent magnets the portion of the hysteresis loop that is of interest is the part that lies in the second quadrant. Let us suppose that a sample of material has been magnetized to saturation, after which H has been reduced. The portion of the hysteresis loop that lies in the second quadrant, as in Figure 10.20, is known as the *demagnetization curve.* As a general rule it is desirable that the remanence, B_r, of permanent-magnet materials be large, in order that the greater part of the magnetization will remain when the magnetizing force is removed. Also, it is desirable that the coercive force, H_c, be large, in order that the magnet will not easily be demagnetized.

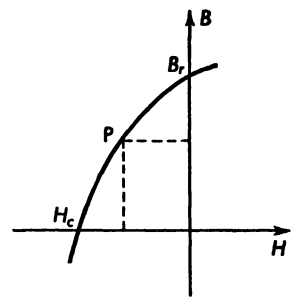


Fig. 10.20

The best single figure of merit of a permanent-magnet material is the maximum value of the product BH along the demagnetization curve. Since the product BH has the dimensions of energy density, it is referred to as the *energy product*, and its maximum value as the *maximum energy product.* The best commercially available permanent-magnet material is *Alnico 5*, a cast alloy containing iron, cobalt, nickel, aluminum, and copper. It has a BH_{max} value of 36,000 joules/m^3. Compacted powders of Fe-Co have yielded similar maximum energy products. Carbon steel, on the other hand, which was for a long period of time a standard material for permanent magnets, has a maximum energy product of only about 1450 joules/m^3.

If the sample of magnetic material under study is in the form of a closed ring, the magnetizing force H is entirely provided by the existing current in the encompassing toroidal solenoid. If, however, the sample is in the form of a rod, the poles that appear at the ends of the rod represent a demagnetizing force. The situation is similar to the electrical case represented in Figure 4.6. The externally applied electric field corresponds to the magnetizing force associated with the current in the exciting coil, while the polarization charges, which oppose

the externally applied field, correspond to the magnetic poles which appear at the ends of the rod-shaped magnet. When the current excitation ceases, these poles remain (unlike the usual electric polarization charges) and, consequently, $\boldsymbol{H}$ becomes somewhat negative. Thus, permanent magnets operate in the second quadrant, that is, $\boldsymbol{B}$ and $\boldsymbol{H}$ are opposite in direction.

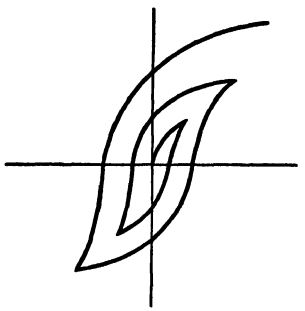


Fig. 10.21

Demagnetization

It is frequently necessary to demagnetize an object, for example, a wrist watch that has been inadvertently brought too close to a magnet. This is most simply done by carrying the object several times around its hysteresis loop (by alternating the sign of $\boldsymbol{H}$), reducing the amplitude of $\boldsymbol{H}$ at each reversal. In this way the object can quickly be brought to a condition in which $\boldsymbol{B}$ is zero for $\boldsymbol{H} = 0$, as suggested in Figure 10.21. In practice, the object can be placed inside a coil carrying an AC current which is then reduced to zero.

10.20 Experimental Tracing of Hysteresis Loops

The best-known, and perhaps the most instructive, method for obtaining the hysteresis loop of a sample of ferromagnetic material is that which employs the *Rowland ring*. The method is indicated in Figure 10.22.

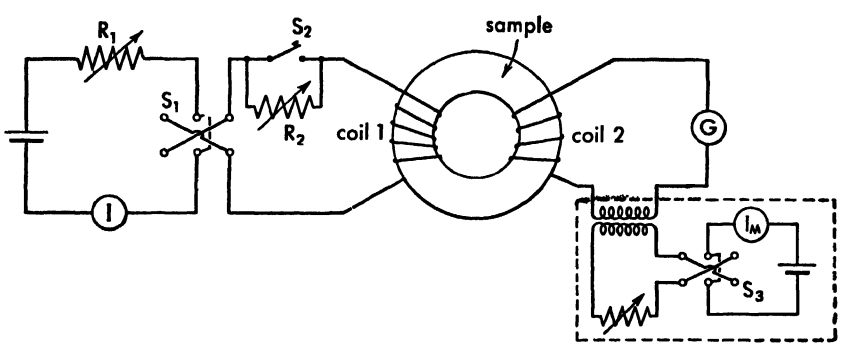


Fig. 10.22

The material to be studied constitutes a torus around which two coils are wound, as shown in the figure. The magnetizing force is applied to the sample

by passing a current through coil *1*. The consequent change in flux through the sample induces an emf in coil *2*, in accordance with Faraday's law of electromagnetic induction. The emf gives rise to a current which causes the galvanometer to deflect. As we shall see, the deflection of the galvanometer is actually proportional to the change in $\boldsymbol{B}$ that results from the application of $\boldsymbol{H}$. The portion of the apparatus shown within the dotted-line box may be ignored until the sub-section, "Calibration of the Ballistic Galvanometer," which concludes this section.

Digression Concerning Ballistic Galvanometers

The galvanometer used in this experiment is a *ballistic* galvanometer. This instrument will be discussed in the next chapter but we should note, in the meantime, that for a current of short duration the deflection of the galvanometer is proportional to the total charge passing through it.

At any instant the induced emf has the value

$$\varepsilon = -N\frac{d\Phi_m}{dt} \qquad [9\text{-}5]$$

If the resistance of the galvanometer circuit is R ohms, the instantaneous current passing through the galvanometer is

$$i = \frac{|\varepsilon|}{R} = \frac{N}{R}\frac{d\Phi_m}{dt} = \frac{NA}{R}\frac{dB}{dt} \qquad [10\text{-}101]$$

The charge carried by this current in time dt is

$$dq = i\,dt = \frac{NA}{R}dB \qquad [10\text{-}102]$$

and the total charge carried during the entire (short) duration of the current is given by

$$Q = \int_0^Q dq = \frac{NA}{R}\int_0^{\Delta B} dB = \frac{NA}{R}\Delta B \text{ coulombs} \qquad [10\text{-}103]$$

Since the galvanometer deflection, d divisions, is proportional to the total charge passing, we have

$$d = K\left[\frac{NA}{R}(\Delta B)\right] \text{ scale divisions} \qquad [10\text{-}104]$$

relating the deflection to the change in flux density within the magnetic sample. K, *the ballistic sensitivity of the galvanometer*, is the reciprocal of the number of coulombs whose passage gives rise to a deflection of one scale division. A method by which it can be determined is described below. It will be realized

that the galvanometer deflection in this case is not a steady one but is, rather, a "throw." The magnitude of this throw can be observed without difficulty. With this background, let us continue the discussion of the Rowland ring experiment.

Let us assume, to begin with, that switch S_2 in Figure 10.22 is closed. Switch S_1 is a reversing switch, by means of which the direction of the current through coil *1* can be changed. In this way the direction of $\boldsymbol{H}$ in the torus can be reversed. With S_1 closed in one direction, the magnitude of the current I is adjusted (by means of R_1) to such a value that, when S_1 is thrown to the reverse position, the galvanometer throw is nearly full-scale. When S_1 is thrown back to its original position, the galvanometer throw has the same value but is in the opposite direction. As we have said above, this throw is a measure of the change in $\boldsymbol{B}$ that results from changing $\boldsymbol{H}$. Each time S_1 is reversed the sample moves from point A to point C (Fig. 10.23) on the hysteresis loop, or vice versa. Thus, a change in $\boldsymbol{H}$ of amount $2H_{\max}$ causes a change in $\boldsymbol{B}$ of amount $2B_{\max}$.

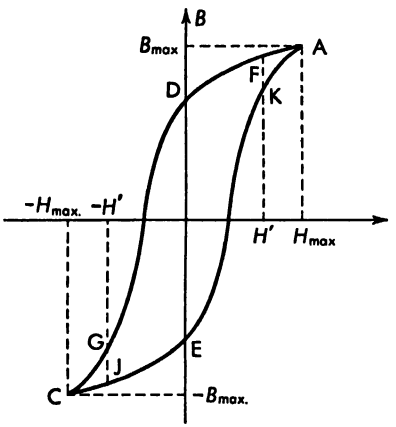


Fig. 10.23

If the sample is at point A, and S_1 is then simply opened, rather than thrown in the opposite direction, the sample moves to point D, because the magnetizing force has dropped to zero. The galvanometer deflection is much smaller than before and its magnitude indicates the position of D relative to $B_{\max}$. Closing S_1 in the opposite direction moves the sample from D to C with a consequently larger galvanometer deflection. Evidently, the *sum* of these two latter deflections is equal to the deflection observed when S_1 is simply reversed. Point E on the loop is located by a procedure similar to that used in locating point D.

Switch S_2, with the parallel resistance R_2, provides a convenient arrangement for locating other points on the hysteresis loop. If the sample is at point A, the opening of S_2 causes the current to drop from $I_{\max}$ to I', representing a change in $\boldsymbol{H}$ from $\boldsymbol{H}_{\max}$ to $\boldsymbol{H}'$. The sample moves from A to F on the loop. Information concerning the magnitude of this vertical displacement, as before, is provided by the galvanometer throw. If S_1 is now reversed, the galvanometer deflection corresponds to the change in B as the sample moves from point F to point G. By closing S_2, the sample is carried to point C, from which, by a similar technique, points J and K can be located. Finally, by varying R_2, other

intermediate points on the loop can be located, as desired. It should be borne in mind that no such excursion, undertaken with a view to ascertaining the position of certain intermediate points, is complete unless it begins at A (or C) and ends at C (or A). Otherwise, the excursionist will find himself upon a hysteresis loop that is different from the original.

The magnetization curve can be obtained by the same general method. It is essential in this case (and, indeed, wise in the experiment above) to demagnetize the specimen first. This is done by reversing the current through the coil many times, at the same time steadily reducing it to practically zero. The sample is then in the condition represented by the origin in Figure 10.18. If $\boldsymbol{H}$ is now increased in steps, the galvanometer deflections indicate the corresponding increases in $\boldsymbol{B}$. In this sequence of operations the direction of $\boldsymbol{H}$ must on no account be changed.

Calibration of the Ballistic Galvanometer

There is no difficulty in calculating the absolute values of $\boldsymbol{H}$ in the above experiments: they are given by nI or $NI/2\pi R$, where R is the mean radius of the toroidal specimen. To obtain the absolute values of $\boldsymbol{B}$ it is necessary to know the sensitivity to passage of charge of the ballistic galvanometer. In this experiment a convenient method of determining the sensitivity is to introduce into the galvanometer circuit, in series with coil *2*, one of the coils of a standard mutual inductance (§ 9.7), as shown in the dotted-line box in Figure 10.22.

When the current I_m in the primary coil of the mutual inductance is changed, an emf is induced in the secondary coil of amount

$$\mathcal{E}_2 = -M\frac{dI_M}{dt} \qquad [9\text{-}23]$$

The instantaneous current flowing in the galvanometer circuit (resistance R ohms) is thus

$$i_2 = \frac{\mathcal{E}_2}{R} = -\frac{M}{R}\frac{dI_M}{dt} \qquad [10\text{-}105]$$

When the primary current is changed from $+I_M$ to $-I_M$ by throwing the reversing switch S_3, the total charge passing through the galvanometer is given by

$$Q_M = \int i_2\,dt = -\frac{M}{R}\int_{I_M}^{-I_M} dI_M = \frac{2MI_M}{R} \qquad [10\text{-}106]$$

and the corresponding deflection by

$$d_M = KQ_M = K\left[\frac{2MI_M}{R}\right] \text{ scale divisions} \qquad [10\text{-}107]$$

Evidently, by observing the deflection that occurs for a reversal of the primary current I_M, the student may ascertain the value of K, the ballistic sensitivity of the galvanometer. The value of the mutual inductance is, of course, known.

Since the secondary coil of the mutual inductance is a series element in the galvanometer circuit throughout the entire experiment, the resistance, R, appearing in equation [10-107] is identical with that appearing in equation [10-104], and the quantity K/R is obtained during calibration, rather than K only.

10.21 The Magnetic Circuit

Kirchhoff's second law (equation [5-49]) relates current, emf, and resistance in an electric circuit. Quantities analogous to these exist in a magnetic "circuit" and are related through a similar type of law. In this section we shall develop this law. Its utility in the design of electromagnets will be demonstrated in the following section.

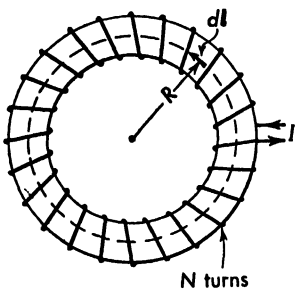


Fig. 10.24

As an example of a magnetic circuit let us consider the arrangement shown in Figure 10.24. Here a ring of iron, of uniform permeability, μ, is surrounded by a close-fitting toroidal coil of N turns carrying a current of I amperes. The cross-sectional area of the iron is A m^2.

In such a circuit the flux, Φ_m, is confined to the iron ring and has the same value through each cross-sectional area of the ring. In these respects it resembles an electric current that is confined to a closed conducting loop and has the same value at every point around the loop. Thus, in drawing a parallel between magnetic and electric circuits, the flux is taken as the magnetic counterpart of the electric current.

The magnetic analogue to the electromotive force is the *magnetomotive force*, abbreviated *mmf*. The emf in a circuit, it will be recalled, is given by the line integral of the electric intensity around the circuit, viz.,

$$\text{emf} = \oint \boldsymbol{E} \cdot d\boldsymbol{l} \qquad [9\text{-}15]$$

In keeping with this relation, and also with our interpretation of $\boldsymbol{H}$ as a magnetizing force, we define the magnetomotive force as the integral of $\boldsymbol{H}$ around the magnetic circuit:

$$\text{mmf} = \oint \boldsymbol{H} \cdot d\boldsymbol{l} \qquad [10\text{-}108]$$

where the value of the right-hand side of the equation is known from equation

[8-40] to be I for each turn of wire that is circumnavigated. The magnetomotive force for the circuit of Figure 10.24 is, therefore,

$$\text{mmf} = NI \qquad [10\text{-}109]$$

Let us evaluate, in terms of other symbols, the line integral of $\boldsymbol{H}$ around the dotted path shown in Figure 10.24. At each point $\boldsymbol{H}$ and $\boldsymbol{dl}$ are parallel and, in addition,

$$H = \frac{B}{\mu} = \frac{\Phi_m}{\mu A} \qquad [10\text{-}110]$$

The line integral can, therefore, be written:

$$\oint \boldsymbol{H} \cdot d\boldsymbol{l} = \oint H\, dl = \Phi_m \oint \frac{dl}{\mu A} \qquad [10\text{-}111]$$

Substituting equation [10-111] in equation [10-108] we have

$$\text{mmf} = \Phi_m \oint \frac{dl}{\mu A} \qquad [10\text{-}112]$$

which, rearranged, becomes

$$\Phi_m = \frac{\text{mmf}}{\oint \frac{dl}{\mu A}} \qquad [10\text{-}113]$$

When this is compared to Kirchhoff's second law ($I = \text{emf}/R$), it is seen that the quantity in the denominator of the right-hand side has the character of magnetic "resistance." It is known as the *reluctance* of the magnetic circuit and is designated by the symbol $\mathcal{R}$: thus,

$$\mathcal{R} = \oint \frac{dl}{\mu A} \qquad [10\text{-}114]$$

For an element of the circuit

$$d\mathcal{R} = \frac{dl}{\mu A} \qquad [10\text{-}115]$$

The parallel that we have been drawing between electric and magnetic circuits is further confirmed by the striking resemblance between equation [10-115] and the corresponding expression for an element of resistance

$$dR = \frac{dl}{\sigma A} \qquad [6\text{-}3]$$

where σ is the conductivity of the material.

In the circuit of Figure 10.24, where μ and A are constant, the reluctance has the simple value

$$\mathcal{R} = \frac{1}{\mu A} \oint dl = \frac{2\pi R}{\mu A} \qquad [10\text{-}116]$$

When the circuit is not a uniform one, but consists of a series combination of two or more dissimilar portions, the total reluctance of the circuit is the sum of the reluctances of the separate portions. Reluctances in parallel are also combined in accordance with the rules for resistances. Thus, in series,

$$\mathcal{R} = \sum_{1}^{n} \mathcal{R}_i \qquad [10\text{-}117]$$

or in parallel,

$$\frac{1}{\mathcal{R}} = \sum_{1}^{n} \frac{1}{\mathcal{R}_i} \qquad [10\text{-}118]$$

In summary, we note that the reluctance of a magnetic circuit can be computed from a knowledge of the geometry and the magnetic properties of the components of the circuit. Knowing the reluctance, we can compute the total flux that will be established in the circuit when a given current flows through the exciting coil. For this purpose we use

$$\Phi_m = \frac{\text{mmf}}{\mathcal{R}} = \frac{NI}{\mathcal{R}} \qquad [10\text{-}119]$$

a relationship which is the working equation in the design of electromagnets.

10.22 Electromagnets

In the apparatus shown in Figure 10.24 the magnetic field is confined to the iron ring. If a section of the ring is removed, as in Figure 10.25, thereby introducing an air gap into the magnetic circuit, the magnetic field in the gap may be used to exert forces upon electric currents, to deflect beams of charged particles, or to perform the other functions required of an electromagnet. The iron faces that bound the air gap constitute the *poles* of the electromagnet. Throughout the iron portion of the circuit the flux is confined to the iron itself, but in the air gap it spreads somewhat, occupying a cross-sectional area, A', which is greater than the cross-sectional area, A, of the iron. But in each case the flux has the same value.

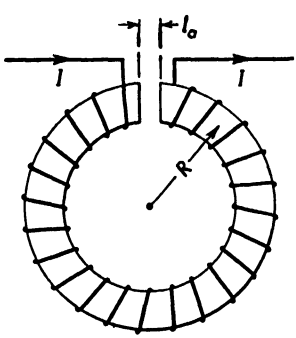


Fig. 10.25

The reluctance of the circuit is given by

$$\mathcal{R} = \mathcal{R}_{\text{iron}} + \mathcal{R}_{\text{air gap}} \qquad [10\text{-}120]$$

or
$$\mathcal{R} = \frac{2\pi R - l_a}{\mu A} + \frac{l_a}{\mu_0 A'} \qquad [10\text{-}121]$$

The flux throughout the circuit is obtained using equation [10-119]:

$$\Phi_m = \frac{NI}{\mathcal{R}} = \frac{NI}{\left[\dfrac{2\pi R - l_a}{\mu A} + \dfrac{l_a}{\mu_0 A'}\right]} \qquad [10\text{-}122]$$

The flux density in the air gap is then given by

$$B_{\text{air gap}} = \frac{\Phi_m}{A'} \qquad [10\text{-}123]$$

In an electromagnet whose air gap is small compared to the diameter of the pole faces, the spreading of the flux in the air gap is negligible, in which case $A' \doteq A$ and the flux density at any point in the magnetic circuit is given by

$$B \doteq \frac{NI}{\left[\dfrac{2\pi R - l_a}{\mu} + \dfrac{l_a}{\mu_0}\right]} \qquad [10\text{-}124]$$

Since μ, the permeability of the iron, is normally about three orders of magnitude greater than μ_0, the permeability of free space, the reluctance of the circuit of Figure 10.25 is largely provided by the air gap, unless the air gap is very narrow.

10.23 Problems

F (10.1) One horizontal rod-shaped magnet is suspended directly above a second horizontal magnet of the same size and pole strength by the repulsion of like poles. Each magnet has a mass of M kilograms and is l meters long. Deduce an expression relating their pole strength to the distance between their axes.

$$\left[\text{Ans.: } Mg = \frac{\mu_0 q_m^2}{2\pi}\left[\frac{1}{d^2} - \frac{d}{(d^2 + l^2)^{3/2}}\right]\right]$$

E (10.2) A uniformly magnetized rod magnet is broken into two parts, one of which is one-third as long as the other. The two pieces are then fastened, with centers together, at right angles to one another. When the combination is floated on a cork, what angle does the longer magnet make with the horizontal component of the magnetic field of the earth? [Ans.: approx. 18.5°.]

M (10.3) A compass needle, C, is situated between two *short* magnets, X and Y, as shown in Figure 10.26. The magnetic moments of the two

magnets are the same, and $\theta = 30°$. Calculate the ratio r_2/r_1. r_1 and r_2 are both large in comparison to the lengths of the magnets. [Ans.: $r_2/r_1 = 1.51$.]

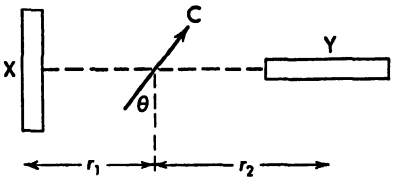


Fig. 10.26

M (10.4) Two identical short magnets, X and Y, are floated on a cork as shown in Figure 10.27. The center of Y is located on the axis of X, and the axes of the two magnets are normal to one another. Compare the torque exerted by X upon Y with that exerted by Y upon X. Is there any torque tending to rotate the cork? Assume $r \gg l$.

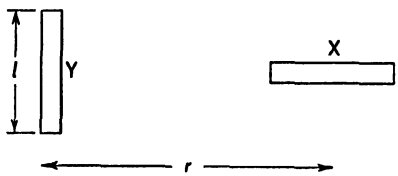


Fig. 10.27

E (10.5) Show that the units of $\boldsymbol{m}$ are $l^2t^{-1}Q$.

M (10.6) A bar magnet of magnetic moment m and moment of inertia $\mathcal{I}$ is suspended from its center by a thin torsionless thread and allowed to take up its equilibrium position with respect to the horizontal component of the earth's magnetic field. If the magnet is then slightly twisted in the horizontal plane from its equilibrium position, show that it executes simple harmonic motion.

M (10.7) $\boldsymbol{r}$ is the vector distance from the center of a circular current I, of radius a, to a point P. $\boldsymbol{r}$ makes an angle θ with the plane of the circular current. Assuming $r \gg a$, show that the r-component of the induction at P arising from the circular current is given by

$$B_r = \frac{\mu I a^2 \sin\theta}{2r^3}$$

H (10.8) A spherical conductor of radius a meters and with total charge Q coulombs rotates with uniform angular velocity ω. Compute the resulting magnetic moment. [Ans.: $\omega Qa^2/3$.]

M (10.9) A close-wound *flat* spiral of outside radius R is comprised of a large number (N) of evenly spaced turns of fine wire. If the wire carries a current of I amperes, what is the magnetic moment of the spiral? [Ans.: $\pi NIR^2/3$.]

E (10.10) (Based on small-type material in § 10.6.) Using expression [10-27] for the magnetic scalar potential, show that V_m changes by an

amount μI as the point P moves from $+\infty$ to $-\infty$, passing through a closed current en route. Is this consistent with the circuital form of Ampère's law (equation [8-39])?

M (10.11) (Based on small-type material in § 10.6.) Use the concept of the magnetic shell to calculate the induction at a point P on the axis of a circular current. Hint: since you are free to choose any shape of shell you wish, choose the one for which P is on the axis of each of the dipoles constituting it.

M (10.12) Show that equation [10-50], giving the force experienced by a magnetic dipole in a nonuniform magnetic field, can be expressed in the vector form

$$\boldsymbol{F} = (\boldsymbol{m}\cdot\nabla)\boldsymbol{B}$$

E (10.13) Prove that $\nabla\cdot(\nabla\times\boldsymbol{A}) = 0$, where $\boldsymbol{A}$ is any vector.

M (10.14) A flat test coil of 100 turns of wire, wound on a circular form of cross-section area 10 cm^2, is connected to a 100-ohm ballistic galvanometer of sensitivity 10^7 cm/coulomb. When the coil is placed in a uniform magnetic field (whose direction is parallel to the axis of the coil) and rotated quickly through 180°, a deflection of 5 cm is observed in the galvanometer. Calculate the value of induction for the field. The coil resistance may be neglected.

[Ans.: 2.5×10^{-4} webers/m^2.]

E (10.15) What is the galvanometer deflection in the preceding problem if the coil is placed with its axis parallel to the magnetic field and is then suddenly withdrawn from the field?

E (10.16) An inductance consists of an air-cored toroidal solenoid containing 20 turns per centimeter. How much larger is the value of the inductance of a similar coil with a core of Armco iron (Table 6). Assume a current of 4 amperes. [Ans.: $\sim$ 180 times.]

E (10.17) Show, by ascertaining the dimensions of H and B, that the product HB has dimensions of energy per unit volume.

E (10.18) An electromagnet is made from a cylindrical piece of iron bent into a complete circuit except for an air gap of 1 cm. If the relative permeability of the iron is 2000, its cross-section area is 20 cm^2, and the mean radius of the circle into which it has been bent is 20 cm, what current in amperes is required in order to establish an induction in the air gap of 0.4 webers/m^2? There are 1000 turns of wire on the magnet. [Ans.: 3.4 amperes.]

M (10.19) Compute the relationship between the oersted and the corresponding mks unit for $\boldsymbol{H}$, the ampere-turn/meter.

[Ans.: One oersted $\doteq$ 79.6 ampere-turns/meter.]

Chapter 11

Galvanometers, Ammeters, Voltmeters, and Related Instruments

11.1 The Moving Coil or d'Arsonval Galvanometer

The galvanometer is an instrument for the detection or measurement of small currents. The most common type is the *moving-coil* or *d'Arsonval* galvanometer, shown semischematically in Figure 11.1.

The current to be measured passes through a rectangular coil, of length l and width w (see left of Fig. 11.1), which is located between the poles of a permanent magnet. The coil is suspended in such a way that it is free to rotate about its long axis and is wound on a soft iron cylinder. This construction, together with the specially shaped poles of the permanent magnet (see right of Fig. 11.1), produces, in the region occupied by the coil, a magnetic field which is radial and uniform over the angular range of the coil. Thus, over a wide angular range, the coil encounters a magnetic field that is both parallel to its face and constant in value. The suspension serves as one lead to the coil, while a loose helical spring serves as the other.[1]

As a result of the force exerted by the field of the magnet upon the current in the coil, the coil experiences a torque which causes it to rotate. This rotation is opposed by a torque exerted by the suspension. As we have seen in § 10.4,

[1] In another type of galvanometer, which will not be described here, the current to be measured flows in a fixed coil. The field produced by this current acts upon a small suspended (or pivoted) permanent magnet, causing it to rotate. The tangent galvanometer (Problem 11.1) is an example of this type.

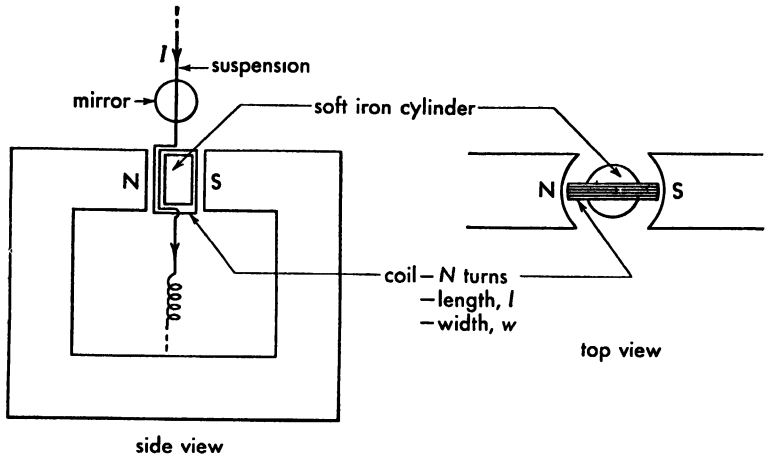


Fig. 11.1

the torque experienced by a dipole in a uniform magnetic field is

$$\boldsymbol{T} = \boldsymbol{m} \times \boldsymbol{B} \qquad [10\text{-}19]$$

When current is present in the coil shown in Figure 11.1, the coil possesses a magnetic moment

$$m = NIA = NIlw \qquad [11\text{-}1]$$

where N is the number of turns in the coil and A $(= lw)$ is its cross-sectional area. Because of the radial magnetic field which exists in the coil region, $\boldsymbol{m}$ and $\boldsymbol{B}$ are normal to one another throughout the range of operation of the galvanometer. Thus, the scalar form of equation [10-19] that is applicable in this case is

$$T = NIlwB \qquad [11\text{-}2]$$

Equilibrium is established when the suspension (torsional constant, τ) exerts an equal and opposite torque. That is, at equilibrium,

$$\tau\theta = NIlwB \qquad [11\text{-}3]$$

where θ is the angular displacement of the coil from its zero position. Thus, the angular displacement resulting from the passage through the galvanometer of a current I is given by

$$\theta = \frac{NIlwB}{\tau} \qquad [11\text{-}4]$$

or

$$\theta = CI \qquad [11\text{-}5]$$

where the value of the constant C is determined by the several parameters of the galvanometer that appear in equation [11-4]. From equation [11-5] we

see that the angular displacement of the galvanometer coil is proportional to the current passing through the galvanometer.

The angular displacement of the coil is not measured directly but, instead, observations are made of a beam of light that is reflected from a mirror attached to the coil suspension (left of Fig. 11.1). When the coil is in its zero position the mirror occupies the position XY (Fig. 11.2), and light originating at A is reflected back on itself. When the coil is rotated through an angle θ, the mirror occupies the position $X'Y'$, and the light beam originating at A is reflected through an angle 2θ, corresponding to a scale reading

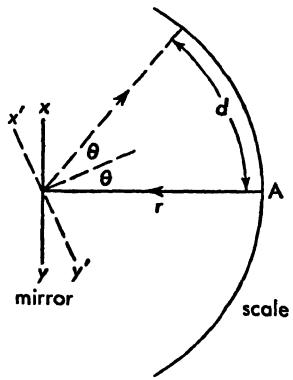


Fig. 11.2

$$d = r(2\theta) \qquad [11\text{-}6]$$

Combining this result with equation [11-5], we have

$$d = 2rCI = C'rI \qquad [11\text{-}7]$$

relating the scale deflection, d, to the galvanometer current, I, and the distance, r, of the scale from the mirror. As with the angular displacement, the scale reading is proportional to the current.

If the galvanometer is to be used for the measurement of current, it is calibrated beforehand by passing certain known currents through it. On the other hand, if it is only to be used for the detection of small currents, as in a Wheatstone bridge circuit, calibration is not necessary. When the direction of the current through the galvanometer changes, the direction of the galvanometer deflection, of course, changes as well.

11.2 Sensitivity of a Galvanometer

There are three commonly employed figures of merit for describing the sensitivity of a galvanometer. These are (a) the *current sensitivity*, (b) the *megohm sensitivity*, and (c) the *voltage sensitivity*. A fourth figure of merit, the *ballistic sensitivity*, which is applicable to certain types of galvanometer, is described in § 11.4.

a. Current Sensitivity: Current sensitivity is the reciprocal of the current required to produce a standard scale deflection.[1] If the galvanometer is equipped with its own scale, the standard deflection is normally taken to be 1 mm on the scale provided. If the galvanometer does not possess a built-in

[1] The current sensitivity is sometimes defined as the current required to produce the standard deflection. This definition has the advantage that it answers the question in the mind of a potential user: "What current is needed to produce a deflection?"

scale arrangement, the standard deflection is taken to be 1 mm on a scale that is located 1 m from the mirror.

From equations [11-4], [11-5], and [11-7] the current sensitivity is seen to be

$$k = \frac{\partial d}{\partial I} = \frac{2rNlwB}{\tau} \qquad [11\text{-}8]$$

Thus, the current sensitivity can be made large by using a thin suspension (small torsional constant), a coil of many turns, and a strong magnet. The coil area (lw) is also a determining factor, but increasing it makes it more difficult to achieve a uniform magnetic field over the required region.

The number of turns on the galvanometer coil is related to the galvanometer resistance in a manner already referred to in § 6.9. For a coil of given size the volume of wire that can be wound into the available space is essentially constant. Hence, the number of turns of wire is inversely proportional to the cross-sectional area, a, of the wire. That is,

$$a = \frac{C''}{N} \qquad [11\text{-}9]$$

The resistance of the coil is given by equation [6-1]:

$$R_G = \frac{\rho \times \text{length of wire}}{\text{cross-sectional area}} = \frac{\rho N(2l + 2w)}{a} = \frac{\rho N^2(2l + 2w)}{C''} \qquad [11\text{-}10]$$

From equation [11-8], we have $k \propto N$, while from equation [11-10] we have $\sqrt{R_G} \propto N$. Thus, the current sensitivity of a moving coil galvanometer (of constant coil size, suspension, and magnetic field, and whose coil is wound with wire of a given resistivity) is directly proportional to the square root of the galvanometer resistance:

$$k \propto \sqrt{R_G} \qquad [11\text{-}11]$$

b. Megohm Sensitivity: Megohm sensitivity is the value of the resistance in megohms which must be placed in series with the galvanometer and an emf of 1 volt in order to secure the standard deflection. The number representing the megohm sensitivity is 10^{-6} times the number representing the current sensitivity.

c. Voltage Sensitivity: Voltage sensitivity is the potential difference across the series combination of the galvanometer and its critical damping resistance (see § 11.3) which produces the standard deflection. It is the product of the reciprocal of the current sensitivity and the combined resistance of the galvanometer and its critical damping resistance.

A few representative values of these figures of merit now follow. For the most sensitive galvanometers, the current sensitivity is $\sim 10^{11}$ mm (at 1 m)

per ampere; the megohm sensitivity is $\sim 10^5$ megohms (for 1 mm deflection at 1 m), and the voltage sensitivity is $\sim 10^{-7}$ volts (for 1 mm deflection at 1 m). For the familiar wall-type galvanometer, whose scale is 50 cm from the mirror, the corresponding values are $5 \times 10^9 - 5 \times 10^8$ mm/ampere, 5000–500 megohms, and $\sim 1.5 \times 10^{-6}$ volts. A much-used galvanometer is the box type. This contains its own lamp and scale and, by virtue of a taut suspension, is portable and does not require careful leveling. The current sensitivity is 4×10^7–2.5×10^7 mm/ampere, the megohm sensitivity is 40–25 megohms, and the voltage sensitivity is $\sim 2.5 \times 10^{-5}$ volts. In low-sensitivity moving-coil galvanometers the angular displacement of the coil is indicated by a pointer attached to the coil. The maximum current sensitivity in this type of instrument is $\sim 10^6$ mm/ampere.

The "Megger"

An interesting embodiment of the moving-coil galvanometer is the "megger," an instrument used for the measurement of insulation resistance. The instrument is shown schematically (from the top) in Figure 11.3. Two coils, Y and Z, at right angles to one another, are located between the poles of a permanent magnet.[3] The hand generator provides an emf which sends current through the two parallel circuits: Coil Z and the known resistance R, and Coil Y and the unknown resistance X. The coils are so wound that the torques which they experience, when carrying current, are in opposite directions. In equilibrium, these torques are equal in magnitude. Thus, assuming that the coils are equal in size and contain equal turns, we have

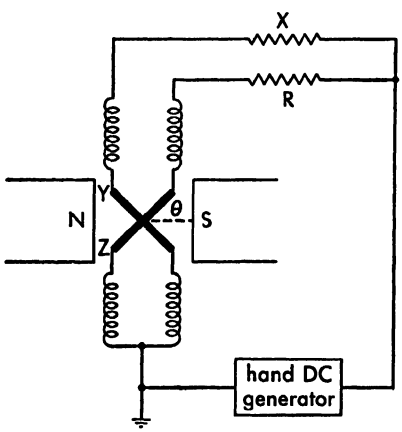


Fig. 11.3

$$|\boldsymbol{m}_Y \times \boldsymbol{B}| = |\boldsymbol{m}_Z \times \boldsymbol{B}| \qquad [11\text{-}12]$$

or
$$NI_Y AB \sin\theta = NI_Z AB \cos\theta \qquad [11\text{-}13]$$

or
$$\frac{\mathcal{E}}{X}\sin\theta = \frac{\mathcal{E}}{R}\cos\theta \qquad [11\text{-}14]$$

or
$$X = R\tan\theta \qquad [11\text{-}15]$$

[3] This differs from the moving coil instrument shown in Figure 11.1 in that the coils are not wound on an iron core and the poles are not curved. Hence, in this case, $\boldsymbol{B}$ is unidirectional rather than radial.

In this way the value of the unknown high resistance is determined. The value of θ is given by a pointer attached to the two-coil assembly.

The connections to the coils exert no torsion, with the result that, when the hand generator is not operated, the pointer needle flops about when the instrument is jarred—no scale position is particularly sacred to it. Once, however, the hand generator is operated, the coils are subject to the opposing torques represented by equation [11-13], and a unique pointer position exists for each value of X. Typical meggers apply 500 volts across the unknown resistance, and can be used to detect insulation resistances of $\sim 10^{10}$ ohms.

11.3 Galvanometer Damping

A moving-coil galvanometer deflects in response to a torque which exists as long as charge is passing through the galvanometer and then returns to its equilibrium position when the flow of charge ceases. For reasons which we shall now discuss, the return to zero may be swift or dilatory.

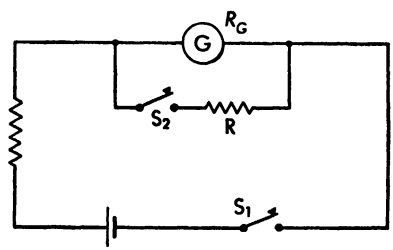


Fig. 11.4

Let us assume that S_2 (in Fig. 11.4) is open and S_1 is closed, in which case the galvanometer registers a deflection. Now let S_1 be opened. The only torque now acting upon the galvanometer coil (we are disregarding frictional effects for the time being) is that exerted by the suspension. Hence, the equation of motion of the galvanometer, given by Newton's second law, is

$$\mathcal{I}\ddot{\theta} = -\tau\theta \qquad [11\text{-}16]$$

where $\mathcal{I}$ is the moment of inertia of the coil, θ is its angular displacement, and τ is the torsional constant of the suspension. Equation [11-16], the equation of simple harmonic motion, has the solution

$$\theta = \theta_0 \cos\left[\sqrt{\frac{\tau}{\mathcal{I}}}\, t - \delta\right] \qquad [11\text{-}17]$$

where θ_0 is the amplitude of the motion, and δ, the phase angle, depends upon the initial conditions. Hence, under these idealized (frictionless) conditions, the open-circuited galvanometer coil swings back and forth with undamped harmonic motion.

In practice, of course, the motion is not frictionless, and equation [11-16] should contain a term representing the frictional torque exerted by the surround-

ing air. This frictional torque is proportional to the angular velocity of the coil and always opposes the motion. Equation [11-16] corrected for this effect, becomes

$$g\ddot{\theta} = -\tau\theta - g\dot{\theta} \qquad [11\text{-}18]$$

or

$$\ddot{\theta} + \frac{g}{g}\dot{\theta} + \frac{\tau}{g}\theta = 0 \qquad [11\text{-}19]$$

Equation [11-19], when solved, gives θ as a function of time. We shall not solve this equation now for the reason that we shall encounter below, in our discussion of electromagnetic damping, a similar but more interesting equation, which we shall solve. We note, however, at this point that the frictional effect of the air is small, with the result that the open-circuited galvanometer swings back and forth many times, with decreasing amplitude, before eventually coming to rest. This type of motion, which is known as underdamped, is discussed further below.

Electromagnetic Damping

If, while the galvanometer is swinging as above, the switch S_2 is closed, the situation is significantly altered. The galvanometer is no longer open-circuited but, instead, is part of a closed circuit consisting of the resistance R and itself. Under these conditions the emf which is induced in the galvanometer coil, as it swings in the field of the galvanometer's permanent magnet, gives rise to a current whose magnetic field opposes the motion by which it was produced (Lenz's law). We shall now derive the expression for the magnitude of this *electromagnetic damping*.

As the coil rotates in the field of uniform induction provided by the permanent magnet, the electric field induced in each length of wire comprising the sides of the coil has the value

$$\boldsymbol{E} = \boldsymbol{v} \times \boldsymbol{B} \qquad [9\text{-}14]$$

where $\boldsymbol{v}$ is the velocity of motion of the length of wire. The magnitude of the back emf induced in each length of wire is

$$|\Delta\varepsilon_B| = \int_0^l \boldsymbol{E}\cdot d\boldsymbol{l} = El \qquad [11\text{-}20]$$

Since the emf's on the sides of the coil add, and since $\boldsymbol{v}$ and $\boldsymbol{B}$ are normal to one another, the total back emf induced in the coil is

$$\varepsilon_B = N2vBl \qquad [11\text{-}21]$$

or

$$\varepsilon_B = N2\left(\frac{w}{2}\frac{d\theta}{dt}\right)Bl = NAB\frac{d\theta}{dt} \qquad [11\text{-}22]$$

where $w/2$, half the width of the coil, is the dimension involved in changing from linear to angular velocity. We now calculate the magnetic moment of the current to which this emf gives rise, and then the torque which it experiences.

The induced current in the galvanometer-plus-R circuit is given by

$$i_B = \frac{\varepsilon_B}{R_G + R} \qquad [11\text{-}23]$$

and the resulting magnetic moment, for the coil of N turns, by

$$m_B = Ni_BA = \frac{N\varepsilon_BA}{R_G + R} \qquad [11\text{-}24]$$

The resulting torque has the magnitude

$$T_B = |\boldsymbol{m}_B \times \boldsymbol{B}| = \frac{N\varepsilon_BAB}{R_G + R} = \frac{N^2A^2B^2}{(R_G + R)}\frac{d\theta}{dt} \qquad [11\text{-}25]$$

Since this torque, as did the frictional one, always opposes the motion, it acts as a damping agent. Thus, the equation of motion of the swinging galvanometer when it is in a series circuit with a resistance R is given by

$$g\ddot{\theta} = -\tau\theta - \left[g + \frac{N^2A^2B^2}{(R_G + R)}\right]\dot{\theta} \qquad [11\text{-}26][4]$$

which is simply equation [11-18] with the electromagnetic damping term added.

Before solving equation [11-26], we note that the electromagnetic damping decreases as the external resistance increases. Indeed, when the external resistance is infinite, the electromagnetic damping is zero: this is the open-circuited, underdamped case studied earlier. On the other hand, when $R = 0$, the electromagnetic damping is a maximum and the motion of the galvanometer is greatly *overdamped*. For a particular intermediate value of R, known as the *critical damping resistance* and designated by the symbols *CDRX*, the motion is said to be *critically* damped. We shall now discuss these matters further.[5]

To facilitate the solving of equation [11-26] we write it in the form

$$\ddot{\theta} + \left[\frac{g}{g} + \frac{N^2A^2B^2}{g(R_G + R)}\right]\dot{\theta} + \frac{\tau}{g}\theta = 0 \qquad [11\text{-}27]$$

[4] For convenience in this section, we shall speak only of the angular deflection of the galvanometer. The actual scale deflection can always be calculated from it by $d = 2r\theta$ (equation [11-6]).

[5] Moving coil galvanometers are often damped by attaching to the coil a coplanar loop of heavy copper wire. This loop is not connected electrically to the coil but serves as an isolated electromagnetic damping agent. Its effect is not taken into account in what follows.

and use the abbreviations $$b = \frac{g}{\mathcal{I}} + \frac{N^2A^2B^2}{\mathcal{I}(R_G + R)} \qquad [11\text{-}28]$$

and $$c = \frac{\tau}{\mathcal{I}} \qquad [11\text{-}29]$$

Thus, we have

$$\ddot{\theta} + b\dot{\theta} + c\theta = 0 \qquad [11\text{-}30]$$

or $$(D^2 + bD + c)\theta = 0 \qquad [11\text{-}31]$$

The roots of the auxiliary equation

$$D^2 + bD + c = 0 \qquad [11\text{-}32]$$

are

$$D_1 = \frac{-b + \sqrt{b^2 - 4c}}{2} \quad \text{and} \quad D_2 = \frac{-b - \sqrt{b^2 - 4c}}{2} \qquad [11\text{-}33]$$

The nature of the solution of equation [11-30] is critically dependent upon the value of the quantity, $b^2 - 4c$, which appears under the root sign in equation [11-33]. There are three possibilities:

(a) $b^2 - 4c = 0$ —this is the critically damped case;

(b) $b^2 - 4c > 0$ —this is the overdamped case; [11-34]

(c) $b^2 - 4c < 0$ —this is the underdamped case.

We shall consider each of these cases in turn.

(a) *Critically Damped Vibration* ($b^2 - 4c = 0$)

Here the two roots of the auxiliary equation are equal, and the solution of equation [11-30] takes the form

$$\theta = (\mathfrak{A} + \mathfrak{B}t)e^{-bt/2} \qquad [11\text{-}35]$$

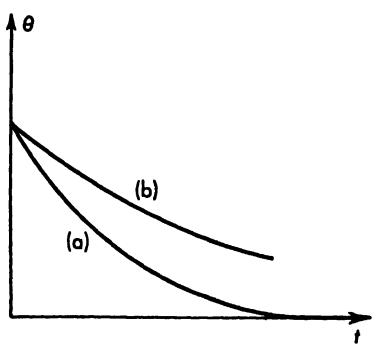


Fig. 11.5

This equation is plotted as curve (*a*) in Figure 11.5 and represents the case in which the galvanometer coil returns to its zero position in the shortest possible time without overshooting—the so-called "deadbeat" case.

Recalling the quantities that constitute b and c (equations [11-28] and [11-29]), we see that the condition for critically damped

motion is

$$\left[\frac{f}{g} + \frac{N^2A^2B^2}{g(R_G + R)}\right]^2 - \frac{4\tau}{g} = 0 \qquad [11\text{-}36]$$

or, since the frictional damping is usually small in comparison with the electromagnetic damping,

$$\frac{N^4A^4B^4}{(R_G + R)^2} = 4\tau g \qquad [11\text{-}37]$$

R, in this equation, is the critical damping resistance. Equation [11-37] is not very useful in ascertaining the value of the critical damping resistance, but it is a helpful guide in selecting or altering a galvanometer to secure critically damped operation in some particular application.

When possible, galvanometers should be operated in the critically damped condition, and manufacturers generally state, when supplying a galvanometer, the value of the critical damping resistance which should be used with it. If the resistance of the associated circuit is small, an extra resistance can be added to bring it up to the *CDRX* value. If the resistance of the associated circuit is large, the galvanometer can be shunted with whatever value of resistance in parallel with the associated circuit produces the *CDRX* value. The loss of voltage sensitivity, on the one hand, and of current sensitivity, on the other, is usually more than offset by the elimination of sluggishness or overshooting. If the *CDRX* value for a given galvanometer is not known, it can soon be determined experimentally by shunting a swinging galvanometer with resistances of various values and observing which results in critically damped motion. The *CDRX* value is often ~10 times the galvanometer resistance.

(b) *Overdamped Vibration* ($b^2 - 4c > 0$)

In this case the solution of equation [11-30] takes the standard form

$$\theta = \mathfrak{A}e^{(-b+\sqrt{b^2-4c})t/2} + \mathfrak{B}e^{(-b-\sqrt{b^2-4c})t/2} \qquad [11\text{-}38]$$

which is plotted in Figure 11.5 as curve (b). Here the coil returns to its zero position at a slower rate than in the critically damped case and does so without overshoot.

(c) *Underdamped Vibration* ($b^2 - 4c < 0$)

Here the galvanometer coil undergoes a damped vibration, as will now be seen.

First, it is advantageous to write

$$\frac{\sqrt{b^2 - 4c}}{2} = \frac{\sqrt{-1}\sqrt{4c - b^2}}{2} = j\frac{\sqrt{4c - b^2}}{2} = jx \qquad [11\text{-}39]$$

where x is a real quantity. The solution of [11-30] may then be written:

$$\theta = \mathfrak{A}e^{(-b/2+[jx])t} + \mathfrak{B}e^{(-b/2-[jx])t} = e^{-bt/2}(\mathfrak{A}e^{jxt} + \mathfrak{B}e^{-jxt}) \qquad [11\text{-}40]$$

or in the following equivalent forms

$$\theta = e^{-bt/2}[\mathfrak{A}(\cos xt + j \sin xt) + \mathfrak{B}(\cos xt - j \sin xt)] \qquad [11\text{-}41]$$

$$\theta = e^{-bt/2}[(\mathfrak{A} + \mathfrak{B}) \cos xt + j(\mathfrak{A} - \mathfrak{B}) \sin xt] \qquad [11\text{-}42]$$

$$\theta = e^{-bt/2}[\mathfrak{C} \cos xt + \mathfrak{D} \sin xt] \qquad [11\text{-}43]$$

Finally, the constants $\mathfrak{C}$ and $\mathfrak{D}$ may be expressed in terms of two other constants, viz.,

$$\mathfrak{C} = \mathfrak{F} \cos \delta \qquad \text{and} \qquad \mathfrak{D} = \mathfrak{F} \sin \delta \qquad [11\text{-}44]$$

Equation [11-43] then becomes

$$\theta = e^{-bt/2}\mathfrak{F}(\cos xt \cos \delta + \sin xt \sin \delta) \qquad [11\text{-}45]$$

or

$$\theta = \mathfrak{F}e^{-bt/2} \cos (xt - \delta) \qquad [11\text{-}46]$$

It will be recalled that the equation for simple harmonic motion has the form

$$\theta = \theta_0 \cos (\omega t - \delta) \qquad [11\text{-}47]$$

where θ_0 is the amplitude of the motion, $\omega = 2\pi f$ is the angular frequency, and δ is the phase angle. Thus, equation [11-46] describes a simple harmonic motion whose amplitude,

$$\mathfrak{F}e^{-bt/2} \qquad [11\text{-}48]$$

decreases exponentially with time. The frequency of the harmonic component is

$$f = \frac{w}{2\pi} = \frac{x}{2\pi} = \frac{\sqrt{4c - b^2}}{4\pi} \doteq \frac{\sqrt{4\tau g - \dfrac{N^4A^4B^4}{(R_G + R)^2}}}{4\pi g} \qquad [11\text{-}49]$$

where the frictional damping term has been neglected in the last expression.

This underdamped vibration is shown in Figure 11.6, where it is bounded, as required by equation [11-46], by the curve representing expression [11-48] and its mirror image.

If the resistance R is large, the amplitude decays slowly, and the vibration persists for a long time. Under these conditions (cf. equation [11-49]) the frequency,

$$f \doteq \frac{1}{2\pi}\sqrt{\frac{\tau}{\mathcal{J}}} \qquad [11\text{-}50]$$

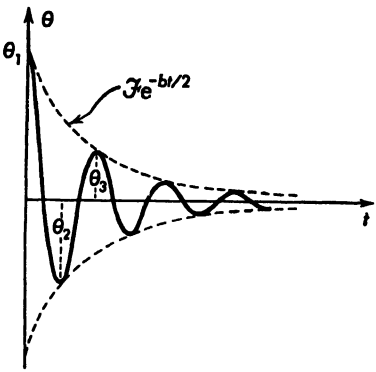


Fig. 11.6

approximates the frictionless, open-circuited value of equation [11-17].

Logarithmic Decrement

The rate at which the amplitude of an underdamped vibration decreases is expressed in terms of the logarithmic decrement, Δ, defined as the natural logarithm of the ratio of two successive maxima. Δ may be determined experimentally by observing any two successive deflections, θ_1 and θ_2 (see Fig. 11.6). Thus:

$$\Delta = \ln\frac{\theta_1}{\theta_2} = \ln\frac{\theta_2}{\theta_3} = \cdots = \ln\frac{\theta_{n-1}}{\theta_n} \qquad [11\text{-}51]$$

Δ may be calculated from equation [11-46] by computing the magnitude of the amplitude at two different times, spaced half a period apart. Thus:

$$\Delta = \ln\left[\frac{\mathcal{F}e^{-bt/2}}{\mathcal{F}e^{-b(t+T/2)/2}}\right] = \ln e^{bT/4} = \frac{bT}{4} \qquad [11\text{-}52]$$

where T is the period of the motion.

The normal way to obtain Δ, however, is to observe it experimentally, as suggested above. A more accurate value for Δ than that obtained from equation [11-51] is got by observing deflections that are several half-periods apart. Thus,

$$(n-1)\Delta = \ln\frac{\theta_1}{\theta_2} + \ln\frac{\theta_2}{\theta_3} + \cdots + \ln\frac{\theta_{n-1}}{\theta_n} \qquad [11\text{-}53]$$

$$= \ln\left[\frac{\theta_1}{\theta_2}\cdot\frac{\theta_2}{\theta_3}\cdots\frac{\theta_{n-1}}{\theta_n}\right] \qquad [11\text{-}54]$$

or

$$\Delta = \frac{1}{n-1}\ln\frac{\theta_1}{\theta_n} \qquad [11\text{-}55]$$

Galvanometer Damping During Initial Deflection

When the switch in the circuit of Figure 11.7 is closed, and the galvanometer coil begins to move so as to take up the angular position appropriate to the current

$$I = \frac{\mathcal{E}}{R_G + R} \qquad [11\text{-}56]$$

the equation of motion of the coil is given by

$$g\ddot{\theta} = -\tau\theta - g\dot{\theta} + NiAB \qquad [11\text{-}57]$$

In the last term on the right-hand side, which gives the net electromagnetic torque (equation [11-2]), i represents the instantaneous current flowing in the circuit. It is given by

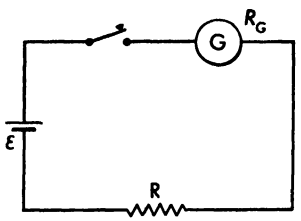


Fig. 11.7

$$i = \frac{\mathcal{E} - NAB\dot{\theta}}{R_G + R} \qquad [11\text{-}58]$$

where the numerator on the right-hand side is the algebraic sum of the applied and back emf's (cf. equation [11-22]). Upon substitution of [11-58], equation [11-57] becomes

$$g\ddot{\theta} + \left[g + \frac{N^2A^2B^2}{(R_G + R)}\right]\dot{\theta} + \tau\theta = \frac{NAB}{(R_G + R)}\mathcal{E} \qquad [11\text{-}59]$$

or, in terms of the abbreviations that we have been using,

$$\ddot{\theta} + b\dot{\theta} + c\theta = \frac{NAB}{(R_G + R)}\frac{\mathcal{E}}{g} \qquad [11\text{-}60]$$

or

$$\ddot{\theta} + b\dot{\theta} + c\theta = d \qquad [11\text{-}61]$$

where d is a constant. This equation is identical with equation [11-30], except for the presence of the constant on the right-hand side.

Equation [11-61] may be solved by substituting the new variable

$$\alpha = \theta - \frac{d}{c} \qquad [11\text{-}62]$$

which transforms it to

$$\ddot{\alpha} + b\dot{\alpha} + c\alpha = 0 \qquad [11\text{-}63]$$

which is identical in form with equation [11-30]. Hence, the motion of the coil as it is deflected from its zero position is underdamped, critically damped, or overdamped in the same way as is its motion during its return to zero. The value of the critical damping resistance is the same for both types of motion.

Combining equation [11-62] with our solutions of equation [11-30], we find the solution of equation [11-63] to be:

(a) *Critically damped motion* (*from equation* [11-35])

$$\theta = \alpha + \frac{d}{c} = \frac{d}{c} + (\mathfrak{A} + \mathfrak{B}t)e^{-bt/2} \qquad [11\text{-}64]$$

(b) *Overdamped motion* (*from equation* [11-38])

$$\theta = \frac{d}{c} + \mathfrak{A}e^{(-b+\sqrt{b^2-4c})t/2} + \mathfrak{B}e^{(-b-\sqrt{b^2-4c})t/2} \qquad [11\text{-}65]$$

(c) *Underdamped motion* (*from equation* [11-46])

$$\theta = \frac{d}{c} + \mathfrak{F}e^{-bt/2}\cos(xt - \delta) \qquad [11\text{-}66]$$

First let us examine the term d/c. Substituting the values of d and c, we find

$$\frac{d}{c} = \frac{\dfrac{NAB\mathcal{E}}{(R_G + R)\mathcal{I}}}{\dfrac{\tau}{\mathcal{I}}} = \frac{NAB\mathcal{E}}{(R_G + R)\tau} = \frac{NABI}{\tau} \qquad [11\text{-}67]$$

where equation [11-56] has been substituted to obtain the final expression. By comparing equation [11-67] with [11-4], we see that the quantity d/c is simply the equilibrium angular deflection. It is the value which θ will ultimately assume, although it may take a variety of routes (equations [11-64] to [11-66]) to achieve it.

Now let us look in detail at equation [11-64], with a view to evaluating the constants $\mathfrak{A}$ and $\mathfrak{B}$. This we shall do by applying the boundary conditions that at $t = 0$, both θ and $\dot{\theta}$ are zero. Thus, at $t = 0$,

$$\theta = 0 = \left[\frac{d}{c} + (\mathfrak{A} + \mathfrak{B}t)e^{-bt/2}\right]_{t=0} \qquad [11\text{-}68]$$

or

$$0 = \frac{d}{c} + \mathfrak{A} \qquad [11\text{-}69]$$

and

$$\dot{\theta} = 0 = \left[\mathfrak{B}e^{-bt/2} + (\mathfrak{A} + \mathfrak{B}t)\left(-\frac{b}{2}\right)e^{-bt/2}\right]_{t=0} \qquad [11\text{-}70]$$

or

$$0 = \mathfrak{B} - \frac{b\mathfrak{A}}{2} \qquad [11\text{-}71]$$

From equations [11-69] and [11-71] we have

$$\mathfrak{A} = -\frac{d}{c} \quad \text{and} \quad \mathfrak{B} = \frac{b\mathfrak{A}}{2} = -\frac{bd}{2c} \qquad [11\text{-}72]$$

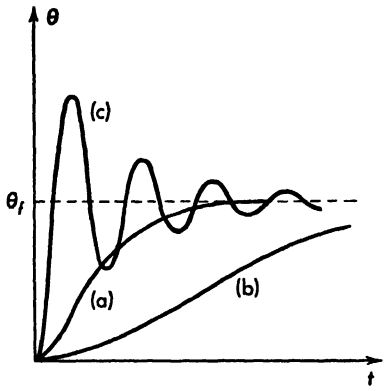


Fig. 11.8

Hence, equation [11-64] takes the form

$$\theta = \frac{d}{c} - \frac{d}{c}\left(1 + \frac{b}{2}t\right)e^{-bt/2} \qquad [11\text{-}73]$$

$$\text{or } \theta = \frac{d}{c}\left[1 - \left(1 + \frac{b}{2}t\right)e^{-bt/2}\right] \qquad [11\text{-}74]$$

This curve is plotted as curve (*a*) in Figure 11.8. Here the coil moves from its zero position to its final position ($\theta_f = d/c$) in the shortest time that is possible without overshooting its mark.

One can similarly evaluate the constants in equations [11-65] and [11-66]. When this is done, and the results plotted, the overdamped motion is typically given by curve (*b*) in Figure 11.8, while the underdamped motion follows curve (*c*). It can be seen that the underdamped coil, having moved from its zero, oscillates about its final position with continually decreasing amplitude.

11.4 The Ballistic Galvanometer

The ballistic galvanometer is used to determine the quantity of charge that passes through a circuit during a transient current. The charge might come, for example, from the discharge of a capacitor or it might be set in motion by an induced emf.

Whereas the usual current galvanometer displays a steady deflection, which is proportional to the current passing through it, the ballistic galvanometer displays a transitory deflection, which is proportional to the amount of charge which has passed through it. Often the same instrument is used for both purposes, although ballistic galvanometers are usually specially constructed to have longer than normal periods of motion ($T \geq 10$ seconds). This is to ensure that the total charge passes through the galvanometer before there is appreciable motion of the galvanometer coil, and also to facilitate the observation of the magnitude of the deflection. Under these conditions, the torque experienced by the coil is an impulsive one, and the coil acquires angular momentum found

by integrating Newton's second law for the period of the impulse, viz.,

$$\int_0^t T\,dt = \int_0^t \frac{d}{dt}(\mathcal{I}\dot{\theta})\,dt \qquad [11\text{-}75]$$

Upon substitution of the value of the torque from equation [11-2], equation [11-75] becomes

$$NAB\int_0^t i\,dt = \mathcal{I}\int_0^{\dot{\theta}_0} d\dot{\theta} \qquad [11\text{-}76]$$

which, when integrated, gives

$$NABQ = \mathcal{I}\dot{\theta}_0 \qquad [11\text{-}77]$$

where Q is the total charge that has passed through the galvanometer, and $\mathcal{I}\dot{\theta}_0$ is the angular momentum gained by the coil.

Following this initial impulse, the subsequent motion of the coil is given by equation [11-26], subject to the initial condition expressed in equation [11-77]. We shall not solve this equation, although it is not difficult to calculate from it the maximum deflection that occurs for critically damped, overdamped, or underdamped conditions. Instead, for simplicity, let us assume that the galvanometer is open-circuited, and that frictional damping may be neglected. This is, of course, an idealized case, but it approximates an important practical situation, namely, that in which a charged capacitor is allowed to discharge through an unshunted galvanometer.

The kinetic energy associated with the angular momentum acquired during the impulse is transformed into potential energy as the coil, in rotating, does work against its suspension. The potential energy associated with the maximum angular displacement θ_m is given by

$$\text{P.E.} = \int_0^{\theta_m} \tau\theta\,d\theta = \frac{\tau\theta_m^2}{2} \qquad [11\text{-}78]$$

where $\tau\theta$, as we have noted earlier, is the restoring torque exerted by the suspension. We shall now use equations [11-77] and [11-78] to relate Q to the maximum angular displacement. From [11-77] the initial kinetic energy is

$$\text{K.E.} = \tfrac{1}{2}\mathcal{I}\dot{\theta}_0^2 = \frac{N^2A^2B^2Q^2}{2\mathcal{I}} \qquad [11\text{-}79]$$

At the extreme point in the swing of the galvanometer the potential energy has this value. Thus,

$$\frac{\tau\theta_m^2}{2} = \frac{N^2A^2B^2Q^2}{2\mathcal{I}} \qquad [11\text{-}80]$$

or $$\theta_m = \frac{NAB}{\sqrt{\tau \mathcal{G}}} Q \qquad [11\text{-}81]$$

or $$d_m = r(2\theta_m) = \frac{2rNAB}{\sqrt{\tau \mathcal{G}}} Q = KQ \qquad [11\text{-}82]$$

where d_m is the maximum scale deflection that results from the passage of Q coulombs of charge. If the frictional effect of the air is taken into account, d_m is slightly smaller than in equation [11-82], but it is still proportional to Q. Also, when the galvanometer is not open-circuited, the d_m is still always proportional to Q, regardless of the magnitude of the electromagnetic damping.

The constant K appearing in equation [11-82] is the *ballistic sensitivity* of the galvanometer. It is the reciprocal of the charge required to produce the standard scale deflection.[6]

In summary: the maximum scale reading or "throw" of a ballistic galvanometer is proportional to the total charge passing through the galvanometer. For this reason the ballistic galvanometer may be used to measure the total charge that passes in a current of short duration.

Calibration of a Ballistic Galvanometer by Capacitor Discharge

The simplest method of calibrating an open-circuited ballistic galvanometer is shown in Figure 11.9. When the switch is thrown to position A, the capacitor is charged to the potential $\mathcal{E}$, thereby acquiring the total charge

$$Q = \mathcal{E}C \qquad [3\text{-}2]$$

Then, when the switch is thrown to position B, this charge flashes through the galvanometer, producing an initial throw d_m. From equation [11-82] we have

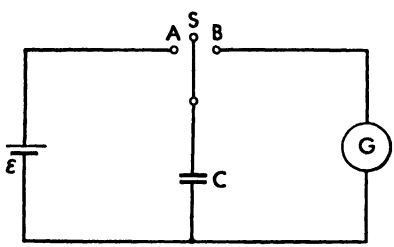


Fig. 11.9

$$K = \frac{d_m}{Q} \qquad [11\text{-}83]$$

or $$K = \frac{d_m}{\mathcal{E}C} \qquad [11\text{-}84]$$

from which the ballistic constant of the galvanometer can be obtained. If C is a standard capacitance and $\mathcal{E}$ is a standard cell, K can be determined with good accuracy, particularly if d_m's are observed for several values of C. One can then make a d_m-versus-$\mathcal{E}C$ plot whose slope gives the value of K.

[6] The ballistic sensitivity is sometimes defined as the charge required to produce the standard deflection. That is, equation [11-82] is written as $Q = K'd_m$, where $K' = 1/K$.

Once the ballistic galvanometer has been calibrated, the arrangement shown in Figure 11.9 provides a convenient method for determining the emf of a battery. That is, the $\mathcal{E}$ in equation [11-84] is regarded as the unknown. The arrangement can thus serve as a voltmeter, with the advantage over the ordinary voltmeter that it indicates the open-circuited voltage (which is equal to the emf) of the battery, no matter what its internal resistance is.

Obviously, the same arrangement can be employed to measure an unknown capacitance. Also, if the switch S, after being moved from position A, is kept in the intermediate position for a period of time, before being moved to position B, one can compute from the reduction in the galvanometer throw the leakage resistance of the capacitor. This experiment is described in greater detail in § 12.2, where its usefulness in the determination of large resistances is pointed out.

Calibration of a Ballistic Galvanometer Using a Mutual Inductance

This operation has been adequately described at the end of § 10.20. It is particularly suited to experiments such as the study of hysteresis using the Rowland ring, in which the ballistic galvanometer is employed to determine the amount of charge that passes as a result of an emf induced in a closed circuit.

The Ballistic Galvanometer as a Fluxmeter

If the ballistic galvanometer is heavily overdamped (say, by reducing the restoring couple exerted by the suspension), $c \doteq 0$ and equation [11-38] then becomes

$$\theta \doteq \mathfrak{A} + \mathfrak{B}e^{-bt} \qquad [11\text{-}85]$$

At $t = 0$, $\theta = 0$, and

$$\dot{\theta}_0 = \frac{NABQ}{\mathcal{I}} \qquad [11\text{-}77]$$

Assuming equation [11-85] is exact and applying these boundary conditions to it, we find

$$\mathfrak{A} = -\mathfrak{B} = \frac{NABQ}{b\mathcal{I}} \qquad [11\text{-}86]$$

which, when substituted in equation [11-85], give for the equation of motion of the coil,

$$\theta = \frac{NABQ}{b\mathcal{I}}(1 - e^{-bt}) \qquad [11\text{-}87]$$

Thus, the coil moves slowly to its maximum deflection,

$$\theta_m = \frac{NABQ}{b\mathcal{G}} \qquad [11\text{-}88]$$

according to an exponential law. Hence, the reading of d_m ($= 2r\theta_m$) can be made leisurely and with care.

The *fluxmeter* is an important application of a ballistic galvanometer acting under these conditions. In the fluxmeter a flat test coil (N' turns of cross-sectional area A'), in series with the galvanometer, is placed in a magnetic field of induction B', with the plane of the coil normal to the direction of the field. The total flux passing through the test coil is

$$\Phi'_m = B'A' \qquad [11\text{-}89]$$

where it is assumed that B' is constant across the face of the coil. When the flux through the coil changes, an emf is induced in it of magnitude

$$|\varepsilon| = N'\frac{d\Phi'_m}{dt} = N'A'\frac{dB'}{dt} \qquad [11\text{-}90]$$

which gives rise to a current in the circuit of magnitude:

$$i = \frac{\varepsilon}{R_G + R} = \frac{N'A'}{R_G + R}\frac{dB'}{dt} \qquad [11\text{-}91]$$

where R is the resistance of the test coil. The total charge flowing during the period of this current is

$$Q = \int_0^t i\,dt = \int_0^{\Delta B'} \frac{N'A'}{R_G + R}\,dB' = \frac{N'A'}{R_G + R}\Delta B' \qquad [11\text{-}92]$$

In practice, once the ballistic galvanometer has subsided to its zero position, the test coil is snatched from the magnetic field. Thus, the change $\Delta B'$ is B' itself.

Once the fluxmeter has been calibrated, it can be used to determine rapidly the strength of a uniform magnetic field, B', in accordance with the relationship:

$$d_m = KQ = \left(\frac{KN'A'}{R_G + R}\right)B' \qquad [11\text{-}93]$$

where the quantities in parentheses are constants of the apparatus. Cali-

bration is accomplished by observing the deflection that occurs when the test coil is snatched from a magnetic field of known induction.

11.5 Meters Operating on the Moving-Coil Principle

Most of the DC ammeters and voltmeters that are in common use operate on the moving-coil principle. As was mentioned in § 6.8, ammeters and voltmeters having various ranges can be constructed from a basic meter. This basic meter is actually a low-sensitivity, portable galvanometer. The coil of the galvanometer is not suspended by a fine wire but rotates, instead, on hardened steel or jewel bearings, with spiral springs providing the restoring torque. A permanent magnet provides the deflecting field in the same manner as in the galvanometer pictured in Figure 11.1. The deflection of the coil is indicated by a pointer attached to it.

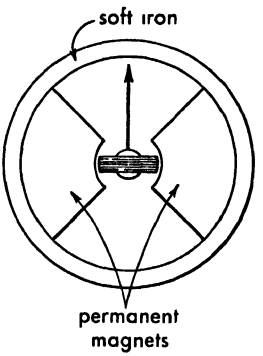


Fig. 11.10

The development in recent years of permanent magnet materials for which the coercive force is large, has made possible the type of construction shown in Figure 11.10. Here, a hollow cylinder of soft iron provides the return path for the magnetic flux. However, the soft iron serves as a magnetic shield as well, with the result that the reading of a meter constructed along the lines indicated by Figure 11.10 is relatively insensitive to the presence of nearby magnetic fields, such as those associated with current-carrying conductors.

Ammeters

The basic galvanometer is converted into an ammeter by shunting it with a resistance which carries a portion, usually most, of the current. If the basic galvanometer, of resistance r, has a full-scale deflection when carrying a current I_r, then when shunted by a resistor of value

$$R = \frac{I_r r}{I - I_r} \qquad [11\text{-}94]$$

full-scale deflection corresponds to a total current I in the parallel combination of galvanometer and shunt (see Fig. 6.12 and equation [6-12]).

Since the coil of the basic meter and its shunt constitute a closed electrical path, no provision is necessary for additional damping.

Voltmeters

The basic galvanometer is converted into a voltmeter by the addition of a series resistor, known as a *multiplier*. As above, let us suppose the basic meter has a resistance r and reads full scale when carrying a current I_r. The full-scale deflection corresponds to a potential drop across the meter of amount

$$V_r = I_r r \qquad [11\text{-}95]$$

When a resistor R is placed in series with the basic meter, full-scale deflection corresponds to a total potential drop across the combination of amount

$$V = I_r(r + R) \qquad [11\text{-}96]$$

Thus, the value of the multiplier that is needed to construct a voltmeter which reads V volts full scale is given by

$$R = \frac{V - I_r r}{I_r} \qquad [11\text{-}97]$$

It is interesting to note, from equation [11-96], that the ratio,

$$\frac{r + R}{V} = \frac{1}{I_r} \qquad [11\text{-}98]$$

is constant for a given basic meter. The quantity $(r + R)/V$ is the ohms-per-volt rating, referred to in § 6.8.

Since the resistance of a voltmeter is large, the effect of electromagnetic damping is small. For this reason the coil of the basic meter is often wound on an aluminum frame in which induced (and, therefore, damping) currents can flow.

For voltmeters which measure high voltages ($V > 5000$ volts), the multiplier is often external to the meter assembly.

11.6 Electrodynamometer Action

In the moving-coil galvanometer which we have been discussing, the magnetic field is provided by a permanent magnet. There exists another important class of galvanometer, also with moving coils, in which the magnetic field is established by fixed coils carrying the current that is to be measured. The fixed coils are located near the suspended one. The arrangement is shown schematically in Figure 11.11.

As with the d'Arsonval type of instrument, the movable coil experiences a torque of magnitude

$$T = |\boldsymbol{m} \times \boldsymbol{B}| \qquad [11\text{-}99]$$

Here $\boldsymbol{m}$ is the magnetic moment of the movable coil and $\boldsymbol{B}$ is the induction at the position of the movable coil arising from the current in the two fixed coils. The same current, I, is present in all three coils. For certain coil configurations the magnitude of the torque experienced by the movable coil can be calculated, but we shall not attempt to do this. Instead, let us investigate the manner in which T depends upon I.

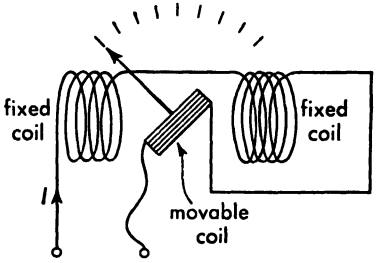


Fig. 11.11

In the first place, the magnetic moment of the movable coil, m, is directly proportional to I. Then, in addition, the induction arising from the fixed coils (cf. equation [8-26]) is directly proportional to I. Hence, the torque experienced by the movable coil is proportional to I^2:

$$T \propto I^2 \qquad [11\text{-}100]$$

in which case it depends not upon the direction of the current, but only upon its magnitude. This characteristic is of the greatest importance, for it allows alternating currents, as well as direct currents, to be measured by this type of instrument. Although the torque is not constant for alternating currents, the natural frequency of the coil is much too low for it to follow the pulsations in the torque, and the coil takes up a position in which the average torque exerted upon it is balanced by the torque exerted by the restoring springs.

The deflection of the pointer needle is not simply proportional to I^2, for the reason that the sine of the angle between $\boldsymbol{m}$ and $\boldsymbol{B}$ also helps to determine the magnitude of the torque (cf. equation [11-99]). The movable coil is generally mounted so that $\boldsymbol{m}$ and $\boldsymbol{B}$ are normal to one another for a mid-scale deflection. Then, for larger deflections, the decline in the value of the sine of the angle counteracts the spreading of the scale which the I^2-dependence would otherwise produce. The linearity of the scale can be further improved by altering the shape of the fixed coils.

Instruments of this type, which operate on the basis of the force between two current-carrying coils, are said to employ the *electrodynamometer* principle.[7]

[7] The current balance, described in § 8.13, is another example of the electrodynamometer principle.

Ammeters

The arrangement shown in Figure 11.11 would itself serve as a low-current ammeter. For higher currents, the entire current is allowed to pass through the fixed coils but the movable coil is shunted with a resistor of proper value. The instrument, when used with alternating currents, reads the *effective current*,[8] or equivalent direct current. Thus, an electrodynamometer meter may be calibrated on direct and then used to measure alternating current. Since the current that is being measured is called upon to excite the magnetic field that in a d'Arsonval instrument is provided by a permanent magnet, the electrodynamometer ammeter consumes more power than an equivalent d'Arsonval instrument. Damping is effected by vanes fastened to the moving coil.

Voltmeters

The basic electrodynamometer meter is converted into a voltmeter by the addition of a series resistor. To obtain the necessary sensitivity for the small currents that are present, both fixed and movable coils are wound with many more turns of wire than in the ammeter. This increases the inductance of the circuit and limits the usefulness of the instrument to the measurement of DC voltages or low-frequency AC voltages. Normal electrodynamometer voltmeters are accurate up to frequencies of ~100 cycles/second.

Wattmeters

The instantaneous power dissipation is given by the product iv. In the wattmeter the fixed coils are made series elements in the circuit, while the movable coil, in series with a large resistance, R, is connected across the line. These connections are shown schematically in Figure 11.12. Thus, the current in the fixed coils is representative of the amperage, and the current in the movable coil is representative of the voltage. Consequently, the torque experienced by the movable coil is proportional to the power. As with voltmeters, the DC calibration of wattmeters is not reliable for AC frequencies above ~100 cycles/sec.

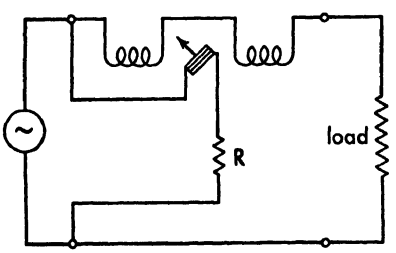


Fig. 11.12

[8] This is the so-called *RMS* (root-mean-square) current: it will be discussed in Chapter 13.

11.7 Moving-Iron Instruments

Moving-iron instruments employ the principle that a piece of iron tends to move from weaker to stronger parts of a magnetic field (cf. footnote to § 10.14).

A fixed coil carries the current, establishing a field of induction B. An iron vane (Fig. 11.13) is pivoted at one end of the coil, and rotates as it is drawn into the region of the stronger field within the coil. The rotation may be opposed by gravity or by a spring.

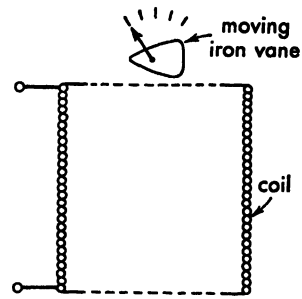


Fig. 11.13

The torque that deflects the iron vane is given by

$$T = \frac{\partial}{\partial\theta}\left(\frac{B^2}{2\mu}\right) \qquad [11\text{-}101]$$

which is proportional to the square of the current in the coil. This suggests that the deflection obeys a square law but, in fact, the law can be altered by shaping the vane. On the other hand, since the expression for the torque always contains the square of the current, the torque is in the same direction regardless of the direction of the current. Hence, as with electrodynamometer instruments, the moving-iron type may be used for alternating currents, although it is less accurate.

Moving-iron ammeters and voltmeters are the most widely used AC meters for low-frequency applications. Since the moving element carries no current, it is not susceptible to damage resulting from current overloads. If the iron comprising the vane exhibits significant hysteresis effects, the reading of the meter will depend upon its previous history. This introduces a source of error in DC measurements but, of course, not in AC work. The error can be almost eliminated, however, by use of permalloy or some other narrow-hysteresis loop material for the vane and for the iron core on which the coil is sometimes wound to secure higher values of B.

11.8 Rectifier Instruments

The d'Arsonval type of movement is more sensitive than the electrodynamometer or moving-iron type. For this reason sensitive AC meters are often made by putting a *rectifier* in series with a d'Arsonval type of instrument. Metallic rectifiers (§ 6.7), arranged in a "bridge" circuit (Fig. 11.14), are used

for this purpose. In the arrow-and-bar representation of a rectifier, used in Figure 11.14, the arrow indicates the permitted direction of current flow.

When Y is positive with respect to X, current passes through the meter from point *2* to point *1*, along the route *3214*. When the polarity reverses, with X now positive with respect to Y, current still passes through the meter from point *2* to point *1*, but now along the route *4213*. The torque experienced by the moving coil in the meter is always in the same direction, albeit pulsating. Because of the inertia of the coil, it does not follow the pulsations of the torque, but takes up a position corresponding to the average torque.

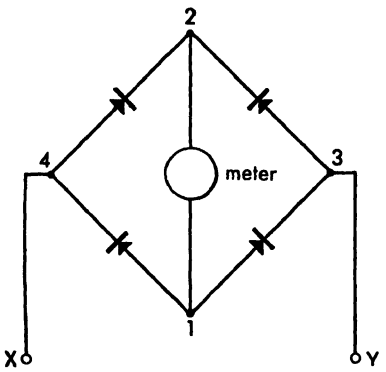


Fig. 11.14

Since this type of meter indicates the average magnitude of the current, rather than the RMS value, it is customary to calibrate it by passing sinusoidal alternating current through it and marking the scale in rms values. An instrument calibrated in this manner may give incorrect values for nonsinusoidal alternating currents. The rectifier-type instrument provides much higher ohm/volt ratings (~2000 ohms/volt) than electrodynamometer or moving-iron types.

11.9 Thermocouple and Hot-Wire Instruments

The *thermocouple* meter is pictured in Figure 7.10 and is discussed in § 7.9, as part of the discussion on thermoelectricity. Since this type of meter utilizes the heating effect of an electric current, it measures rms current and can be used for both direct and alternating currents. Further, since the resistive element can be made from a straight piece of wire, inductive effects are negligible, and the meter can be used as an ammeter for AC currents of high frequency (up to ~50 megacycles). In the most sensitive meters, the resistive and thermocouple elements are housed in a vacuum in order to reduce heat losses. The resistive or heater element will not survive excessive currents, usually burning out for ~50 percent overload. Voltmeters can be constructed from thermocouple meters but, because of the

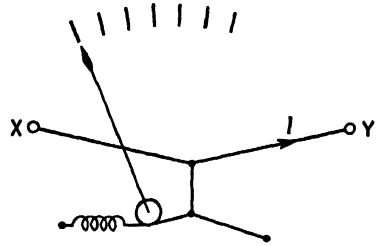


Fig. 11.15

inductance associated with the series resistance, they do not provide the wide frequency range characteristic of thermocouple ammeters.

The *hot-wire* meter utilizes the linear expansion that occurs when a wire is heated. The resistance wire XY (Fig. 11.15) carries the current to be measured. As the wire expands, its mid-point is pulled to one side by the action of a spring which, passing around a pivoted drum, exerts a force on a wire which is fastened to the resistance wire. Thus, an expansion of the resistance wire results in a rotation of the drum, and the consequent motion of a pointer fastened to it. As with thermocouple meters, the resistance wire is easily burned out. The chief application of hot-wire meters is in the measurement of high-frequency currents. Hot-wire and thermocouple meters are usually calibrated with DC currents of known value.

11.10 Electrometers

The term *electrometer*, as now used, refers to any piece of equipment, with high input resistance, that can measure small differences in potential. An important use of such an instrument is shown in Figure 11.16, where it is employed to determine the magnitude of a small current, I. This current is a constant current, such as that reaching the ion collector in a mass spectrometer (Fig. 8.17) or that reaching the plate of a photoelectric tube. When I is *very* small, the response of even the most sensitive galvanometer is either slight or negligible. However, if it is allowed to flow through a sufficiently high resistance, R, the resulting potential drop,

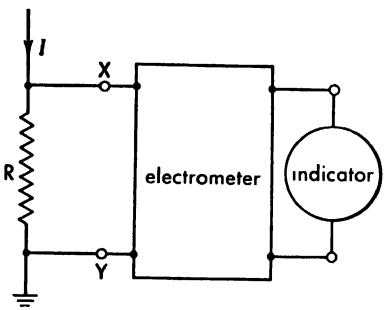


Fig. 11.16

$$V = IR \qquad [11\text{-}102]$$

is large enough to be measured by the electrometer.

It is imperative, of course, that the resistance which the electrometer presents between points X and Y be much larger than R. Otherwise, R is shunted, and the relationship given in equation [11-102] no longer holds.

Quadrant Electrometer

The original electrometers were electrostatic devices in which a conductor took up an equilibrium position under the action of an electrical force and some opposing force. The electrostatic voltmeter described in § 3.11 answers this description, although it is not properly called an electrometer since it is primarily used to measure large differences in potential.

In the *quadrant electrometer*, which was an integral component in many of the classic experiments in atomic physics, a thin metallic vane is suspended inside four "quadrants," made by quartering a shallow metallic pillbox (Fig. 11.17). Two opposite quadrants are grounded, and the other two are connected to the potential, V ($= IR$), to be measured. The vane, or "needle," is maintained at the potential V_n. As the potential V changes, the needle rotates so as to reduce the electrical energy of the system, and reaches equilibrium when the torque arising from this cause is balanced by the restoring torque exerted by the suspension. The angular position of the vane is observed by reflecting light from a mirror fastened to the bottom end of its suspension. Detailed descriptions of this and other types of electrostatic electrometer are given in most older textbooks on electricity and magnetism. Steady currents of $\sim 10^{-14}$ amperes can be measured with such instruments.

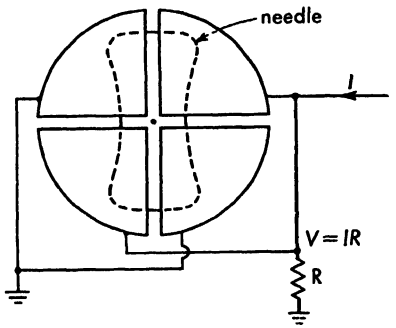


Fig. 11.17

Vacuum-Tube Electrometer

This is a vacuum-tube device, for the amplification of DC voltages, in which special pains have been taken to increase the input resistance. The potential difference across the resistance R (Fig. 11.18) is applied between the grid and cathode of the vacuum tube which forms the first stage of the amplifier. For the reason mentioned earlier, the input resistance (or *grid resistance*) of this tube must be substantially higher than R, the *grid-leak resistor*. Tubes are available with input resistances $\sim 10^{14}$ ohms. These can be

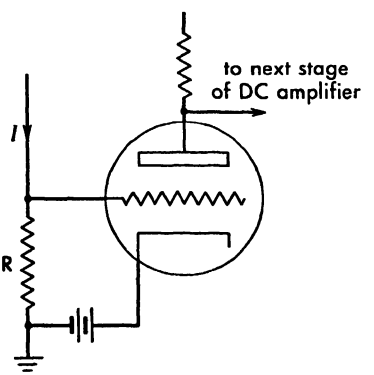


Fig. 11.18

used with grid-leaks of $\sim 10^{11}$ ohms to measure currents of $\sim 10^{-15}$ amperes. The output of a DC amplifier is subject to zero drift.

Vibrating-Reed Electrometer

At the present time the most useful electrometer for the determination of small currents is the vibrating-reed type. Its operation will be explained with reference to Figure 11.19.

The zero-drift of an AC amplifier is very much less than that of a DC amplifier. The vibrating-reed electrometer provides a means of effectively transforming a DC voltage into an AC voltage, after which it can be amplified by an AC amplifier. The DC voltage, V_0, is developed across the high resistance, R, when the current, I, passes through it. Under equilibrium conditions, points X and Y are both at this same potential, V_0, with respect to ground. If a difference in potential is created between X and Y, the action of R' and C is to delay the passage of charge that is necessary to remove the potential difference.[9] One plate of the capacitor, C, is a vibrating reed, with the result that the capacitance of C varies periodically with amplitude dC. However, for the reason given above, the charge on the capacitor does not change appreciably during the periodic variation in the capacitance. That is, at any time,

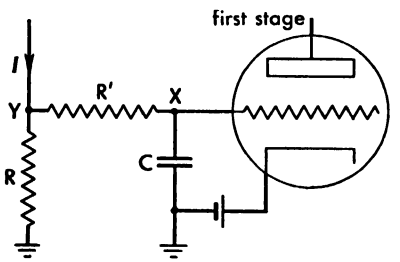


Fig. 11.19

$$VC = Q \doteq \text{constant} \qquad [11\text{-}103]$$

Consequently, an AC voltage is developed at point X of amplitude

$$dV \doteq -\frac{V_0}{C_0}\,dC \qquad [11\text{-}104]$$

found by differentiating equation [11-103]. This AC voltage, whose amplitude is proportional to the DC voltage, V_0, is then amplified by the AC amplifier.

Vibrating-reed electrometers are available with input resistances of 10^{15} ohms. They can detect currents of less than 10^{-15} amperes.

Limit to Value of the Voltage-Dropping Resistor

Mention has already been made that R, the voltage-dropping resistor, must be smaller than the input resistance of the electrometer. It should be at

[9] We shall see in the next chapter that, when resistances *and* capacitances are involved, the passage of charge is not instantaneous. Hence, equilibrium conditions, following a change in potential, are not immediately established.

least two orders of magnitude smaller. There is an additional consideration which limits the value of R.

As we have hinted above, and will establish in the next chapter, the presence of resistance and capacitance prevents changes in potential from taking place instantly. In fact, the larger the product RC, the slower the change. In the case of an electrometer, the capacitance that is involved is that between the terminals across which R is connected. Hence, for a given electrometer, there is a maximum usable value of R, unless one can tolerate a painfully slow response to changes in the magnitude of the incoming current.

11.11 Problems

E (11.1) The tangent galvanometer, which is now rarely used, consists of a short coil, of radius r, whose face is parallel to the magnetic field of the earth, $\boldsymbol{B}_E$, and at the center of which is pivoted a small magnet, which is free to rotate in the horizontal plane (Fig. 11.20). Show that when the coil is carrying a current of I amperes the equilibrium position of the magnet is given by $\tan\theta = \mu_0 NI/2rB_E$, where N is the total number of turns in the coil.

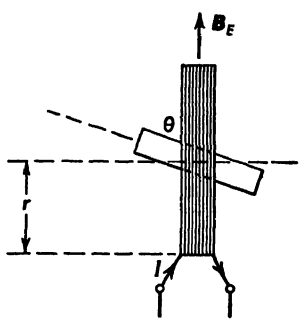


Fig. 11.20

M (11.2) Calculate the time at which the coil of an overdamped galvanometer ($\theta = \theta_0$ at $t = 0$) is moving with maximum velocity.

E (11.3) Show that the current sensitivity, k, and the undamped ballistic sensitivity, K, are related as follows:

$$K = \frac{2\pi k}{T}$$

where T is the period of the galvanometer's undamped motion.

E (11.4) Assume that the motion described by equation [11-35] begins from rest when $\theta = \theta_0$. Determine the value of the constants $\mathcal{A}$ and $\mathcal{B}$.

Chapter 12

Transient Currents

If a circuit contains resistance only, the application of a DC emf results in the current rising instantaneously to its final steady value. The response of the circuit, however, is not instantaneous if it contains inductive and/or capacitative elements. Here a change in the applied emf results in transient effects which differ greatly from the steady-state behavior. Several such cases will now be studied.

12.1 Discharge of a Capacitor Through a Resistor

Consider, for example, the discharge of a capacitor through a resistor, a process shown diagrammatically in Figure 12.1. When the switch S is in position A, the capacitor C is charged until the potential difference across it, V_0, is numerically equal to the emf of the battery.[1] When the switch is thrown to position B, the capacitor is free to discharge through the resistance, R.

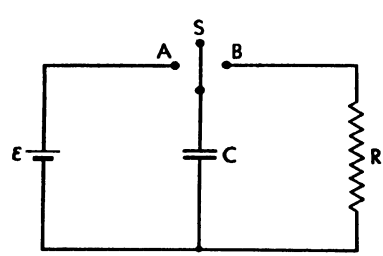


Fig. 12.1

Kirchhoff's second law can be applied to the RC circuit to obtain an expression which describes the situation at any time during this discharge. This expression is

[1] It will be recalled from § 5.8 that the potential difference across a battery which is not supplying current is equal to the emf of the battery.

$$0 = iR + \frac{q}{C} \qquad [12\text{-}1]^{8}$$

or $$0 = R\frac{dq}{dt} + \frac{q}{C} \qquad [12\text{-}2]$$

whence $$\frac{dq}{q} = -\left(\frac{1}{RC}\right)dt \qquad [12\text{-}3]$$

whose solution is $$\ln q = -\frac{t}{RC} + \text{constant} \qquad [12\text{-}4]$$

When the boundary conditions ($t = 0$, $q = q_0$) are applied to equation [12-4], the constant of integration is seen to be $\ln q_0$. Equation [12-4], rearranged, then becomes

$$q = q_0 e^{-t/RC} \qquad [12\text{-}5]$$

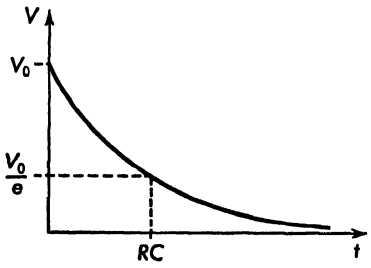


Fig. 12.2

Upon substitution of $q = CV$ and $q_0 = CV_0$ ($= C\mathcal{E}$) equation [12-5] has the alternative form

$$V = V_0 e^{-t/RC} \qquad [12\text{-}6]$$

Thus, when a charged capacitor is allowed to discharge through a resistor, the potential (or charge) of the capacitor decays with time in an exponential manner. This behavior is shown in Figure 12.2.

Time Constant

It is customary to describe the rate of discharge of the capacitor by the *time constant*, defined as *the time required for the potential (or charge) to decay to* 1/e

[8] The derivation of Kirchhoff's second law in § 5.8 was essentially based on the principle that the sum of the potential differences around any closed loop is zero. In that section, however, potential differences only manifested themselves in two ways: either as potential rises produced by electromotive forces or as potential drops across current-carrying resistors. Hence the law took the form

$$\sum_{1}^{n} \mathcal{E}_i = \sum_{1}^{m} I_j R_j \qquad [5\text{-}49]$$

Equation [12-1] differs from [5-49] in the fact that recognition is given to the potential drop ($V = q/C$) that occurs across a capacitance. There is the additional feature that equation [12-1] describes instantaneous behavior, but this is also in the spirit of equation [5-49], since Kirchhoff's laws must hold at all times.

of its initial value. It will be seen from equation [12-6] that when $t = RC$, the potential of the capacitor is in fact $V = V_0/e$. Thus the time constant of a resistance-capacitance circuit is RC seconds. This may be of appreciable duration if R, C, or both are sufficiently large. Thus, for example, if a $1\mu f$-capacitor is allowed to discharge through a 10^7-ohm resistor, its potential drops to $1/e(= 1/2.71)$ of its initial value in $RC = 10^7 \times 10^{-6} = 10$ seconds.

12.2 Determination of the Value of a High Resistance by the Rate of Discharge of a Capacitor

The measurement of the rate at which a capacitor discharges through a resistance provides a convenient method for determining the resistance value of large resistors. With the circuit employed for this purpose, shown in Figure 12.3, it is possible to determine the charge of the capacitor at various stages of the discharge.

When the switch is in position A, the capacitor is at full charge $q_0 = CV_0$ $(= C\mathcal{E})$. When the switch is thrown to position B, the capacitor discharges through the parallel combination of the resistor, R, and the ballistic galvanometer, G. If R is large, as in this application, practically the entire charge passes through the galvanometer, and the amount of the charge is indicated by the galvanometer throw. If, however, the switch is not moved directly from A to B but, instead, from A to the intermediate position, as in the diagram, the capacitor is free to discharge through the resistor alone. Then, after a known time, the switch may be thrown to position B to allow the amount of the charge remaining in the capacitor to be ascertained by discharging it through the galvanometer.

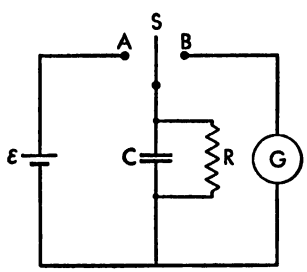


Fig. 12.3

To determine the value of a resistance, one makes a series of readings in which, in each case, the capacitor is given the initial charge $q_0 = CV_0$. The capacitor is then allowed to discharge through R for t seconds, at which time its reduced charge, q, is noted. With these data, the value of R may be calculated from equation [12-5]. Typical data obtained in such an experiment are given in Problem 12.3.

12.3 Charging a Capacitor Through a Resistor

This operation is indicated in Figure 12.4. When the switch is closed, charge flows in the circuit until the final potential difference, V_0, across the

capacitor is equal to the emf of the battery. At any stage in this operation

$$\mathcal{E} = iR + \frac{q}{C} = R\frac{dq}{dt} + \frac{q}{C} \qquad [12\text{-}7]^3$$

Since $\mathcal{E} = V_0 = q_0/C$, where q_0 is the final charge on the condenser, equation [12-7] may be rewritten:

$$\frac{q_0}{C} = R\frac{dq}{dt} + \frac{q}{C} \qquad [12\text{-}8]$$

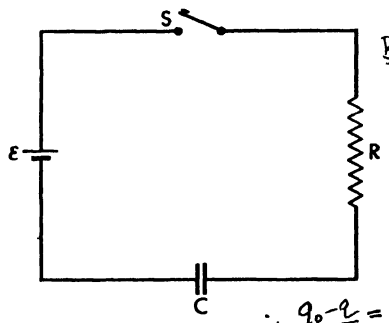


Fig. 12.4

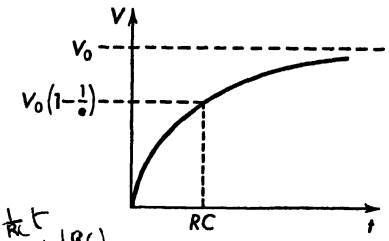


Fig. 12.5

The solution of equation [12-8], after proper cognizance has been taken of the boundary conditions, is easily shown to be

$$q = q_0(1 - e^{-t/RC}) \qquad [12\text{-}9]$$

or

$$V = V_0(1 - e^{-t/RC}) \qquad [12\text{-}10]$$

Thus, the potential of the condenser builds up as shown in Figure 12.5. The time constant, RC, here represents the time required for the potential to rise to a value V_0/e short of its ultimate value.

12.4 Relaxation Oscillator

In Figure 12.6, which shows the circuit for a simple *relaxation oscillator*, the glow tube, G, is a small discharge tube which "strikes" at a potential lower than the potential difference, V_0 ($= \mathcal{E}$), across the battery.

When the switch is closed, the potential across the capacitor builds up, according to equation [12-10], to the breakdown or *ignition* potential, V_i, of the glow tube, which then strikes as at point A in Figure 12.7. Since the resistance of the glow tube when discharging is negligible, the capacitor quickly discharges

[3] It is assumed here, and elsewhere in the chapter, that the internal resistance of the battery is negligible in comparison with the resistance, R.

through it along the route AB. When the capacitor potential drops to the extinction potential, V_e, of the glow tube, as at B in Figure 12.7, the discharge ceases and the capacitor begins to recharge along the path BC. Problem 12.4 is an exercise in calculating the frequency of such a relaxation oscillator.

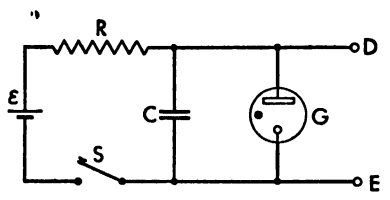


Fig. 12.6

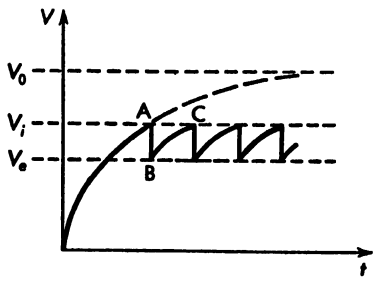


Fig. 12.7

The *saw-tooth* voltage output, which appears between terminals D and E in Figure 12.6, has been used as a sweep voltage for oscilloscopes, although the fact that BC is slightly curved, rather than straight, introduces some distortion of the wave form appearing on the oscilloscope screen. The device is also occasionally employed as an audio-frequency oscillator, despite the fact that its output is far from harmonic.

12.5 Circuits Containing Inductance and Resistance

The presence of inductance in a circuit, as in Figure 12.8, giving rise as it does to back emf's as the current changes, prevents the current in the circuit from immediately attaining its steady value.

Suppose the switch, S_1, in Figure 12.8, is closed at time $t = 0$. Applying Kirchhoff's second law to the circuit, we obtain

$$\mathcal{E} - L\frac{di}{dt} = iR \qquad [12\text{-}11]$$

which has the solution

$$i = i_0(1 - e^{-(R/L)t}) \qquad [12\text{-}12]$$

where $i_0 = \mathcal{E}/R$ is the final steady-state current in the circuit. Thus, the current in an LR circuit rises with time in the same manner as does the potential across a capacitor in an RC circuit (Fig. 12.5).

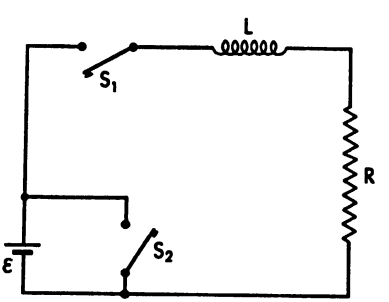


Fig. 12.8

Furthermore, it is easy to show that if S_2 is now closed (with complete disregard for the welfare of the battery), the current decays according to the law:

$$i = i_0 e^{-(R/L)t} \qquad [12\text{-}13]$$

corresponding to equation [12-6] and Figure 12.2 for the potential across a capacitor in an RC circuit. It will be noted, from equation [12-13], that the time constant in an LR circuit is L/R seconds.

12.6 Discharge of a Capacitor Through a Resistor and Inductor

This case is of particular interest since the LCR combination behaves as a damped oscillator. It is shown schematically in Figure 12.9.

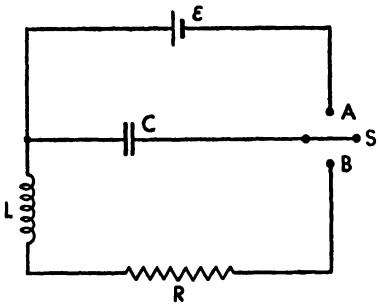


Fig. 12.9

When the switch, S_1, is in position A, the capacitor is charged to the potential V_0 ($= \mathcal{E}$). Then, after the switch is thrown to position B, the capacitor discharges through L and R. During this discharge the state of affairs in the LCR circuit, given by Kirchhoff's second law, is

$$-L\frac{di}{dt} = iR + \frac{q}{C} \qquad [12\text{-}14]$$

Substituting $i = dq/dt$ and $di/dt = d^2q/dt^2$, we have

$$\frac{d^2q}{dt^2} + \frac{R}{L}\frac{dq}{dt} + \frac{q}{LC} = 0 \qquad [12\text{-}15]$$

or

$$\ddot{q} + b\dot{q} + cq = 0 \qquad [12\text{-}16]$$

where

$$b = \frac{R}{L} \quad \text{and} \quad c = \frac{1}{LC} \qquad [12\text{-}17]$$

Equation [12-16] is identical in form with equation [11-30], which describes the motion of a damped galvanometer coil and which has been studied in detail in § 11.3. Instead of repeating the analysis here, we shall simply make use of the solutions there derived, and point out their physical significance in the present context.

As with equation [11-30], the nature of the solution of equation [12-16] is critically dependent upon the value of $b^2 - 4c$ [$= (R^2C^2 - 4LC)/L^2C^2$]. Three categories of solution exist.

Critically Damped Discharge

This occurs when $b^2 - 4c = 0$. Under these conditions the net charge on each of the plates of the capacitor decreases to zero as rapidly as possible without subsequently changing sign. That is, there is no overshoot. The value of the capacitor charge as a function of time is given by the analogue of equation [11-35]:

$$q = (\mathfrak{A} + \mathfrak{B}t)e^{-bt/2} \qquad [12\text{-}18]$$

This discharge would be represented by curve (*a*) in Figure 11.5 if q were plotted as the ordinate rather than θ.

Overdamped Discharge

This occurs when $b^2 - 4c$ is positive. As in the above case, the sign of the net charge on each of the capacitor plates remains unchanged during the discharge. However, the discharge does not take place so rapidly as when the circuit is critically damped. Instead, the time-dependence of the discharge is given by an expression similar to equation [11-38], viz.,

$$q = \mathfrak{A}e^{[(-b+\sqrt{b^2-4c})/2]t} + \mathfrak{B}e^{[(-b-\sqrt{b^2-4c})/2]t} \qquad [12\text{-}19]$$

A plot of q versus t would be similar to curve (*b*) in the θ versus t plot of Figure 11.5.

Underdamped Discharge

This effect occurs when $b^2 - 4c$ is negative. Under these conditions, the solution of equation [12-16], by analogy to equation [11-46], is given by

$$q = \mathfrak{F}e^{-bt/2}\cos(xt - \delta) \qquad [12\text{-}20]$$

where $x = \frac{1}{2}\sqrt{4c - b^2}$. Equation [12-20] represents a damped oscillatory discharge in which q, the capacitor charge, changes with time in the same manner as does θ in Figure 11.6. Since $V = q/c$, the voltage across the capacitor also alternates with time in a damped harmonic manner. An example of a damped oscillator of this type is given in § 12.8.

The angular frequency of the harmonic component of this damped oscillation of charge is given by

$$\omega = x \qquad [12\text{-}21]$$

and the frequency itself by

$$f = \frac{\omega}{2\pi} = \frac{x}{2\pi} = \frac{\sqrt{4c - b^2}}{4\pi} = \frac{\sqrt{4LC - R^2C^2}}{4\pi LC} \qquad [12\text{-}22]$$

If the resistance in the circuit is small, so that $4LC \gg R^2C^2$, the period of the oscillation is given by

$$T = \frac{1}{f} \doteq \frac{4\pi LC}{\sqrt{4LC}} = 2\pi\sqrt{LC} \qquad [12\text{-}23]$$

As with the underdamped galvanometer oscillation, the rate of decay of the oscillation of charge can be described in terms of the *logarithmic decrement*, Δ, defined as the natural logarithm of the ratio of the amplitudes of two successive charge maxima. Thus (cf. equation [11-52]),

$$\Delta = \frac{bT}{4} = \frac{RT}{4L} \qquad [12\text{-}24]$$

or, after substitution of the exact expression for the period (reciprocal of equation [12-22]),

$$\Delta = \frac{R}{4L}\frac{4\pi LC}{\sqrt{4LC - R^2C^2}} = \frac{\pi RC}{\sqrt{4LC - R^2C^2}} \qquad [12\text{-}25]$$

For values of R small compared to L/C the amplitude of the oscillation decays slowly and the oscillation persists for a long time. Under these conditions, the logarithmic decrement is given approximately by

$$\Delta \doteq \frac{\pi R}{2}\sqrt{\frac{C}{L}} \qquad [12\text{-}26]$$

which approaches zero in value as R becomes vanishingly small. The reciprocal of Δ is a measure of the persistence of oscillation, and is closely related to the *quality* or "Q" of the circuit, a term which is defined and discussed in Chapter 13.

For low-frequency oscillations, the heat dissipated in the resistance in the circuit is the chief means by which energy is lost. At higher frequencies additional energy is lost through radiation. The foregoing discussion neglects this latter type of loss.

12.7 Charge of a Capacitor Through a Resistor and Inductor

When the switch in Figure 12.10 is closed, the capacitor is charged through the resistor and inductor. Eventually, the potential across the capacitor becomes equal to the emf of the battery. At any stage during this charging process, Kirchhoff's second law may be written, viz.,

$$\mathcal{E} - L\frac{di}{dt} = iR + \frac{q}{C} \qquad [12\text{-}27]$$

This, written in terms of q and rearranged, becomes

$$\frac{d^2q}{dt^2} + \frac{R}{L}\frac{dq}{dt} + \frac{q}{LC} = \frac{\mathcal{E}}{L} \qquad [12\text{-}28]$$

or

$$\ddot{q} + b\dot{q} + cq = d \qquad [12\text{-}29]$$

where b and c have the same significance as in the preceding section, and $d = \mathcal{E}/L$.

Equation [12-29] is identical in form with equation [11-61], which we studied in detail in the preceding chapter. Making use of the results there obtained, we conclude that the process of charging a capacitor is also critically damped, overdamped, or underdamped when $b^2 - 4c$ is zero, positive, or negative, respectively.

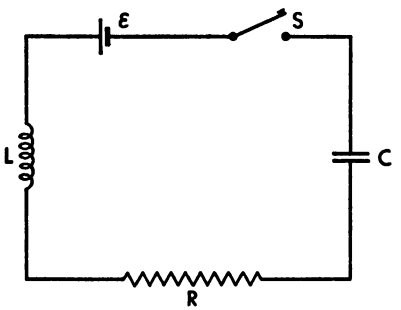


Fig. 12.10

In the case of the two first-named cases, the charging process is aperiodic, while in the case of the third it is oscillatory. The exact forms of the q versus t dependence can be written down by analogy to equations [11-64], [11-65], and [11-66]. In addition, Figure 11.8, with q substituted for θ, provides a graphical representation of the manner in which q varies with t in each of the three cases.

12.8 The Tesla Oscillator

A spectacular example of the underdamped oscillator discussed in § 12.6 is that devised by Nicolai Tesla, the Serbo-American electrical engineer who, among many other things, designed the Niagara power system. This oscillator is shown schematically in Figure 12.11.

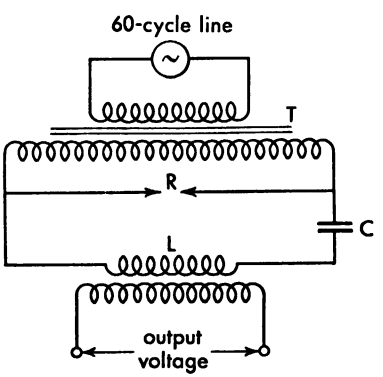


Fig. 12.11

The resistive element, R, is in this case a spark gap which breaks down when the voltage across it (supplied by the transformer, T) reaches a sufficiently high value. When a breakdown occurs, the value of R drops sharply and the LCR circuit oscillates, as in Figure 11.6, until the breakdown across the gap has ceased.

The high-frequency alternating volt-

age which appears across the inductor as a result of this oscillation may be stepped up as shown, and attain a very high value indeed. This sequence of operations occurs twice during each cycle of the line voltage.

Tesla oscillators are usually designed to operate at frequencies of a few megacycles/sec. In this frequency range energy is readily radiated, thereby facilitating corona discharge and contributing in large part to the high voltage sparks which are associated with this type of oscillator.

Essentially the same type of oscillator was used by Hertz (1888) to demonstrate the existence of electromagnetic waves and by Marconi (1895) as a transmitter for wireless telephony. In the latter case one of the output terminals was grounded and the other connected to an antenna from which the high-frequency energy was radiated. The ground lead contained a variable series inductance which could be tuned so as the bring the natural frequency of the antenna circuit into coincidence with that of the oscillating circuit. The capacitance between antenna and ground served as the capacitative element in the antenna circuit.

12.9 Problems

E (12.1) A 3-μf capacitor is charged to 271.8 volts and then discharged through a 1-megohm resistor. What is the voltage across the capacitor after three seconds? Recall that e = 2.718.
[Ans.: 100 volts.]

M (12.2) Calculate the heat dissipated in the resistor during the entire discharge described in Problem 12.1. Compare this value with the energy stored in the capacitor prior to the commencement of the discharge.

E (12.3) The capacitor discharge method is used to determine the value of a high resistance, as described in § 12.2. At various stages of the discharge, the deflections of the ballistic galvanometer are found to be

Time (sec)	*Deflection* (cm)
0.0	20.0
60	15.0
120	11.3
240	6.4
360	3.7
600	1.4

Plot these results on semi-log paper and determine therefrom the value of the resistance in question. A 1.00-μf capacitor is used. [Ans.: 2.1×10^8 ohms.]

E (12.4) Show for the relaxation oscillator of Figure 12.6 that the period of oscillation is given approximately by

$$T \doteq RC \ln \frac{V_0 - V_e}{V_0 - V_i}$$

E (12.5) Derive the expression for the current flowing through the resistance during the oscillatory discharge described by equation [12-20]. Show that the frequency of this current is the same as that of the charge oscillation.

Chapter 13

Alternating Currents

An alternating current is one that changes direction periodically. In practice, the most important alternating currents are those which alternate according to a harmonic law. Unless it is otherwise stated, we shall assume in this chapter that we are dealing with alternating currents of this type.[1]

13.1 The Simple AC Generator [2]

When a loop of wire rotates with constant angular velocity in a uniform magnetic field, as in Figure 13.1, the emf induced in the loop follows a harmonic law. This statement will now be proved.

The total magnetic flux, Φ_m, passing through the loop of wire at any stage during its rotation is given by

$$\Phi_m = \boldsymbol{B} \cdot \boldsymbol{A} \qquad [13\text{-}1]$$

where $\boldsymbol{B}$ is the induction associated with the uniform magnetic field and $\boldsymbol{A}$ is

[1] Actually, *any* periodic function can be written as a series of terms consisting of a fundamental and its harmonics, each of which follows a harmonic law. For this reason, results that are obtained for simple harmonic alternations can subsequently be applied to non-harmonic cases. A series of the type referred to is a Fourier Series.

[2] Although "AC" is literally an abbreviation for "alternating current," it has come to be used as an adjective that is synonymous with "alternating." Thus, alternating voltages and alternating currents are described as AC voltages and AC currents, respectively, despite the fact that the former usage is a literal incongruity and the latter is a literal redundancy. Similarly, "DC" is commonly used as a synonym of "direct."

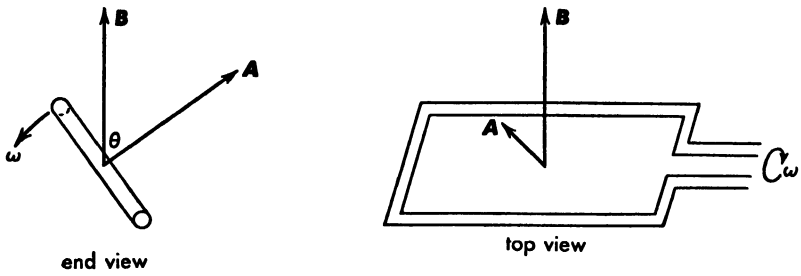


Fig. 13.1

the vector representing the area enclosed by the loop, thus,

$$\Phi_m = BA \cos \theta \qquad [13\text{-}2]$$

where θ is the angle between $\boldsymbol{B}$ and $\boldsymbol{A}$.

The instantaneous value of the induced emf is given by Faraday's law, viz.,

$$\varepsilon = -\frac{d\Phi_m}{dt} \qquad [9\text{-}4]$$

which becomes, upon substitution of equation [13-2],

$$\varepsilon = -\frac{d}{dt}(BA \cos \theta) = BA \sin \theta \frac{d\theta}{dt} = BA\omega \sin \theta \qquad [13\text{-}3]$$

Here ω $(= d\theta/dt)$, the angular velocity of the coil, is a constant.

We notice from equation [13-3] that the maximum value of the induced emf, which we shall write as $\mathcal{E}_0$, occurs when $\theta = \pi/2$ and is given by

$$\mathcal{E}_0 = BA\omega \qquad [13\text{-}4]$$

Also, at any time, t, the orientation, θ, of the coil is given by

$$\theta = \omega t + \delta \qquad [13\text{-}5]$$

where δ represents the value of θ at $t = 0$.

When equations [13-4] and [13-5] are substituted in equation [13-3], the expression for the instantaneous value of the emf induced in the rotating loop becomes

$$\varepsilon = \mathcal{E}_0 \sin (\omega t + \delta) \qquad [13\text{-}6]$$

Thus, as asserted above, the induced emf goes through harmonic alternations.

13.2 Graphical Representation of Alternating emf's and Currents

Equation [13-6] is plotted in Figure 13.2, as a solid curve, with ε as the ordinate and ωt as the abscissa. With reference to this curve, let us recall some of the terms that are used in the description of harmonic motion.

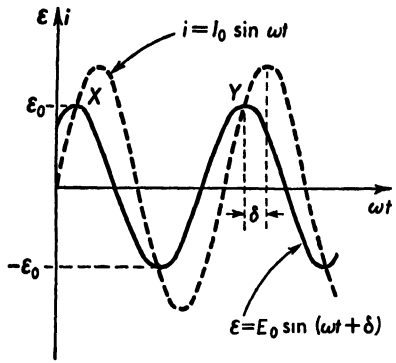


Fig. 13.2

As shown by the solid curve in Figure 13.2, the instantaneous value of the harmonic emf varies between the limits $\pm\mathcal{E}_0$, where $\mathcal{E}_0$ is known as the *amplitude* of the emf. The quantity δ appearing in equation [13-6] is the *initial phase angle:* thus, the initial value of ε (that is, at $t = 0$) is given by $\mathcal{E}_0 \sin \delta$. The phase of ε at any subsequent time is determined by the *instantaneous phase angle*, $\omega t + \delta$. One complete alternation of the emf is called one *cycle* and is accomplished in a time known as one *period*, T. The *frequency*, f, of the alternation, or the number of cycles executed in one second, is the reciprocal of the period. The symbol, ω, appearing in equation [13-6] represents the *angular frequency*. This can be readily verified by computing the time difference between successive maxima, X and Y, as we shall now do.

Suppose the maximum X occurs at $t = t_1$. Then,

$$\sin(\omega t_1 + \delta) = 1 \qquad [13\text{-}7]$$

or

$$\omega t_1 + \delta = (2n - 1)\frac{\pi}{2} \qquad [13\text{-}8]$$

where n is any odd integer. The next maximum, Y, occurs at an angle which is 2π larger, and a time which is a period later. Hence

$$\omega(t_1 + T) + \delta = (2n - 1)\frac{\pi}{2} + 2\pi \qquad [13\text{-}9]$$

where n has the same value as in equation [13-8]. Subtracting equation [13-8] from [13-9], we have

$$\omega T = 2\pi \qquad [13\text{-}10]$$

or

$$T = \frac{2\pi}{\omega} \qquad [13\text{-}11]$$

Also, it follows that $$f = \frac{1}{T} = \frac{\omega}{2\pi} \qquad [13\text{-}12]$$

or $$\omega = 2\pi f \qquad [13\text{-}13]$$

thus establishing that ω is the angular frequency of the alternating emf.

When an AC emf, of the type described by equation [13-6], is included in a circuit, the resulting current is also harmonic, although not necessarily in phase with the emf. This assertion is proved in § 13.5 *et seq.* Let us suppose that the initial conditions and circuit parameters are such that the instantaneous current and emf may be written

$$i = I_0 \sin(\omega t - \delta_1) \qquad [13\text{-}14]$$

and $$\varepsilon = \mathcal{E}_0 \sin(\omega t + \delta_2) \qquad [13\text{-}15]$$

where δ_1 and δ_2 are both positive. Here the emf passes through each phase at an earlier time than does the current. It is said to "lead" the current by the phase angle

$$\delta = \delta_2 - (-\delta_1) = \delta_2 + \delta_1 \qquad [13\text{-}16]$$

It is this quantity, the *relative phase angle* between the emf and the current, which is of practical importance. For this reason, it is customary to choose the zero of time so as to make either δ_2 or δ_1 equal to zero.[3] For example, if by this device $\delta_1 = 0$, the instantaneous current and emf would be written

$$i = I_0 \sin \omega t \qquad [13\text{-}17]$$

and $$\varepsilon = \mathcal{E}_0 \sin(\omega t + \delta) \qquad [13\text{-}18]$$

This combination is somewhat neater than equations [13-14] and [13-15] and, in addition, the relative phase angle between the current and the emf is more simply expressed. Equation [13-17], representing the current, is shown in Figure 13.2 as a dotted curve. As required by equations [13-17] and [13-18], it "lags" the emf curve by the relative phase angle, δ.[4] Hereafter, in speaking of the relative phase angle between emf and current, the adjective "relative" will usually be omitted.

---

[3] If the loop in Figure 13.1 is horizontal ($\theta = 0$) at $t = 0$, one can see from equation [13-5] that $\delta = 0$. In this way, by proper choice of the zero of time, the phase angle associated with the emf can be made to vanish. Alternatively, we could define the zero of time so as to make the phase angle of the current vanish.

[4] Although we shall occasionally express phase angles in degrees, the reader should not forget that δ in equations like [13-18] is measured in radians.

13.3 Definition of Unit AC Current

The instantaneous value of an AC current has been discussed above. For many applications it is desirable to express in some way an "effective" behavior of such a current. Clearly a simple average is not useful, since it has the value zero for any complete cycle. On the other hand, the heating effect of an AC current does not change sign as the direction of the current changes. For this reason, the magnitude of an AC current is expressed in terms of the heat which it produces in a resistance. Thus, 1 ampere of AC current is defined as that current which produces the same average heating effect as 1 ampere of DC current.

The DC equivalent of an AC current is known as the *effective* value, I_{eff}, of the AC current.

13.4 Effective or RMS Value of an AC Current

We shall now calculate the relationship between the effective value of an AC current and its maximum value. This is done, as suggested above, by computing the *average* heating effect of the alternating current.

The rate of heat production in a current-carrying resistance is given at any instant by

$$\text{heat/sec} = i^2R \qquad \text{[5-37]}$$

We are interested in the average or mean value of this quantity, which we shall write as $\overline{i^2}R$, where $\overline{i^2}$ represents the mean value of i^2.

By definition,

$$I_{\text{eff}}^2R = \overline{i^2}R \qquad \text{[13-19]}$$

or

$$I_{\text{eff}} = \sqrt{\overline{i^2}} \qquad \text{[13-20]}$$

Hence, the effective value of an AC current is the square root of the mean of the squares of the instantaneous values. For this reason, the effective current is frequently termed the RMS (root-mean-square) current.

The value of $\overline{i^2}$ is calculated in the manner suggested in § 4.11 in the "Digression Concerning Average Values" :

$$\overline{i^2} = \frac{\int_0^T i^2\,dt}{\int_0^T dt} = \frac{\int_0^T i^2\,dt}{T} \qquad \text{[13-21]}$$

As indicated, an appropriate interval over which to calculate this time average is one period. A larger number of periods is not necessary since each is a repetition of the preceding. On the other hand, no fraction of a period (except $T/2$) is suitable since it would not provide a representative picture of the variation of i^2.

Substituting equation [13-17] in equation [13-21], we have

$$\overline{i^2} = \frac{I_0^2 \int_0^T \sin^2 \omega t \, dt}{T} \qquad [13\text{-}22]$$

Making use of the identity

$$\sin^2 \omega t = \tfrac{1}{2}(1 - \cos 2\omega t) \qquad [13\text{-}23]$$

we can evaluate equation [13-22], viz.,

$$\overline{i^2} = \frac{I_0^2}{2T}\left[t - \frac{\sin 2\omega t}{2\omega}\right]_0^T = \frac{I_0^2}{2T} T \qquad [13\text{-}24]$$

or

$$\overline{i^2} = \frac{I_0^2}{2} \qquad [13\text{-}25]$$

It will be seen that, in equation [13-24], the term $\sin 2\omega t$ vanishes for both limits. For the upper limit, $\sin 2\omega t = \sin 2\omega T = \sin 4\pi f(1/f) = \sin 4\pi = 0$, while, of course, for the lower limit, $\sin 2\omega t = \sin 0 = 0$.

Combining equations [13-20] and [13-25], we have

$$I_{\text{eff}} = \sqrt{\overline{i^2}} = \frac{I_0}{\sqrt{2}} = 0.707 I_0 \qquad [13\text{-}26]$$

or, in words, the effective value of an AC current is 0.707 times its amplitude or maximum value.

Effective and maximum AC emf's are similarly related, viz.,

$$\mathcal{E}_{\text{eff}} = \frac{\mathcal{E}_0}{\sqrt{2}} = 0.707 \mathcal{E}_0 \qquad [13\text{-}27]$$

This can easily be shown by considering a circuit containing only $\mathcal{E}$ and R. The above calculation involving the current can then be carried out with $\mathcal{E}/R$ substituted for the instantaneous current. A similar relationship exists between the effective value of a potential drop and its maximum value.

AC ammeters and voltmeters, unless it is specifically stated to the contrary, read effective or RMS values. As a result, effective values are the ones normally

dealt with. From this point on, therefore, effective currents, emf's, and potential drops will be written with the "effective" subscript omitted: that is $I_{\text{eff}} = I$, $\mathcal{E}_{\text{eff}} = \mathcal{E}$ and $V_{\text{eff}} = V$.

13.5 AC Current and Voltage in a Series LR Circuit

Let us assume that the instantaneous current in the circuit of Figure 13.3 is given by equation [13-17]:

$$i = I_0 \sin \omega t \qquad [13\text{-}17]$$

At any instant, according to Kirchhoff's second law,

$$\mathcal{E} = L\frac{di}{dt} + iR \qquad [13\text{-}28]^{6}$$

When the expression for the current is substituted in this equation and the resulting equation rearranged, we have

$$\mathcal{E} = I_0(R \sin \omega t + \omega L \cos \omega t) \qquad [13\text{-}29]$$

Let us express the constants ωL and R in terms of the two new constants, Z and δ. Thus,

$$R = Z \cos \delta \qquad [13\text{-}30]$$

and

$$\omega L = Z \sin \delta \qquad [13\text{-}31]$$

When these substitutions are made, equation [13-29] takes the form

$$\mathcal{E} = I_0 Z(\sin \omega t \cos \delta + \cos \omega t \sin \delta) \qquad [13\text{-}32]$$

or

$$\mathcal{E} = I_0 Z \sin (\omega t + \delta) \qquad [13\text{-}33]$$

a result which is similar in form to the general expression which we earlier wrote for the instantaneous emf in the circuit, viz.,

$$\mathcal{E} = \mathcal{E}_0 \sin (\omega t + \delta) \qquad [13\text{-}18]$$

We draw two conclusions from equation [13-33]. First, by comparing equations [13-33] and [13-18], we see that

$$\mathcal{E}_0 = I_0 Z \qquad [13\text{-}34]$$

or

$$\mathcal{E} = IZ \qquad [13\text{-}35]$$

---

[6] Unless a statement is made to the contrary, we shall assume in this and following sections that the resistance of the generator is negligible. Hence, equation [13-28] contains no term representing the iR drop across the generator. Also, under these conditions, the potential difference across the LR combination is numerically equal to the emf of the generator.

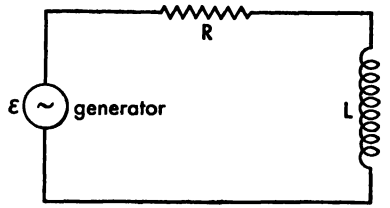


Fig. 13.3

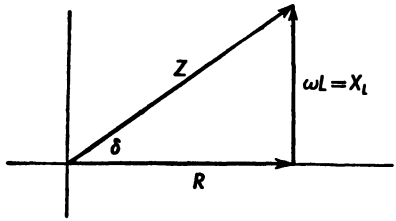


Fig. 13.4

This expression is similar in form to equation [5-45],

$$\mathcal{E} = IR \qquad [5\text{-}45]$$

relating the current in a circuit to the circuit resistance and the impressed emf, except that the resistance, R, appearing in equation [5-45], is replaced in [13-35] by the new constant, Z, defined above. We therefore conclude that a relation of the type of [5-45] is applicable in calculating the value of an alternating current in a series LR circuit provided the resistance is replaced by Z. Z is known as the *impedance* of the circuit. Its relationship to the other circuit parameters implicit in equations [13-30] and [13-31] will be discussed below.

The second conclusion concerns the relative phases of the current and emf in the series LR circuit. It is evident from equation [13-33] that the constant, δ, is a relative phase angle. Furthermore, by comparing the expressions for the instantaneous current and emf (equations [13-17] and [13-33]), it is seen that in a series circuit consisting of inductance and resistance *the emf leads the current* at all times *by the phase angle* δ.

The relationships between R, ωL, Z, and δ, given in equations [13-30] and [13-31], are shown graphically in Figure 13.4. As indicated, the quantity ωL is frequently written as X_L. It is called the *inductive reactance*. Evidently, Z and δ are given by the following expressions:

$$Z = \sqrt{R^2 + \omega^2 L^2} = \sqrt{R^2 + X_L^2} \qquad [13\text{-}36]$$

and

$$\delta = \tan^{-1}\left(\frac{\omega L}{R}\right) = \tan^{-1}\left(\frac{X_L}{R}\right) \qquad [13\text{-}37]$$

The impedance, Z, represents the total opposition, both inductive and resistive, which the circuit offers to an alternating current. The two components of the impedance must be added vectorially, as shown in Figure 13.4 and indicated in equation [13-36]. The impedance and its inductive component, the inductive reactance, have the same units as resistance, namely, ohms. Figure 13.4, which shows the manner in which the resistive and reactive components of the

impedance combine to form the impedance itself, is known as an *impedance diagram.*

Components of the Total emf and Their Phases Relative to the Current

We shall show that the impressed emf in an LR circuit consists of two components. These, together with total impressed emf and the phases of all three relative to the current, are shown graphically in Figure 13.5. The current is, of course, identical in all parts of the circuit.

The relationship between the current and the potential drop in the resistive portion of the circuit is given at all times by Ohm's law. Hence, the instan-

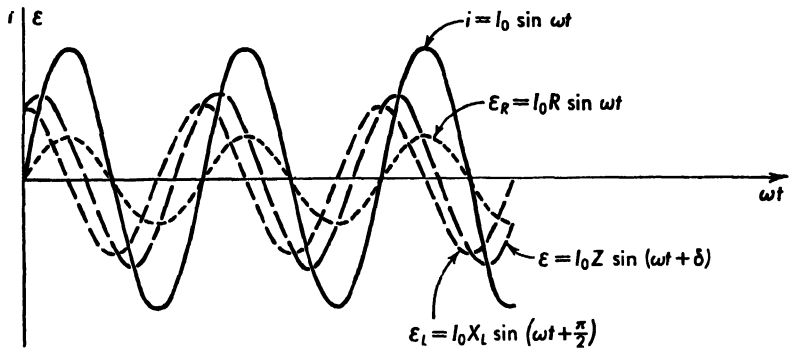


Fig. 13.5

taneous current in the resistance and the instantaneous voltage drop across it, v_R, are related as follows:

$$i = \frac{v_R}{R} \qquad [13\text{-}38]$$

or

$$v_R = iR = I_0 R \sin \omega t \qquad [13\text{-}39]$$

According to Kirchhoff's second law (equation [5-49]) the existence of this potential drop necessitates the presence in the circuit of an emf of the same value. Consequently, in order to maintain the current [13-17] in the resistive portion of the LR circuit, an emf

$$\varepsilon_R = I_0 R \sin \omega t \qquad [13\text{-}40]$$

must be present in the circuit. This particular component of the total impressed emf is in phase with the current, as shown in Figure 13.5.

Consider now the emf that is responsible for current in the inductive element in the circuit. The back emf generated in the inductance as a result

of the changing current is

$$-L\frac{di}{dt} = -I_0\omega L\cos\omega t = -I_0X_L\cos\omega t \qquad [13\text{-}41]$$

This back emf must be counterbalanced by an equal and opposite impressed emf if the current is to continue. This impressed emf must, therefore, have the value

$$\varepsilon_L = I_0X_L\cos\omega t \qquad [13\text{-}42]$$

Since $\cos\omega t = \sin(\omega t + \pi/2)$, we can rewrite [13-42] as

$$\varepsilon_L = I_0X_L\sin\left(\omega t + \frac{\pi}{2}\right) \qquad [13\text{-}43]$$

Hence, this component of the total generator emf leads the current by $\pi/2$ radians or 90°. This fact is shown in Figure 13.5 in the plot of ε_L.

Equations [13-40] and [13-42], when added, give the total impressed emf:

$$\varepsilon = \varepsilon_R + \varepsilon_L \qquad [13\text{-}44]$$

or

$$\varepsilon = I_0R\sin\omega t + I_0X_L\cos\omega t \qquad [13\text{-}45]$$

When equations [13-30] and [13-31] are substituted in this expression, it becomes

$$\varepsilon = I_0Z(\sin\omega t\cos\delta + \cos\omega t\sin\delta) \qquad [13\text{-}46]$$

or

$$\varepsilon = I_0Z\sin(\omega t + \delta) \qquad [13\text{-}33]$$

the expression that we have already obtained for the total applied emf. Since the total impressed emf is the sum of two components, one which is in phase with the current and the other which leads the current by 90°, we conclude that the total impressed emf, ε, in an LR circuit leads the current by a phase angle, δ, which lies between 0° and 90°. This is shown in Figure 13.5. As we have noted above, the value of δ is given by

$$\delta = \tan^{-1}\left(\frac{X_L}{R}\right) \qquad [13\text{-}37]$$

If $R \to 0$, $\delta \to \pi/2$: that is, in a circuit containing pure inductance, the impressed emf leads the current by 90°. If $L \to 0$, $\delta \to 0$: that is, in a circuit containing pure resistance, the impressed emf is in phase with the current.

Voltage Diagrams

As we have several times observed in applying Kirchhoff's second law to a DC circuit, the presence of an emf in a closed circuit results in a difference in

potential. In DC circuits the differences in potential appear as IR drops, viz.,

$$\sum_{1}^{n} \mathcal{E}_i = \sum_{1}^{m} I_j R_j \qquad [5\text{-}49]$$

Although the situation is not as simple in AC circuits because of the differences in phase that can exist among the various impressed emf's and the various potential differences, one can still apply Kirchhoff's second law, as we have done, for example, in obtaining the equation describing the LR circuit of Figure 13.3. Thus,

$$\mathcal{E} - L\frac{di}{dt} = iR \qquad [13\text{-}28]$$

or

$$\mathcal{E} = iR + L\frac{di}{dt} \qquad [13\text{-}47]$$

Here iR is the instantaneous potential difference, v_R, across the resistance. When $-L\,di/dt$, the back emf, is moved to the right side of the equation it appears as an instantaneous potential difference, v_L, across the inductance.

Normally, we are more interested in the potential differences that are measured by voltmeters than in the instantaneous values appearing on the right-hand side of equation [13-47]. In short, we are interested in effective values of potential difference. We shall now work out such values for the LR circuit under consideration, and show how their vector characters can be presented graphically.

We note from equation [13-39] that the instantaneous potential difference across the resistance is

$$v_R = I_0 R \sin \omega t \qquad [13\text{-}39]$$

and is in phase with the current. The effective value of this potential difference is

$$V_R = \frac{I_0 R}{\sqrt{2}} = IR \qquad [13\text{-}48]$$

Likewise the instantaneous potential difference across the inductance is seen from equation [13-43] to be

$$v_L = I_0 X_L \sin\left(\omega t + \frac{\pi}{2}\right) \qquad [13\text{-}43]$$

This potential difference leads the current by 90° and has an effective value

$$V_L = \frac{I_0 X_L}{\sqrt{2}} = IX_L \qquad [13\text{-}49]$$

Similarly, the total potential difference across the LR combination is seen from

equation [13-33] to be

$$v = I_0 Z \sin(\omega t + \delta) \qquad [13\text{-}50]$$

and to have an effective value

$$V = \frac{I_0 Z}{\sqrt{2}} = IZ \qquad [13\text{-}51]$$

It is customary to present the relationships between the potential differences on a *voltage diagram* (Fig. 13.6). In this diagram the voltage drops which are in phase with the current are drawn as vectors pointing in the positive X-direction. Hence, the voltage drop ($V_R = IR$) across a pure resistance is so represented. Other voltage drops are drawn as vectors which are inclined to the X-axis at the angle by which they lead the current. Hence, the voltage drop ($V_L = IX_L$) across a pure inductance is represented by a vector pointing in the positive Y-direction (that is, inclined at an angle of $+\pi/2$ radians to the X-axis). The vector representing the total voltage ($V = IZ$) is the vector sum of the other two and is, consequently, inclined to the X-axis at some angle $0 < \delta < \pi/2$. The triangle formed by these three vectors is similar to the impedance triangle shown in Figure 13.4. Each side of the former is simply I times the corresponding side of the latter.

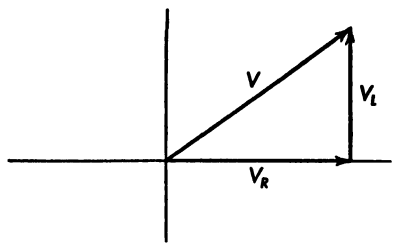


Fig. 13.6

Neglect of Transient Currents

We have assumed in the foregoing discussion that the current has the steadily alternating value

$$i = I_0 \sin \omega t \qquad [13\text{-}17]$$

The reader will realize, from what has been said in the previous chapter, that certain transient effects occur when an emf is first applied to a circuit. These effects are deliberately neglected by assuming the existence of the above steady-state condition.

The correct, but less simple, procedure would be to solve equation [13-28], which was obtained by applying Kirchhoff's second law to the circuit consisting of the harmonic generator, the resistance, and the inductance. This equation is

$$\varepsilon = L\frac{di}{dt} + iR \qquad [13\text{-}28]$$

We can write the generator emf in a standard harmonic form,

$$\varepsilon = \mathcal{E}_0 \sin \omega t \qquad [13\text{-}52]$$

to obtain

$$\mathcal{E}_0 \sin \omega t = L\frac{di}{dt} + iR \qquad [13\text{-}53]$$

The solution of this equation, which we shall not work out here, is

$$i = \frac{\mathcal{E}_0}{\sqrt{R^2 + \omega^2 L^2}} \sin(\omega t - \delta) + Ae^{-Rt/L} \qquad [13\text{-}54]$$

where δ, the phase angle, has its usual value, and A is a constant whose value depends upon the initial conditions. Clearly, for a circuit with a small time constant, L/R (cf. § 12.5), the second term on the right-hand side of equation [13-54] decays rapidly. Its effect can be neglected unless one is especially interested in the effects that occur very soon after the harmonic emf is applied to the LR combination. This we are not, and so our neglect of transient effects in this chapter is justified. Transient currents due to the presence of capacitance in the circuit will similarly be ignored.

13.6 AC Current and Voltage in a Series RC Circuit

This circuit, shown in Figure 13.7, can be treated in much the same way as was the LR circuit of the previous section. Applying Kirchhoff's second law to it, we obtain

$$\varepsilon = iR + \frac{q}{C} \qquad [13\text{-}55]$$

If, as before, we assume

$$i = I_0 \sin \omega t \qquad [13\text{-}17]$$

we have

$$\frac{dq}{dt} = i = I_0 \sin \omega t \qquad [13\text{-}56]$$

or

$$dq = I_0 \sin \omega t \, dt \qquad [13\text{-}57]$$

whence

$$q = -\frac{I_0}{\omega} \cos \omega t + \text{constant of integration} \qquad [13\text{-}58]$$

The constant of integration must be zero since otherwise the capacitor would possess a fixed charge in addition to its fluctuating one. Equation [13-55] then takes the form:

$$\varepsilon = I_0 R \sin \omega t - \frac{I_0}{\omega C} \cos \omega t \qquad [13\text{-}59]$$

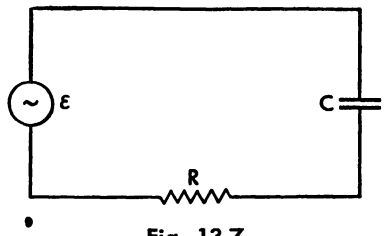


Fig. 13.7

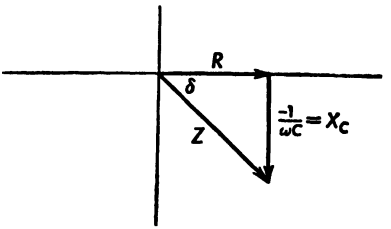


Fig. 13.8

Writing

$$R = Z \cos \delta \qquad \text{[13-30]}$$

and

$$-\frac{1}{\omega C} = X_C = Z \sin \delta \qquad \text{[13-60]}$$

we can express equation [13-59] in the form

$$\varepsilon = I_0 Z(\sin \omega t \cos \delta + \cos \omega t \sin \delta) \qquad \text{[13-61]}$$

or

$$\varepsilon = \mathcal{E}_0 \sin (\omega t + \delta) \qquad \text{[13-62]}$$

One might at first conclude, after comparing equations [13-17] and [13-62], that the impressed emf in an RC circuit leads the current, as in the LR circuit. However, as we shall now see, the phase angle δ has a negative value in the RC circuit, and the impressed emf actually lags the current rather than the reverse.

Definitions [13-30] and [13-60] lead to the impedance diagram shown in Figure 13.8. The quantity $-1/\omega C = X_C$ is known as the *capacitative reactance*. Evidently, the impedance of the series RC combination is given by

$$Z = \sqrt{R^2 + \frac{1}{\omega^2 C^2}} = \sqrt{R^2 + X_C^2} \qquad \text{[13-63]}$$

and the phase angle by which the emf leads the current is given by

$$\delta = \tan^{-1} \frac{\left(-\frac{1}{\omega C}\right)}{R} = \tan^{-1} \left(\frac{X_C}{R}\right) \qquad \text{[13-64]}$$

The phase angle is thus negative, as stated above, so that the current leads the emf.

Across the capacitor alone, the value of the series resistance is zero. Hence

$$\delta = \tan^{-1} \frac{\left(-\frac{1}{\omega C}\right)}{0} = \tan^{-1}(-\infty) = -\frac{\pi}{2} \qquad \text{[13-65]}$$

or the difference in potential lags the current by 90°. The physical reason for this phase shift is easily understood. When the potential difference across the capacitor is zero (no charge on the condenser), the current flowing to the capacitor has its maximum value. As the charge builds up in the capacitor, however, the current falls off, and becomes zero when the voltage across the capacitor reaches its maximum value. For this reason the potential difference across the capacitor, v_C, lags the current by 90°. Across the resistance alone,

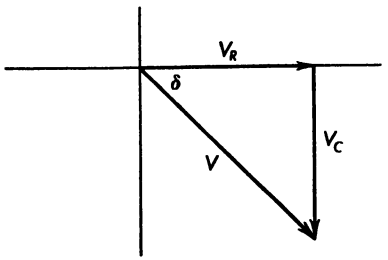


Fig. 13.9

$$\delta = \tan^{-1}(0) = 0 \qquad [13\text{-}66]$$

or the potential drop is in phase with the current. Obviously the potential difference across the RC combination lags the current by some phase angle which lies in the range $-\frac{\pi}{2} < \delta < 0$.

The voltage diagram for the series RC circuit is shown in Figure 13.9. The lengths of the three effective voltages shown are $V_R = IR$, $V_C = I/\omega C = IX_C$, and $V = IZ$. The vector representing V_C points in the negative Y-direction since the voltage across the capacitor "leads" the current by the angle $-\pi/2$.

13.7 AC Current and Voltage in a Series LCR Circuit

Making use of the results of the two previous sections, we may write for the circuit of Figure 13.10

$$\varepsilon = L\frac{di}{dt} + iR + \frac{q}{C} \qquad [13\text{-}67]$$

which becomes, upon substitution of $i = I_0 \sin \omega t$,

$$\varepsilon = I_0\omega L \cos \omega t + I_0 R \sin \omega t - \frac{I_0}{\omega C}\cos \omega t \qquad [13\text{-}68]$$

or

$$\varepsilon = I_0 R \sin \omega t + I_0\left(\omega L - \frac{1}{\omega C}\right)\cos \omega t \qquad [13\text{-}69]$$

By substituting

$$R = Z \cos \delta \qquad [13\text{-}30]$$

and

$$\left(\omega L - \frac{1}{\omega C}\right) = (X_L + X_C) = X = Z \sin \delta \qquad [13\text{-}70]$$

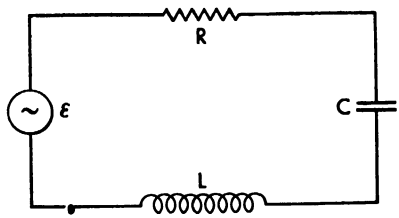


Fig. 13.10

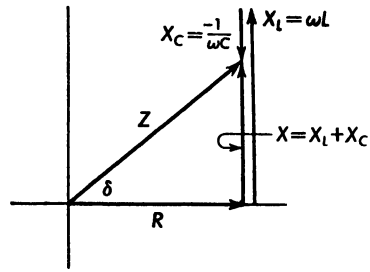


Fig. 13.11

we may rewrite equation [13-68] in the form

$$\varepsilon = I_0 Z \sin(\omega t + \delta) = \mathcal{E}_0 \sin(\omega t + \delta) \qquad [13\text{-}71]$$

where

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2} = \sqrt{R^2 + (X_L + X_C)^2}$$

$$= \sqrt{R^2 + X^2} \qquad [13\text{-}72]$$

and

$$\delta = \tan^{-1} \frac{\omega L - \frac{1}{\omega C}}{R} = \tan^{-1} \frac{(X_L + X_C)}{R}$$

$$= \tan^{-1}\left(\frac{X}{R}\right) \qquad [13\text{-}73]$$

as shown graphically in the impedance diagram, Figure 13.11.

The quantity, $X = X_L + X_C = \omega L - \frac{1}{\omega C}$, which appears in the above expressions, is the *reactance* of the circuit. The emf leads or lags the current depending upon whether X is positive or negative. Equations [13-72] and [13-73] are general expressions for the impedance and phase angle, respectively, of a series circuit. The corresponding expressions for series LR and RC circuits, which were given in the two preceding sections, will be recognized as special cases of [13-72] and [13-73].

Finally, the voltage diagram for the series LCR circuit is shown in Figure 13.12. The effective voltage drop across each of the circuit elements is given by the product

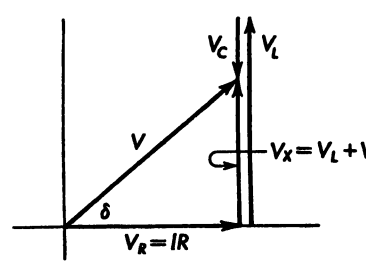


Fig. 13.12

of the effective current and the impedance of the element. Thus, across R, L, and C, the effective potential differences are IR, IX_L, and IX_C, respectively, whereas the total potential difference across the entire combination is $V = IZ$.

Illustrative Example: In the circuit of Figure 13.13, the applied AC emf has a frequency of 60 cycles/sec. Compute the readings of the voltmeter, V, and the ammeter, A, and

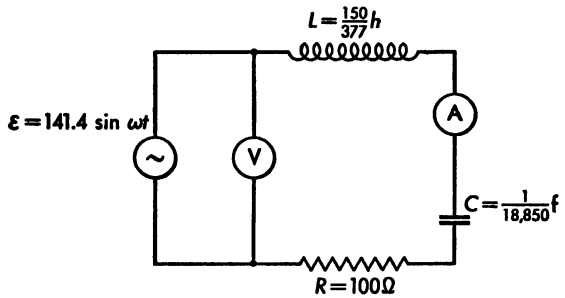


Fig. 13.13

write the expression for the instantaneous current in the circuit. Assume the impedance of the generator to be negligibly small.

a. The angular frequency corresponding to 60 cycles is

$$\omega = 2\pi(60) = 377 \text{ radians/sec}$$

b. The voltmeter, V, reads the effective value of the potential difference across the LCR combination. This is equal to the effective value of the impressed emf. Thus,

$$V = \frac{V_0}{\sqrt{2}} = \frac{\mathcal{E}_0}{\sqrt{2}} = \frac{141.4}{\sqrt{2}} = 100 \text{ volts}$$

c. $$X_L = \omega L = 377 \times \frac{150}{377} = 150 \text{ ohms}$$

$$X_C = -\frac{1}{\omega C} = -\frac{18{,}850}{377} = -50 \text{ ohms}$$

$$Z = \sqrt{R^2 + (X_L + X_C)^2} = \sqrt{100^2 + (150 - 50)^2} = 100\sqrt{2} \text{ ohms}$$

d. Ammeter A reads the effective current. This is computed from a knowledge of the potential drop across, and the impedance of, the LCR combination. Thus:

$$I = V/Z = \frac{100}{100\sqrt{2}} = 0.707 \text{ amperes}$$

e. The phase angle between the voltage and current

$$\delta = \tan^{-1}\left(\frac{X}{R}\right) = \tan^{-1}\left(\frac{100}{100}\right) = \frac{\pi}{4}$$

f. Since δ is positive, the impressed emf leads the current. Hence,

$$i = I_0 \sin(\omega t - \delta)$$

$$= \sqrt{2}\,(0.707) \sin\left(377t - \frac{\pi}{4}\right)$$

$$= \sin\left(377t - \frac{\pi}{4}\right) \text{ amperes}$$

13.8 Power Dissipation in an AC Circuit

At any instant the rate of energy loss in a circuit element is given by the product of the current through it and the potential difference across it (cf. equation [5-36]).

$$\text{Instantaneous power loss} = iv \qquad [13\text{-}74]$$

$$\text{Average power loss} = \overline{iv} \qquad [13\text{-}75]$$

$$= \frac{1}{T}\int_0^T iv\,dt \qquad [13\text{-}76]$$

$$= \frac{1}{T}\int_0^T I_0 \sin \omega t\, V_0 \sin(\omega t + \delta)\,dt \qquad [13\text{-}77]$$

$$= \frac{I_0 V_0}{T}\int_0^T \sin \omega t\,[\sin \omega t \cos \delta + \cos \omega t \sin\delta]\,dt \qquad [13\text{-}78]$$

$$= \frac{I_0 V_0}{T}\int_0^T \left[\frac{(1 - \cos 2\omega t)}{2}\cos \delta + \sin \omega t \cos \omega t \sin \delta\right] dt \qquad [13\text{-}79]$$

$$= \frac{I_0 V_0}{T}\left[\frac{t}{2}\cos \delta - \frac{\sin 2\omega t \cos \delta}{4\omega} + \frac{\sin^2 \omega t \sin \delta}{2\omega}\right]_0^T \qquad [13\text{-}80]$$

$$= \frac{I_0 V_0}{2}\cos \delta \qquad [13\text{-}81]$$

$$= IV \cos \delta \qquad [13\text{-}82]$$

Cos δ, which appears in equation [13-82], is known as the *power factor*. For purely resistive circuits, in which $\delta = 0$, the power factor is unity and the power is given by the product of the effective current and the effective voltage. For purely reactive circuits, in which $\delta = \pm\pi/2$, the power factor is zero, as is also the power loss.

It will be noted in Figure 13.12 that $V \cos \delta = V_R = IR$. An alternative expression for the power loss is, therefore,

$$\text{Power} = I^2R \qquad [13\text{-}83]$$

the familiar expression for the heat generated in a resistance as a result of current flow. Thus, in any circuit, whether it contains reactive elements or not, the only power loss occurs in the resistive portions, and has the value I^2R watts. The effect of reactance is to provide emf's ($-L\,di/dt$) or charge concentrations (q of the capacitance) which oppose the applied emf, in this way limiting the current without energy dissipation. Inductances that are used to control current, say for dimming theater lights, are called *choke coils*, and can, if substituted for resistances, effect a substantial saving in power (see Problem 13.10).

In AC machinery it is important to know not only the power consumption but also the voltage and current requirements. For this reason the volt-ampere or kilovolt-ampere (KVA) rating of a piece of machinery is usually given, reserving the term watt or kilowatt (KW) for the actual power dissipation. As in equation [13-82], the latter (KW) is given by the product of the former (KVA) and the power factor. Reference is also occasionally made to the wattless component of volt-amperes, known, rather inappropriately, as the *wattless power*.

$$\text{Wattless power} = IV \sin \delta \qquad [13\text{-}84]$$

13.9 Resonance in Series Circuits

Let us recall, from § 13.7, the expressions for the impedance and phase angle in a series LCR circuit:

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2} = \sqrt{R^2 + (X_L + X_C)^2} = \sqrt{R^2 + X^2} \qquad [13\text{-}72]$$

and

$$\delta = \tan^{-1} \frac{\left(\omega L - \dfrac{1}{\omega C}\right)}{R} = \tan^{-1} \frac{(X_L + X_C)}{R} = \tan^{-1} \frac{X}{R} \qquad [13\text{-}73]$$

The unique situation which exists when $\omega L = 1/\omega C$ is known as *resonance*. Evidently, series resonance may be defined in a variety of ways:

a. the reactance is zero ($X = X_L + X_C = 0$);
b. the impedance is a minimum;
c. the circuit acts as a pure resistance ($Z = R$);
d. for a given impressed emf the current in the circuit is a maximum;

e. the phase angle is zero ($\delta = \tan^{-1} 0 = 0$);
f. the current and voltage are in phase ($\delta = 0$);
g. the power factor is unity ($\cos \delta = \cos 0 = 1$).

Of these several equivalent definitions the most useful to keep in mind, as a *general* definition of resonance, is f.: the current and voltage are in phase.

A circuit may be brought into resonance by varying L, C, or ω. The series resonant frequency is given by

$$\omega_r L = \frac{1}{\omega_r C} \qquad [13\text{-}85]$$

or

$$f_r = \frac{1}{2\pi\sqrt{LC}} \qquad [13\text{-}86]$$

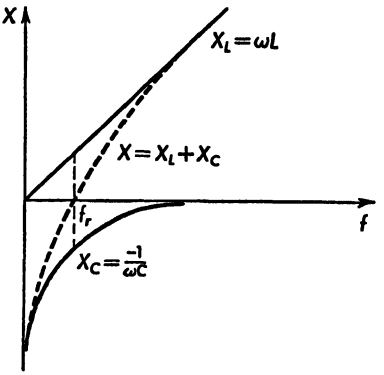


Fig. 13.14

For $f < f_r$, the capacitative reactance overshadows the inductive reactance, with the result that the applied voltage lags the current. For $f > f_r$, the inductive reactance takes the leading role, and the voltage across the circuit leads the current. For $f = f_r$, the effects of the capacitative and inductive reactances exactly cancel one another, with the result that the current in the circuit is in phase with the applied voltage. The manner in which the reactance changes with f is shown in Figure 13.14.

13.10 Potential Differences at Resonance

Students are frequently surprised and occasionally pained (in a literal sense) to discover that the potential differences across reactive elements can, as the circuit approaches resonance, exceed greatly the applied voltage. Consider a 60-cycle, 110-volt AC applied to a series circuit consisting of $R = 55$ ohms, $L = 1$ henry, and $C = 7$ microfarads. Hence,

$$X_C = -1/\omega C = -1/377 \times 7 \times 10^{-6} = -379 \text{ ohms}$$

and

$$X_L = \omega L = 377 \text{ ohms}$$

$$I = 110/\sqrt{55^2 + 2^2} \doteq 110/55 = 2 \text{ amperes}$$

Consequently, $V_L \doteq 2(377) = 754$ volts, and $V_C \doteq 758$ volts—each approximately seven times the applied emf.

It would seem from this that a voltmeter across the LC combination would read only 4 volts. In practice it is not possible to obtain a pure inductance

because of the resistance, R_L, of the wire comprising it. Thus, in addition to the 754 volts mentioned above, there exists an IR_L drop (in phase with the current) across the inductance which has its maximum value at resonance. This leads to a V_{LC} value which is substantially higher than 4 volts.

13.11 Sharpness of Resonance: Quality of a Circuit

At resonance the current through a series circuit is at its maximum value, I_r, limited only by the resistance that is present. The manner in which the current varies in the vicinity of resonance is shown in Figure 13.15, where it is plotted as a function of frequency.

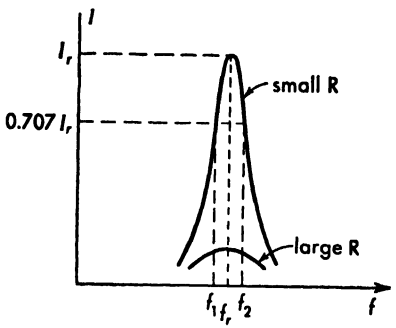


Fig. 13.15

As the resistance in the circuit is reduced, the resonance curve becomes sharper, a property which is often desirable, particularly in radio circuits. The sharpness of resonance of a circuit is described by its *quality* or Q, defined as follows:

$$Q = \frac{f_r}{f_2 - f_1} = \frac{\omega_r}{\omega_2 - \omega_1} \qquad [13\text{-}87]$$

where f_1 and f_2 are the frequencies at which the power dissipation in the circuit has dropped to one-half its resonance value, $I_r R^2$. The power dissipation at f_1 is, therefore,

$$I_1^2 R = \tfrac{1}{2} I_r^2 R \qquad [13\text{-}88]$$

corresponding to

$$I_1 = \frac{I_r}{\sqrt{2}} = 0.707 I_r \qquad [13\text{-}89]$$

as indicated in Figure 13.15. From this it follows that the impedance at f_1 (or f_2) must be $\sqrt{2}$ times its value at resonance ($Z_r = R$). This is so when the resistance and reactance are equal in magnitude, viz.,

$$Z_1 = Z_2 = \sqrt{R^2 + X^2} = \sqrt{R^2 + R^2} = \sqrt{2}R \qquad [13\text{-}90]$$

We may write

$$X_2 = \left(\omega_2 L - \frac{1}{\omega_2 C}\right) = +R \qquad [13\text{-}91]$$

and

$$X_1 = \left(\omega_1 L - \frac{1}{\omega_1 C}\right) = -R \qquad [13\text{-}92]$$

Multiplying [13-91] by ω_2 and [13-92] by ω_1 and subtracting, one obtains

$$L(\omega_2^2 - \omega_1^2) = R(\omega_2 + \omega_1) \qquad [13\text{-}93]$$

or

$$\omega_2 - \omega_1 = \frac{R}{L} \qquad [13\text{-}94]$$

Thus, equation [13-87] becomes

$$Q = \frac{\omega_r}{\omega_2 - \omega_1} = \frac{\omega_r L}{R} \qquad [13\text{-}95]$$

High Q circuits are those that resonate at high frequencies and that have low resistance. In addition to speaking of the quality of a circuit, it is the practice to refer to the Q of a coil alone as $\omega_r L/R$. These two usages of the term are compatible since in a resonant circuit the series resistance associated with the capacitance is negligible. Thus, if a capacitance be chosen to resonate with the coil at the frequency f_r, the Q of the combination will in fact be $\omega_r L/R$.

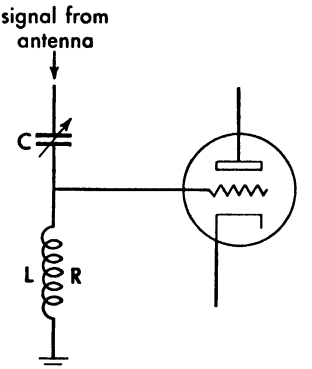


Fig. 13.16

Figure 13.16 shows schematically the manner in which a series LCR circuit can be used to "tune in" a radio signal. The variable element, usually C, is tuned until the circuit is resonant at the desired frequency. At this point the current from the antenna to ground is a maximum, as is also the signal voltage on the grid. If the circuit possesses a high Q, the only significant current flowing in it is that belonging to the desired signal. This discrimination with respect to nearby frequencies is a *sine qua non* if the circuit is to be used in a radio receiver.

13.12 Parallel Circuits

In the analysis of parallel circuits the same methods which we have so far employed are applicable, although they are frequently cumbersome. The analysis is greatly simplified by the use of complex algebra, as described later in the chapter. In the meantime, however, it is instructive to examine cursorily a parallel circuit, such as that shown in Figure 13.17.

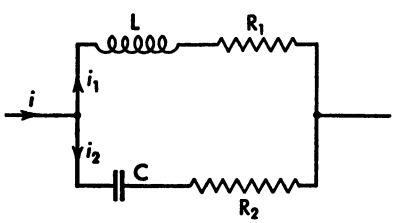


Fig. 13.17

Here the potential difference across each branch has the same value, say, $v = V_0 \sin \omega t$. One may, therefore, compute the current in each branch and add these vectorially to obtain the total current, i. For the upper branch,

$$Z_1 = \sqrt{R_1^2 + \omega^2 L^2} \qquad [13\text{-}96]$$

$$\delta_1 = \tan^{-1} \frac{\omega L}{R_1} \qquad [13\text{-}97]$$

and

$$i_1 = \frac{V_0}{Z_1} \sin (\omega t - \delta_1) \qquad [13\text{-}98]$$

Similarly, for the lower branch,

$$Z_2 = \sqrt{R_2^2 + \frac{1}{\omega^2 C^2}} \qquad [13\text{-}99]$$

$$\delta_2 = \tan^{-1} \frac{-1/\omega C}{R} \qquad [13\text{-}100]$$

and

$$i_2 = \frac{V_0}{Z_2} \sin (\omega t - \delta_2) \qquad [13\text{-}101]$$

Consequently,

$$i = i_1 + i_2 = \frac{V_0}{Z_1} \sin (\omega t - \delta_1) + \frac{V_0}{Z_2} \sin (\omega t - \delta_2) \qquad [13\text{-}102]$$

$$= \frac{V_0}{Z_1} (\sin \omega t \cos \delta_1 - \cos \omega t \sin \delta_1) + \frac{V_0}{Z_2} (\sin \omega t \cos \delta_2 - \cos \omega t \sin \delta_2) \qquad [13\text{-}103]$$

$$= \left[\frac{V_0}{Z_1} \cos \delta_1 + \frac{V_0}{Z_2} \cos \delta_2\right] \sin \omega t - \left[\frac{V_0}{Z_1} \sin \delta_1 + \frac{V_0}{Z_2} \sin \delta_2\right] \cos \omega t \qquad [13\text{-}104]$$

$$= I_0 \cos \delta \sin \omega t - I_0 \sin \delta \cos \omega t \qquad [13\text{-}105]$$

where

$$I_0 \cos \delta = V_0 \left[\frac{\cos \delta_1}{Z_1} + \frac{\cos \delta_2}{Z_2}\right] \qquad [13\text{-}106]$$

$$I_0 \sin \delta = V_0 \left[\frac{\sin \delta_1}{Z_1} + \frac{\sin \delta_2}{Z_2}\right] \qquad [13\text{-}107]$$

and

$$\delta = \tan^{-1} \frac{I_0 \sin \delta}{I_0 \cos \delta} \qquad [13\text{-}108]$$

Thus

$$i = I_0 \sin (\omega t - \delta) \qquad [13\text{-}109]$$

As in the series case, resonance occurs when the total current through the circuit is in phase with the applied voltage. For the phase angle to be zero, one sees from equation [13-107] that

$$\frac{\sin \delta_1}{Z_1} = -\frac{\sin \delta_2}{Z_2} \qquad [13\text{-}110]$$

This is the condition for parallel resonance.

Let us simplify the problem by assuming that the resistance, R_2, in the capacitative branch of Figure 13.17 is zero. We are studying, therefore, the case of a capacitor shunted by a coil and its associated series resistance. Equation [13-110] may now be written

$$\frac{\omega_r L}{(R_1^2 + \omega_r^2 L^2)} = \omega_r C \qquad [13\text{-}111]$$

or

$$\omega_r = \frac{\sqrt{LC - R_1^2 C^2}}{LC} \qquad [13\text{-}112]$$

which reduces, for an ideal inductance ($R_1 = 0$), to the same value as was sound for the series resonance case:

$$\omega_r = \frac{1}{\sqrt{LC}} \qquad [13\text{-}113]$$

Further discussion of parallel circuits will be delayed until the complex-algebra methods have been developed.

13.13 Complex Algebra

The use of *complex algebra* simplifies the treatment of many AC problems. In this section some of the properties of complex numbers will be briefly reviewed. This knowledge will then be employed in subsequent sections to reconsider some aspects of alternating currents which we have already discussed, and to study others not yet touched upon.

Complex Quantities

Consider the quantity

$$Z = R + jX \qquad [13\text{-}114]$$

where R and X are *real* quantities and $j = \sqrt{-1}$. Z is comprised of a real component and an *imaginary* component, and is said to be *complex*.

All complex quantities may be expressed in the standard form shown in equation [13-114]. For example, suppose

$$Z = \frac{a + jb}{c + jd} \qquad [13\text{-}115]$$

This can be put into the standard form by multiplying by $(c - jd)/(c - jd)$. Thus,

$$Z = \frac{a + jb}{c + jd} \times \frac{c - jd}{c - jd} = \frac{ac + bd}{c^2 + d^2} + j\frac{bc - ad}{c^2 + d^2} \qquad [13\text{-}116]$$

which is now in the standard form. This operation is known as *rationalization*, and the quantity $c - jd$ is termed the *complex conjugate* of the original denominator, $c + jd$. A complex quantity and its complex conjugate differ only in the signs of their imaginary components, and their product, as above, is always real.

Vector Character of a Complex Quantity and Its Graphical Representation

Complex quantities are represented graphically by plotting their real components as abscissae and their imaginary components as ordinates. The quantity $Z = R + jX$ is shown so plotted in Figure 13.18.

Evidently, in this type of plot, the complex quantity is represented by a vector, whose length indicates the magnitude of the complex quantity, and whose direction indicates the relative proportions of its real and imaginary constituents. This vector character of a complex quantity will be indicated from now on by writing all complex quantities in bold-face, the same symbolism that has been employed to represent true vectors. This is not to say that complex quantities are manipulated according to the same rules as vectors. Although they are added and subtracted according to the same rules, the similarity ends here. For example, complex quantities are not multiplied as *dot* or *cross* products, but in the ordinary algebraic manner. Since we shall not encounter any true vectors in this chapter, the use of bold-face to designate complex quantities should not cause confusion.

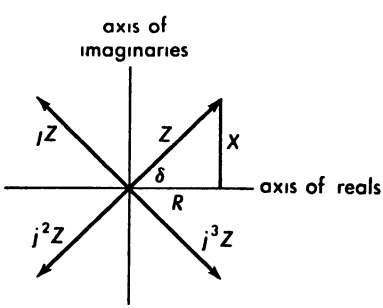


Fig. 13.18

Multiplication of a complex vector by j is a vector operation producing a 90° counterclockwise rotation of the complex vector without changing its

magnitude. Thus,

$$j\boldsymbol{Z} = j(R + jX) = jR - X \qquad [13\text{-}117]$$

has the direction shown in Figure 13.18, where the vectors resulting from two further multiplications by j are also shown.

Complex quantities may often be expressed to advantage in an exponential form. Thus,

$$A = B + jC$$

$$= A\cos\theta + jA\sin\theta \qquad [13\text{-}118]$$

$$= Ae^{j\theta} \qquad [13\text{-}119]^{6}$$

which is a vector that rotates in a counterclockwise direction as θ increases in value.

The *argument*, or *phase*, of the complex quantity, $\boldsymbol{Z} = R + jX$, is $\delta = \tan^{-1}(X/R)$. The *complex conjugate* of $\boldsymbol{Z}$ is $\boldsymbol{Z}^* = R - jX$. The *modulus*, or absolute value, of $\boldsymbol{Z}$ is given by

$$Z = \sqrt{\boldsymbol{Z}\boldsymbol{Z}^*} = \sqrt{R^2 + X^2} \qquad [13\text{-}120]$$

Addition, Subtraction, Multiplication, and Division of Complex Quantities

Addition of complex quantities is performed according to the usual rules for vectors. Thus, in adding two complex quantities graphically, the second complex quantity is shown with its tail at the tip of the first one. The complex quantity representing the sum is then drawn from the tail of the first vector to the tip of the second. This operation is equivalent to the following

$$\boldsymbol{Z} = \boldsymbol{Z}_1 + \boldsymbol{Z}_2 \qquad [13\text{-}121]$$

$$= (R_1 + jX_1) + (R_2 + jX_2) \qquad [13\text{-}122]$$

$$= (R_1 + R_2) + j(X_1 + X_2) \qquad [13\text{-}123]$$

Thus, the real (imaginary) part of the sum of two complex quantities is the sum of their real (imaginary) parts. This corresponds to the usual vector addition in which the X-, Y-, and Z-components of the two vectors are separately added. Subtraction of complex numbers is similarly handled by vector methods.

---

[6] The identity $e^{j\theta} = \cos\theta + j\sin\theta$ can be proved by expanding $e^{j\theta}$ in an infinite series, thus,

$$e^{j\theta} = 1 + j\theta + (j\theta)^2/2! + (j\theta)^3/3! + \cdots$$
$$= (1 - \theta^2/2! + \theta^4/4! - \cdots) + j(\theta - \theta^3/3! + \theta^5/5! - \cdots)$$
$$= \cos\theta + j\sin\theta.$$

Multiplication of complex quantities is done according to the usual scalar rules. Thus,

$$\boldsymbol{Z}_1\boldsymbol{Z}_2 = (R_1 + jX_1)(R_2 + jX_2) = (R_1R_2 - X_1X_2) + j(R_2X_1 + R_1X_2) \qquad [13\text{-}124]$$

Division of complex numbers is the rationalization operation described above. Thus,

$$\frac{\boldsymbol{Z}_1}{\boldsymbol{Z}_2} = \frac{R_1 + jX_1}{R_2 + jX_2} = \frac{R_1 + jX_1}{R_2 + jX_2} \times \frac{R_2 - jX_2}{R_2 - jX_2} = \frac{(R_1R_2 + X_1X_2) + j(R_2X_1 - R_1X_2)}{R_2^2 + X_2^2} \qquad [13\text{-}125]$$

The exponential form for expressing complex quantities (equation [13-119]) is particularly well suited for performing multiplications and division. Thus,

$$\boldsymbol{A}_1\boldsymbol{A}_2 = (A_1e^{j\theta_1})(A_2e^{j\theta_2}) = A_1A_2e^{j(\theta_1+\theta_2)} \qquad [13\text{-}126]$$

and

$$\frac{\boldsymbol{A}_1}{\boldsymbol{A}_2} = \frac{A_1}{A_2}e^{j(\theta_1-\theta_2)} \qquad [13\text{-}127]$$

13.14 Application of Complex Algebra to Series Circuits

It will be recalled that in the impedance diagrams (Figs. 13.4, 13.8, and 13.11) for series LR, RC, and LCR circuits, resistance was plotted along the X-axis, inductive reactance along the positive Y-axis, and capacitative reactance along the negative Y-axis. In the language of complex algebra (cf. Fig. 13.18), resistances are real quantities, whereas inductive and capacitative reactances are, respectively, positive and negative imaginary quantities. It has not been a coincidence, therefore, that in the foregoing discussion the typical complex quantity has usually been written as $\boldsymbol{Z} = R + jX$. This is, in fact, the complex algebra expression for the total impedance of a series combination of R and X.

Illustrative Example: Compute the current flowing in the circuit of Figure 13.19. First, we evaluate the separate reactances:

$$X_{L_1} = \omega L_1 = 377(0.2) = 75.4 \text{ ohms}$$

$$X_{L_2} = 150.8 \text{ ohms}$$

$$X_C = -1/\omega C = -\frac{10^6}{377 \times 10} = -265.2 \text{ ohms}$$

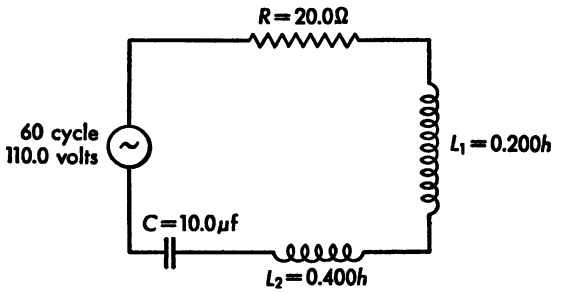


Fig. 13.19

Then we work out the complex expression for the impedance:

$$\begin{aligned} \boldsymbol{Z} &= R + jX \\ &= 20 + j(75.4 + 150.8 - 265.2) \\ &= 20 - j39 \text{ ohms} \end{aligned}$$

We can now calculate the current from the expression

$$\boldsymbol{I} = \frac{\boldsymbol{V}}{\boldsymbol{Z}}$$

where $\boldsymbol{V}$ is the total potential drop across the RL_1L_2C combination. This gives

$$\boldsymbol{I} = \frac{110}{20 - j39} = \frac{110}{20 - j39} \times \frac{20 + j39}{20 + j39} = 1.13 + j2.23 \text{ amperes}$$

The effective value of the current is the modulus of its complex effective value, viz.,

$$I = \sqrt{\boldsymbol{I}\boldsymbol{I}^*} = \sqrt{1.13^2 + 2.23^2} = 2.50 \text{ amperes}$$

We should make a few other remarks to demonstrate the further usefulness of the above results. Since the complex expressions for the voltage drop and the current are given by

$$\boldsymbol{V} = 110 \text{ volts}$$

and

$$\boldsymbol{I} = 1.13 + j2.23 \text{ amperes}$$

the current leads the voltage by the angle between these two vectors, namely, $\tan^{-1}(2.23/1.13) = 63°$. This piece of intelligence could have been gained by inspecting the complex expression for the impedance, which indicates that the voltage lags the current by the phase angle $\delta = \tan^{-1}(-39/20) = -63°$.

The complex expression for the power is given by the product of I and V, in this case

$$\begin{aligned} \boldsymbol{P} &= \boldsymbol{V}\boldsymbol{I} = 110(1.13 + j2.23) \\ &= 124 + j245 \text{ watts} \end{aligned}$$

From this the real power is seen to be 124 watts, while the wattless power is 245 watts. This value for the real power, of course, agrees with the product

$$IV \cos \delta = 2.50(110) \cos 63°$$

13.15 Application of Complex Algebra to Parallel Circuits

The ease with which the current through a parallel circuit can be calculated will be demonstrated by means of an example.

Illustrative Example: Calculate the instantaneous total current in the circuit of Figure 13.20. For the upper branch, $\boldsymbol{Z}_1 = 3 + j4$. For the lower branch, $\boldsymbol{Z}_2 = 8 - j6$. The joint impedance is computed using the rule for combining parallel resistances (see § 13.18). Thus,

$$\frac{1}{\boldsymbol{Z}} = \frac{1}{\boldsymbol{Z}_1} + \frac{1}{\boldsymbol{Z}_2} = \frac{1}{3 + j4} + \frac{1}{8 - j6}$$

or

$$\frac{1}{\boldsymbol{Z}} = \frac{11 - j2}{48 + j14}$$

Hence,

$$\boldsymbol{I} = \frac{\boldsymbol{V}}{\boldsymbol{Z}} = 100 \times \frac{11 - j2}{48 + j14} = 100 \times \frac{11 - j2}{48 + j14} \times \frac{48 - j14}{48 - j14} = 20 - j10 \text{ amperes}$$

Comparing the complex expressions for $\boldsymbol{V}$ and $\boldsymbol{I}$, namely $\boldsymbol{V} = 100$ and $\boldsymbol{I} = 20 - j10$, we see that the voltage leads the current by the phase angle $\theta = \tan^{-1}(10/20) = 26.6° \doteq 0.46$ radians. Also, from the complex expression for the current, the effective and maximum values of the current are $I = \sqrt{20^2 + 10^2} = 22.4$ and $I_0 = 22.4\sqrt{2}$.

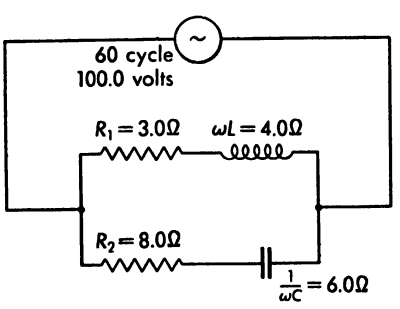


Fig. 13.20

If we assume the instantaneous voltage to be

$$v = V_0 \sin \omega t = 100\sqrt{2} \sin 377t$$

then the instantaneous current is

$$i \doteq 22.4\sqrt{2} \sin (377t - 0.46)$$

Incidentally, we could easily have obtained the total current by calculating the current through each of the two branches and adding them vectorially. Thus, in the upper branch,

$$\boldsymbol{I}_1 = 100/\boldsymbol{Z}_1 = 100/(3 + j4) = 12 - j16 \text{ amperes}$$

and in the lower branch,

$$\boldsymbol{I}_2 = 100/\boldsymbol{Z}_2 = 100/(8 - j6) = 8 + j6 \text{ amperes}$$

Adding these, we obtain

$$\boldsymbol{I} = \boldsymbol{I}_1 + \boldsymbol{I}_2 = 20 - j10 \text{ amperes}$$

as above.

13.16 Parallel Resonance

We have already discussed, in § 13.12, the condition for resonance in a parallel LC circuit, such as that shown in Figure 13.21. In the figure, R is the resistance of the wire of which the coil constituting the inductance is wound. This circuit will now be analyzed using complex algebra.

The joint impedance of the two branches is given by

$$\frac{1}{\boldsymbol{Z}} = \frac{1}{R + j\omega L} + \frac{1}{-j/\omega C} = \frac{1}{R + j\omega L} - \frac{\omega C}{j} \qquad [13\text{-}128]$$

$$= \frac{\omega CR + j(\omega^2 LC - 1)}{\omega L - jR} \qquad [13\text{-}129]$$

Hence, if the potential difference between points A and B is $\boldsymbol{V}$, the total current through the parallel network is

$$\boldsymbol{I} = \frac{\boldsymbol{V}}{\boldsymbol{Z}} = V\left[\frac{\omega CR + j(\omega^2 LC - 1)}{\omega L - jR} \times \frac{\omega L + jR}{\omega L + jR}\right]$$

$$= V\left[\frac{R + j(\omega^3 L^2 C - \omega L + \omega CR^2)}{\omega^2 L^2 + R^2}\right] \qquad [13\text{-}130]$$

At resonance, $\boldsymbol{I}$ and $\boldsymbol{V}$ are in phase. Since we have assumed that $\boldsymbol{V} = V$ (i.e., that it has no imaginary component), the imaginary component of I in equation [13-130] must vanish at resonance. The condition for resonance is, therefore,

$$\omega_r^2 L^2 C - L + CR^2 = 0 \qquad [13\text{-}131]$$

or

$$\omega_r = \frac{\sqrt{LC - C^2R^2}}{LC} \qquad [13\text{-}112]$$

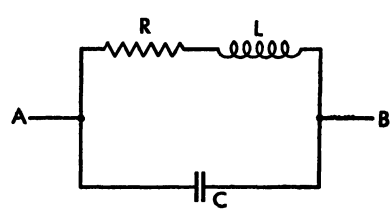


Fig. 13.21

as obtained previously.

The impedance at resonance has the value,

$$Z_r = \frac{\omega_r^2 L^2 + R^2}{R} = \frac{L}{RC} \qquad [13\text{-}132]$$

When this expression is compared with the general expression for impedance,

$Z = R + jX$, it is seen that the impedance at resonance is entirely resistive. Furthermore, the effective resistance at resonance approaches infinity as the resistance of the coil approaches zero. In practice, of course, the resistance of the coil never vanishes, but it is still possible for the effective resistance of an actual parallel resonant circuit to be very high—very much larger than the actual resistance in the circuit.

The behavior of a parallel resonant circuit is strikingly different from that of a series resonant circuit. In both cases the impedance is resistive but, whereas parallel resonance implies maximum impedance, series resonance implies minimum impedance. Parallel resonance is sometimes termed *antiresonance*, in order to distinguish it from series resonance.

As with series resonant circuits, parallel resonant circuits are employed at the input of radio and television receivers, as well as between amplifier stages, in order to discriminate against adjacent but undesired frequencies. A typical input circuit was shown in Figure 3.6. The maximum voltage signal appearing on the grid of the first tube of the receiver corresponds to the frequency for which the parallel circuit is resonant and to which it, therefore, offers the maximum impedance. The sharpness of resonance of a parallel circuit is described by its *quality*, or Q, as defined in § 13.11.

13.17 Admittance, Conductance, and Susceptance

Conductance was defined in § 6.4 as the reciprocal of resistance, and the comment was made at that time that the concept of conductance is especially useful in dealing with parallel DC circuits. For a similar reason the reciprocal of the impedance, defined as the *admittance*, is useful in handling parallel AC circuits.

Admittance is designated by the symbol $\mathbf{Y}$. It consists of two components, the *conductance*, G, and the *susceptance*, B. The relationship between $\mathbf{Y}$, G, and B is expressed in complex form as follows:

$$\mathbf{Y} = G - jB \qquad [13\text{-}133]$$

The relationship between G and B on the one hand and R and X on the other is easily derived. Thus,

$$\mathbf{Y} = \frac{1}{\mathbf{Z}} = \frac{1}{R + jX} = \frac{R - jX}{R^2 + X^2} \qquad [13\text{-}134]$$

By comparing equations [13-133] and [13-134], we see that

$$G = \frac{R}{R^2 + X^2} \quad \text{and} \quad B = \frac{X}{R^2 + X^2} \qquad [13\text{-}135]$$

Evidently, for purely resistive elements, the conductance is the reciprocal of the resistance (in keeping with our earlier definition), while, for purely reactive elements, the susceptance is the reciprocal of the reactance. Also, since

$$\boldsymbol{Y} = \frac{1}{\boldsymbol{Z}} = \sum_1^n \frac{1}{\boldsymbol{Z}_i} = \sum_1^n \boldsymbol{Y}_i \qquad [13\text{-}136]$$

the total admittance of a group of impedances in parallel is the sum of their separate admittances.

Useful analogies between series and parallel circuits are established by describing the former in terms of impedance and its components and the latter in terms of admittance and its components. For example, the total impedance (admittance) of a series (parallel) circuit is the sum of the individual complex impedances (admittances). Also, the impedance (admittance) of a series (parallel) resonant circuit is a minimum.

Problem 13.24 provides an opportunity, for those who wish to do so, to develop some fluency in the use of the admittance language.

13.18 Kirchhoff's Laws in Complex Form

Kirchhoff's two laws (§ 5.11) for a DC circuit are: (1) the algebraic sum of the currents meeting at any point is zero and (2) the sum of the emf's in any closed circuit is equal to the sum of the potential drops. The first law is sometimes called the *current law*, and the second the *voltage law*. Since these two laws hold at any instant, they apply equally well to AC circuits; indeed, we have already made frequent use of the voltage law to solve problems involving transient or fluctuating currents.

Using the language of complex algebra, Kirchhoff's two laws can be restated in a form that is especially suited for handling AC problems.

1. The algebraic sum of the complex currents meeting at any point in a circuit is zero. Thus,

$$\sum_1^n \boldsymbol{I}_i = 0 \qquad [13\text{-}137]^7$$

[7] If we deal with a group of parallel complex impedances in the same manner as was done with parallel resistances in § 5.8, the complex current through each component has the form $\boldsymbol{I}_i = \boldsymbol{V}/\boldsymbol{Z}_i$. Combining these constituent currents, as on the left side of equation [13-137], we see that the total current through the array is given by

$$\boldsymbol{I} = \sum_1^n \boldsymbol{I}_i = \boldsymbol{V} \sum_1^n 1/\boldsymbol{Z}_i = \boldsymbol{V}/\boldsymbol{Z}.$$

That is, complex impedances in parallel are added according to the reciprocal rule, as was done in § 13.15.

2. The algebraic sum of the complex applied emf's in any closed circuit is equal to the algebraic sum of the IZ drops. Thus,

$$\sum_1^n \mathcal{E}_i = \sum_1^m I_j Z_j \qquad [13\text{-}138]$$

This statement of the voltage law differs from preceding statements of the law in that only *applied* emf's are shown on the left-hand side of equation [13-138]. Any back emf's that occur in the circuit are not entered on the left-hand side, but are included instead, with reversed sign, on the right-hand side of the equation, among the voltage drops. This method of dealing with back emf's is, of course, equivalent to what has been done before. It has the advantage, which we shall reap, that the effect of inductive reactance is taken into account, in equation [13-138], in the same way as capacitative reactance and resistance.

13.19 AC Bridges

Let us now apply Kirchhoff's laws in complex form to a bridge circuit, shown in Figure 13.22, in which the applied emf, $\mathcal{E}$, is harmonic and the bridge elements are Z_1, Z_2, Z_3, and Z_4. Z represents the impedance of whatever detecting device (say, a set of earphones) is used to ascertain when points C and D are at the same potential.

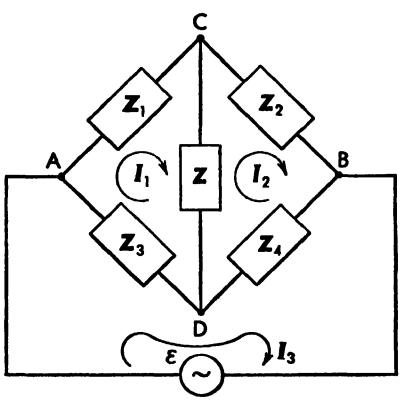


Fig. 13.22

Using the method of cyclic currents (§ 5.12) we can write Kirchhoff's second law for the ACD and CBD circuits, respectively, as

$$0 = I_1 Z_1 + (I_1 - I_2)Z + (I_1 - I_3)Z_3 \qquad [13\text{-}139]$$

and

$$0 = I_2 Z_2 + (I_2 - I_3)Z_4 + (I_2 - I_1)Z \qquad [13\text{-}140]$$

When the bridge is balanced, the current through Z is zero. That is,

$$I_1 - I_2 = 0 \qquad [13\text{-}141]$$

Equations [13-139] and [13-140] then simplify to

$$I_1 Z_1 = -(I_1 - I_3)Z_3 \qquad [13\text{-}142]$$

and

$$I_1 Z_2 = -(I_1 - I_3)Z_4 \qquad [13\text{-}143]$$

which yield, when the former is divided by the latter,

$$\frac{Z_1}{Z_2} = \frac{Z_3}{Z_4} \qquad [13\text{-}144]$$

This condition of balance has the same form as for a DC bridge (equation [6-23]) except that impedances replace the resistances.

The Inductance Bridge

The inductance bridge, shown in Figure 13.23, is used to compare the value of an unknown inductance with that of a standard.

The condition of balance (equation [13-144]) takes, in this case, the form:

$$\frac{R_1 + j\omega L_1}{R_2} = \frac{R_3 + j\omega L_3}{R_4} \qquad [13\text{-}145]$$

Equating the real parts of this equation we obtain

$$\frac{R_1}{R_2} = \frac{R_3}{R_4} \qquad [13\text{-}146]$$

while equating imaginary parts we obtain

$$\frac{L_1}{R_2} = \frac{L_3}{R_4} \qquad [13\text{-}147]$$

Hence, balance of the bridge involves two separate balances: the usual *DC balance* described by equation [13-146] and, in addition, an *AC balance* described by equation [13-147].

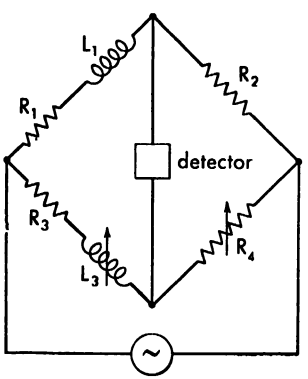


Fig. 13.23

It usually simplifies the achieving of the AC balance to obtain the DC balance first. This can be done by using a battery as the source of emf and a galvanometer as the detector. These are replaced by the harmonic emf and the AC detector in order to secure the AC balance. As can be seen from equation [13-146], at least one of the resistances must be variable in order to obtain the DC balance. Then, if one of the inductances, say, L_3, is a variable standard inductance, it can be adjusted to obtain the AC balance. The value of the unknown inductance, L_1, is finally got from equation [13-147].

The double balance can be obtained simultaneously by using an AC emf throughout. This, however, calls for a patient, and often discouraging, series of alternating adjustments of L_3 and R_4. The achievement of the double balance is indicated by a null in the sound heard in the earphones.

A *capacitance bridge* can be constructed by replacing L_1 and L_3 in Figure 13.23 with two capacitors, C_1 and C_3. It is left as an exercise for the reader to show that at balance

$$\frac{R_1}{R_2} = \frac{R_3}{R_4} \quad \text{and} \quad \frac{C_1}{C_3} = \frac{R_4}{R_2} \qquad [13\text{-}148]$$

In this bridge it is obviously not possible to secure a DC balance first. Instead, the double balance must be achieved with an AC emf. If, as is often the case, a variable standard capacitor is not available for use, the double balance can be achieved by adjusting the resistors alone.

Many different AC bridges have been devised. For the details of these, and the particular advantages which each offers, the reader is referred to a text, such as M. B. Stout, *Basic Electrical Measurements* (Prentice-Hall, 1950), which deals with electrical measurements.

13.20 Condition for Maximum Power Transfer in AC Circuits

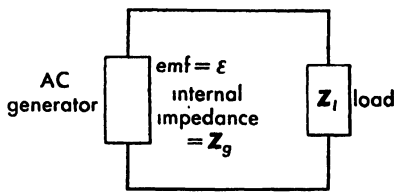


Fig. 13.24

We have shown in § 5.9 that the condition for maximum power transfer in a DC circuit is that the resistance of the load and that of the battery or generator be equal. An analogous condition holds for AC circuits.

Let us assume, in the circuit of Figure 13.24, that $\boldsymbol{Z}_g$ and $\boldsymbol{Z}_l$ are the impedances of the generator and load, respectively. The current in the circuit is

$$\boldsymbol{I} = \frac{\mathcal{E}}{\boldsymbol{Z}_g + \boldsymbol{Z}_l} = \frac{\mathcal{E}}{(R_g + jX_g) + (R_l + jX_l)} = \frac{\mathcal{E}}{(R_g + R_l) + j(X_g + X_l)} \qquad [13\text{-}149]$$

and the power dissipated in the load is given by

$$I^2R_l = \boldsymbol{I}\boldsymbol{I}^*R_l = \frac{\mathcal{E}^2R_l}{(R_g + R_l)^2 + (X_g + X_l)^2} \qquad [13\text{-}150]$$

As far as reactance is concerned, this expression has its maximum value when

$$X_g = -X_l \qquad [13\text{-}151]$$

in which case equation [13-150] reduces to

$$I^2R_l = \frac{\mathcal{E}^2R_l}{(R_g + R_l)^2} \qquad [13\text{-}152]$$

which has been shown in § 5.9 to have its maximum value when $R_g = R_l$.

Hence, *the condition for maximum power transfer from an AC generator to a load is that the impedance of the load be the complex conjugate of the generator impedance:*

$$R_l + jX_l = R_g - jX_g \qquad [13\text{-}153]$$

or

$$\mathbf{Z}_l = \mathbf{Z}_g^* \qquad [13\text{-}154]$$

13.21 The Transformer

The widespread use of AC power, in preference to DC power, is due to the fact that it can be transmitted at high voltages and low currents, with consequent small power loss (I^2R), and can subsequently be transformed to modest voltages and larger currents near the locations where it is to be utilized. This AC voltage transformation is accomplished by *transformers*. Furthermore, the wide variety of voltages which are often required at a single location can be obtained, by means of transformers, from a single source of AC power. Transformers are also employed, as we shall see, as impedance-matching devices in order to secure maximum power transfer from generator to load. In addition, they can provide convenient AC coupling between two circuits which are not at the same DC potential.

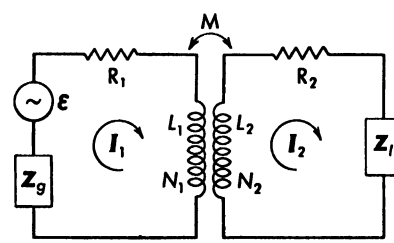


Fig. 13.25

Let us consider the transformer shown in the circuit of Figure 13.25. The primary coil, of N_1 turns, possesses inductance L_1 and resistance R_1. N_2, L_2, and R_2 represent similar properties of the secondary coil. The mutual inductance of the two coils is M. $\mathbf{Z}_g$ and $\mathbf{Z}_l$ are the impedances of the generator and load, respectively.

Applying Kirchhoff's second law to the two meshes in Figure 13.25, we

obtain

$$\mathcal{E} = \boldsymbol{I}_1(\boldsymbol{Z}_g + R_1 + j\omega L_1) \pm j\omega M\boldsymbol{I}_2 = \boldsymbol{I}_1\boldsymbol{Z}_{11} \pm \boldsymbol{I}_2\boldsymbol{Z}_{12} \quad [13\text{-}155]$$

and

$$0 = \boldsymbol{I}_2(\boldsymbol{Z}_l + R_2 + j\omega L_2) \pm j\omega M\boldsymbol{I}_1 = \boldsymbol{I}_2\boldsymbol{Z}_{22} \pm \boldsymbol{I}_1\boldsymbol{Z}_{21} \quad [13\text{-}156]$$

where

$$\boldsymbol{Z}_{11} = \boldsymbol{Z}_g + R_1 + j\omega L_1 = \boldsymbol{Z}_g + \boldsymbol{Z}_1 \quad [13\text{-}157]$$

$$\boldsymbol{Z}_{12} = \boldsymbol{Z}_{21} = j\omega M \quad [13\text{-}158]$$

and

$$\boldsymbol{Z}_{22} = \boldsymbol{Z}_l + R_2 + j\omega L_2 = \boldsymbol{Z}_l + \boldsymbol{Z}_2 \quad [13\text{-}159]$$

The choice of signs preceding the final terms in equations [13-155] and [13-156] depends upon whether the two coils are wound in the same or opposite senses. For simplicity, we shall assume the positive sign hereafter.

Solving equations [13-155] and [13-156] simultaneously, we find

$$\boldsymbol{I}_1 = \frac{\mathcal{E}\boldsymbol{Z}_{22}}{\boldsymbol{Z}_{11}\boldsymbol{Z}_{22} - \boldsymbol{Z}_{12}^2} \quad [13\text{-}160]$$

and

$$\boldsymbol{I}_2 = \frac{-\mathcal{E}\boldsymbol{Z}_{12}}{\boldsymbol{Z}_{11}\boldsymbol{Z}_{22} - \boldsymbol{Z}_{12}^2} \quad [13\text{-}161]$$

From these expressions we can compute the voltage drop, $\boldsymbol{V}_1$, across the primary coil, and the voltage drop across the load in the secondary circuit. This latter is the voltage, $\boldsymbol{V}_2$, appearing across the secondary coil. Thus,

$$\boldsymbol{V}_1 = \mathcal{E} - \boldsymbol{I}_1\boldsymbol{Z}_g \quad [13\text{-}162]$$

and

$$\boldsymbol{V}_2 = \boldsymbol{I}_2\boldsymbol{Z}_l \quad [13\text{-}163]$$

Voltage and Current Transformation in an Ideal Transformer

In the ideal transformer the two coils are so wound that the magnetic coupling is complete. Hence, the coefficient of coupling, K, of the coils is unity, and their mutual inductance is given by

$$M = \sqrt{L_1 L_2} \quad [9\text{-}61]$$

The resistances of the two coils of the ideal transformer are zero. In addition, the reactances of the coils are so large in comparison with the generator and load impedances that we may write

$$\boldsymbol{Z}_{11} = \boldsymbol{Z}_g + R_1 + j\omega L_1 \doteq \boldsymbol{Z}_g + j\omega L_1 \doteq j\omega L_1 \quad [13\text{-}164]$$

and

$$\boldsymbol{Z}_{22} = \boldsymbol{Z}_l + R_2 + j\omega L_2 \doteq \boldsymbol{Z}_l + j\omega L_2 \doteq j\omega L_2 \quad [13\text{-}165]$$

When these simplifying assumptions are made, the ratio $\boldsymbol{I}_1/\boldsymbol{I}_2$ becomes

$$\frac{\boldsymbol{I}_1}{\boldsymbol{I}_2} = -\frac{\boldsymbol{Z}_{22}}{\boldsymbol{Z}_{12}} \doteq -\frac{j\omega L_2}{j\omega M} = -\frac{L_2}{M} = -\sqrt{\frac{L_2}{L_1}} \qquad [13\text{-}166]$$

or

$$\frac{\boldsymbol{I}_1}{\boldsymbol{I}_2} \doteq -\frac{N_2}{N_1} \qquad [13\text{-}167]$$

Equation [13-167] follows from the fact (§ 9.10) that the inductance of an ideally coupled coil is proportional to the square of the number of turns.

At the same time the ratio $\boldsymbol{V}_1/\boldsymbol{V}_2$ becomes

$$\frac{\boldsymbol{V}_1}{\boldsymbol{V}_2} = \frac{\mathcal{E} - \boldsymbol{I}_1\boldsymbol{Z}_g}{\boldsymbol{I}_2\boldsymbol{Z}_l} \doteq \frac{\mathcal{E}}{\boldsymbol{I}_2\boldsymbol{Z}_l} \qquad [13\text{-}168]$$

$$\doteq -\frac{\mathcal{E}[(j\omega L_1)(j\omega L_2 + \boldsymbol{Z}_l) - (j\omega M)^2]}{\mathcal{E}(j\omega M)\boldsymbol{Z}_l} \qquad [13\text{-}169][8]$$

$$\doteq -\frac{[-\omega^2 L_1 L_2 + j\omega L_1\boldsymbol{Z}_l + \omega^2 M^2]}{j\omega M\boldsymbol{Z}_l} \qquad [13\text{-}170]$$

$$\doteq -\frac{L_1}{M} \doteq -\sqrt{\frac{L_1}{L_2}} \qquad [13\text{-}171]$$

or

$$\frac{\boldsymbol{V}_1}{\boldsymbol{V}_2} \doteq -\frac{N_1}{N_2} \qquad [13\text{-}172]$$

This equation indicates the voltage transformation property of a transformer. When $N_2 > N_1$, the voltage across the secondary coil is greater than the voltage across the primary. Such a transformer is said to be a *step-up* transformer. When $N_2 < N_1$, the arrangement constitutes a *step-down* transformer. We also note, from equation [13-167], that a transformer which steps up (steps down) the voltage steps down (steps up) the current by the same factor. Finally, we note from [13-167] and [13-172] that

$$\boldsymbol{V}_1\boldsymbol{I}_1 = \boldsymbol{V}_2\boldsymbol{I}_2 \qquad [13\text{-}173]$$

in keeping with the fact that no power is dissipated in an ideal transformer—the power output is equal to the power input.

The negative signs appearing in equations [13-167] and [13-172] indicate that the secondary voltage and current are out of phase with their primary counterparts when the two coils are wound in the same sense.

[8] Here $\boldsymbol{Z}_l + j\omega L_2$, rather than $j\omega L_2$, is substituted for $\boldsymbol{Z}_{22}$. Otherwise, the numerator would vanish.

Actual Transformers

As it happens, the ideal transformer can be closely approximated in practice, and the results derived above provide a reasonable description of actual transformers. Close coupling between primary and secondary coils is secured by winding both coils on an iron yoke, as in Figure 9.13. The flux passing through the primary coil is led by the iron to pass also through the secondary coil. Further, the amount of the flux is greatly enhanced by the magnetic polarization of the iron, as in any magnetic circuit. In addition, the presence of the iron increases the values of L_1 and L_2, with the effect that impedances elsewhere in the circuit are negligible in comparison to them. In short, the conditions described by equations [9-61], [13-164], and [13-165] are largely achieved.

The iron forming the core of a transformer is carried once through the hysteresis cycle during each cycle of the applied voltage. It will be recalled from § 10.19 that the energy required to carry a sample of iron through its magnetic cycle is proportional to the area of its hysteresis loop. It is therefore imperative that transformer cores be made of iron with narrow hysteresis loops. This loss, as well as that due to eddy currents in the core, has been neglected in the preceding discussion of the transformer.

Autotransformer

A widely used type of transformer is the autotransformer shown in Figure 13.26. Here the primary and secondary coils are overlapping portions of the same coil. The number of turns on the secondary winding, and therewith the output voltage, is changed by moving the variable contact, C. The range of V_2 is from zero (C at position A) to a value somewhat higher than V_1 (C at position B). Commercially available autotransformers are usually known by their trade names, such as "Variac" or "Powerstat." The autotransformer suffers from the limitation that the primary and secondary coils are not DC-isolated. In many applications this is no disadvantage whatsoever.

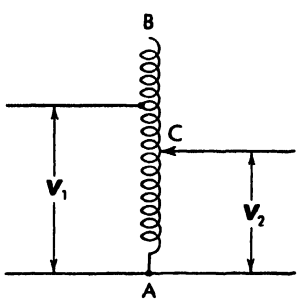


Fig. 13.26

Impedance Matching by Means of a Transformer

Let us now examine, in some detail, the expression (equation [13-160]) for the current in the primary circuit of Figure 13.25:

$$I_1 = \frac{\mathcal{E}Z_{22}}{Z_{11}Z_{22} - Z_{12}^2} = \frac{\mathcal{E}}{Z_{11} - \dfrac{Z_{12}^2}{Z_{22}}} \qquad [13\text{-}174]$$

Here the denominator of the final expression represents the total effective impedance, Z'_{eff}, that is encountered by the primary circuit, I_1:

$$Z'_{\text{eff}} = Z_{11} - \frac{Z_{12}^2}{Z_{22}} \qquad [13\text{-}175]$$

Making use of the characteristics of an ideal transformer expressed in equations [9-61], [13-164], and [13-165], we can rewrite equation [13-175] as follows:

$$Z'_{\text{eff}} \doteq Z_g + j\omega L_1 - \frac{(j\omega M)^2}{Z_l + j\omega L_2} \qquad [13\text{-}176]$$

which simplifies to

$$Z'_{\text{eff}} \doteq Z_g + \frac{L_1}{L_2} Z_l \doteq Z_g + \left(\frac{N_1}{N_2}\right)^2 Z_l \qquad [13\text{-}177]$$

This expression suggests the impedance-matching potentialities of the transformer. The impedance of the primary circuit consists of the generator impedance plus $(N_1/N_2)^2$ times the load impedance. In brief, the load, which is located in the secondary circuit, is reflected into the primary circuit where its effective value, $(N_1/N_2)^2 Z_l$ is determined by the turns ratio of the transformer. In this way it can be arranged for the generator to "see" the load as an impedance which is considerably different from its actual value. For example, if both generator and load are pure resistances, the load can be matched to the generator by means of a transformer, with the result that more power is delivered to it than would be the case if the connection were made directly. Problems 13.25, 13.26, and 13.27 provide examples of impedance transformation and impedance matching. The question of impedance matching is of vital importance in all communications engineering.

13.22 AC Voltages and Currents as Rotating Vectors

It is likely that the reader was first introduced to simple harmonic motion by means of the "circle of reference." Here the instantaneous displacement of a one-dimensional harmonic oscillator, whose motion takes place along the X-axis, is given by the abscissa of a point that moves with constant speed in the X-Y plane around a circle whose center is at the origin of coordinates. The radius of the circle is equal to the maximum displacement or amplitude, A, of

the simple harmonic motion. This is shown in Figure 13.27 where the X-displacement at any time is

$$x = A \cos \theta = A \cos (\omega t + \delta) \qquad [13\text{-}178$$

In the language of complex algebra, A is a rotating vector of constant length whose real component (projection along the X-axis) gives the instantaneous value of x. We have seen in equation [13-119] that the exponential form of a complex quantity represents such a rotating vector. An equivalent expression for an AC voltage, $\varepsilon = \mathcal{E}_0 \cos \omega t$, is, therefore, the real part of $\mathcal{E}_0 e^{j\omega t}$. Similarly, the component of $\mathcal{E}_0 e^{j\omega t}$ along the axis of the imaginaries is equivalent to $\varepsilon = \mathcal{E}_0 \sin \omega t$.

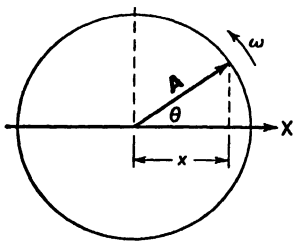


Fig. 13.27

In many calculations involving AC currents, emf's, voltages, and fields the necessary manipulations are simplified by writing the instantaneous values in this exponential form with the understanding that, following the required manipulations, the instantaneous quantity under study is *either* the real *or* the imaginary component of the quantity obtained. This is particularly true when harmonic quantities are to be multiplied or divided, as in Problem 13.29. The exponential form is widely employed in the study of electromagnetic waves since the constituent electric and magnetic fields are both harmonic functions of time.

13.23 Problems

E (13.1) A rectangular coil of dimensions 5 x 20 cm and containing 800 turns of wire is rotated at 3600 rpm in a uniform magnetic field whose flux density is 0.05 weber/m^2. The axis of rotation is parallel to the long side of the coil and perpendicular to the direction of the magnetic field. Assuming the plane of the coil to be parallel to the magnetic field at $t = 0$, write the expression for the instantaneous value of the induced emf.
[Ans.: $\varepsilon = 151 \sin (377t \pm \pi/2)$.]

E (13.2) Show that the average value of an alternating current during the half cycle that it is positive is equal to $2I_0/\pi$.

E (13.3) An ideal half-wave rectifier transmits the positive half of an AC current while suppressing the negative half. Calculate both RMS and MRS values of such a current.

M (13.4) Calculate the effective or "virtual" voltage for $\varepsilon = 100 \sin 2\pi ft + 30 \sin 6\pi ft$. [Ans.: 73.8 volts.]

E (13.5) A 400-cycle AC voltage of 100 volts is applied across the terminals of a series circuit consisting of 10 ohms resistance and 2 milli-

henrys inductance. Calculate the current in the circuit and the phase angle by which the current lags the voltage.

[Ans.: 8.9 amp., 26.7°.]

E (13.6) A DC voltage of 100 volts, when impressed across the terminals of a series circuit, gives rise to a DC current of 0.50 amperes. When the DC voltage is replaced by a 60-cycle AC voltage of 100 volts an AC meter in the circuit reads 0.40 ampere. What are the values of R, L, and C in the circuit?

E (13.7) An AC voltage of amplitude 150 volts is impressed across a pure resistance of 100 ohms. Calculate the current and power dissipation when the frequency of the voltage is (a) 60 cycles/sec; (b) 400 cycles/sec.

E (13.8) Suppose in Problem 13.7 the resistance is replaced by a pure inductance of 0.25 henry. Calculate the current and power loss at the two frequencies given.

E (13.9) A 60-cycle/sec source of emf is impressed across a series circuit consisting of a 10-ohm resistance and a capacitance of unknown value. A meter across the source reads 100 volts and another in the circuit reads 0.2 amperes. Calculate (a) the impedance of the circuit; (b) the power consumption; (c) the power factor; and (d) the value of the capacitance.

E (13.10) A group of theater lights requiring a total current of 20 amperes is connected to a 550-volt line (60~) through a 20-ohm variable resistance. How much money would be saved over a period of 1000 hours by replacing the 20-ohm resistance with a suitable choke coil? What value of inductance should the choke possess? Electrical energy costs 4¢ per KWH. [Ans.: $320, 70 mh.]

E (13.11) The voltage across the inductance and resistance in a series LR circuit is $v_{LR} = 141.4 \cos \omega t$ and the current through the resistive portion of the circuit is $i_R = 2.80 \cos (\omega t - \pi/6)$. Calculate the power loss for the circuit. [Ans.: 173 watts.]

E (13.12) A 60-cycle AC voltage of 100 volts impressed across an LR circuit results in a current of 2 amperes. If the power dissipation is 100 watts, calculate the maximum value of the back emf arising in the inductance. [Ans.: 122 volts.]

E (13.13) Assume Figure 13.16 represents the input to a radio receiver. The resistance R is 100 ohms and L has the value 0.25 millihenry. If two signals of equal strength, of 1000 KC/sec and 950 KC/sec, are received at the antenna, and the receiver is tuned to the former, what fraction of the total signal voltage on the grid is caused by the 950-KC component? [Ans.: 0.32.]

E (13.14) The extent to which a voltage has been attenuated, or the difference in level between two voltages, is conveniently expressed in decibels $= 20 \log_{10} (V_1/V_2)$. In the preceding problem, by how many db is the 950-KC grid signal down with respect to that of the tuned frequency? [Ans.: 6.4 db.]

M (13.15) A parallel circuit consists of three parallel branches. The first branch is pure resistance and carries a current of 5 amperes, the current in the second branch is 4 amperes and lags $\pi/3$ behind the voltage, while the current in the third branch is 7.5 amperes and lags the voltage by $\pi/4$. Calculate the total current and the power factor. [Ans.: 15.2 amperes, 0.8.]

E (13.16)[9] How does division by j affect the direction and magnitude of the vector representing $\boldsymbol{Z} = R + jX$?

E (13.17) Calculate the modulus of the complex quantity $\boldsymbol{Z} = (2 + j4)/(5 - j7)$.

E (13.18) What is the modulus of the complex quantity $Ae^{j\omega t}$?

E (13.19) Show, using the exponential form of a complex quantity, that the modulus of a product or quotient of two complex quantities is equal to the product or quotient of the moduli.

M (13.20) Consider a parallel circuit of a pure inductance, L, shunted by a pure capacitance, C. Use complex algebra to calculate the frequency of resonance and the resistance of the circuit at resonance.

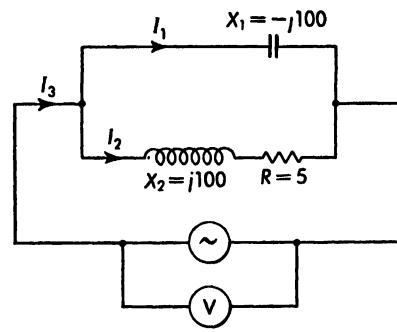


Fig. 13.28

E (13.21) In the circuit of Figure 13.28 the voltmeter, V, reads 100 volts. Compute the complex expressions for the currents I_1, I_2, and I_3 and also their RMS values. (A moral to be noted in this problem is the fact that substantial currents may exist *within* a parallel resonant circuit even though the current *through* it is small. This large internal current is associated with oscillations within the parallel network.)

[Ans.: $\boldsymbol{I}_1 = j$ amperes, $\boldsymbol{I}_2 \doteq 0.05 - j$ amperes, $\boldsymbol{I}_3 \doteq 0.05$ ampere, $I_1 = 1$ ampere, $I_2 \doteq 1$ ampere, $I_3 \doteq 0.05$ ampere.]

M (13.22) What value of the impedance, $\boldsymbol{Z}$, in the circuit of Figure 13.29 will lead to the largest reading of the ammeter, A? [Ans.: $\boldsymbol{Z} = -j0.15$.]

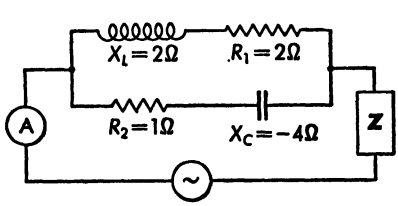


Fig. 13.29

[9] The remaining problems involve the use of complex algebra.

M (13.23) Show that the effective inductance of the parallel network shown in Figure 13.30 is given by $L = (L_1L_2 - M^2)/(L_1 + L_2 \mp 2M)$.

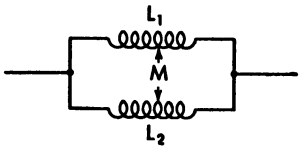


Fig. 13.30

E (13.24) Four admittances of $0.04 + j0.06$, $0.07 - j0.09$, $0.03 - j0.04$, and $0.06 + j0.07$ mhos, respectively, are connected in parallel across an AC voltage of 100 volts. What is the total current in the circuit and what is the phase angle between it and the applied voltage? [Ans.: 20 amperes, 0°.]

E (13.25) The primary of a transformer is connected to a 60-cycle AC generator of peak voltage 141 volts and internal impedance equal to 23 ohms pure resistance. If the RMS secondary voltage is 500 volts, what resistance should the secondary load possess if maximum power is to be delivered to it? [Ans.: 575 ohms.]

E (13.26) In the circuit of Figure 13.31, what value of $\boldsymbol{Z}_l$ results in maximum power being delivered to the secondary circuit? Under these circumstances, what power is actually delivered? [Ans.: $80 + j160$; 125 watts.]

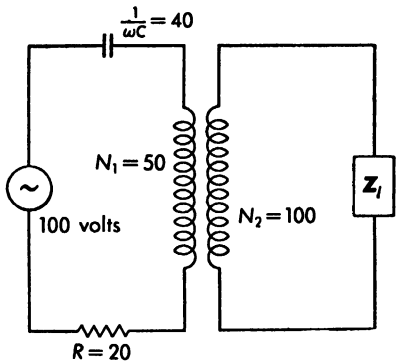


Fig. 13.31

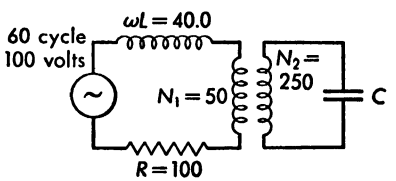


Fig. 13.32

M (13.27) In the circuit of Figure 13.32, what value of C leads to maximum current in the primary circuit? [Ans.: $C = 2.65$ μf.]

M (13.28) Show that the sum of two AC currents of the same frequency is also harmonic with the same frequency.

E (13.29) Write $\varepsilon = \mathcal{E}_0 \cos \omega t$ and $i = I_0 \cos(\omega t + \pi/2)$ in the exponential form and show, with the help of these expressions, that the power is also harmonic but with a frequency which is twice that of the voltage or current.

Chapter 14

Maxwell's Equations and Electromagnetic Waves

We have briefly digressed, at four different points in the text, to draw attention to Maxwell's equations. These four differential equations express, in a completely general way, the several relationships which exist between the field vectors, $\boldsymbol{E}$, $\boldsymbol{D}$, $\boldsymbol{B}$ and $\boldsymbol{H}$. The equations are

$$\nabla \cdot \boldsymbol{D} = \rho \qquad [4\text{-}26]$$

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} \qquad [8\text{-}49]$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad [9\text{-}18]$$

and

$$\nabla \cdot \boldsymbol{B} = 0 \qquad [10\text{-}55]$$

Any of the subjects that we have discussed so far, whether in electrostatics, current electricity, or electromagnetism, could be treated as special applications of Maxwell's equations. In practice, in introductory courses in electricity in magnetism, this elegant procedure is not followed for two reasons. First, it is only toward the end of such a course, after studying many special aspects of electricity and magnetism, that the student is enlightened to the point where he can appreciate the significance of Maxwell's equations. Second, certain portions of the subject, in particular those dealing with current electricity

(either DC or AC), are more conveniently handled by circuit theory than by using Maxwell's equations as the starting point. In this chapter the equations will be used to investigate certain phenomena, notably the existence and propagation of electromagnetic waves, which cannot be studied in any other context.

14.1 Plane Electromagnetic Waves in Free Space

We shall consider the collective implications, in free space, of Maxwell's four equations. We shall discover that a "wave equation" exists for each of the four field vectors. From these wave equations we shall surmise the existence of electromagnetic waves and deduce certain of their properties.

In free space the charge density ρ is zero. Also, since the conductivity of free space is zero, the current density, $\boldsymbol{J} = \sigma \boldsymbol{E}$ (from equation [5-13]—note that ρ in this equation represents the resistivity $= 1/\sigma$), is zero. Maxwell's equations then reduce, in free space, to the form

$$\nabla \cdot \boldsymbol{D} = 0 \qquad [14\text{-}1]$$

$$\nabla \times \boldsymbol{H} = \frac{\partial \boldsymbol{D}}{\partial t} \quad \text{or} \quad \nabla \times \boldsymbol{H} = \epsilon_0 \frac{\partial \boldsymbol{E}}{\partial t} \qquad [14\text{-}2]$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \quad \text{or} \quad \nabla \times \boldsymbol{E} = -\mu_0 \frac{\partial \boldsymbol{H}}{\partial t} \qquad [14\text{-}3]$$

and

$$\nabla \cdot \boldsymbol{B} = 0 \qquad [14\text{-}4]$$

We may eliminate either $\boldsymbol{H}$ or $\boldsymbol{E}$ from equations [14-2] and [14-3]. For example, to eliminate $\boldsymbol{H}$, we take the curl of both sides of equation [14-3]:

$$\nabla \times (\nabla \times \boldsymbol{E}) = -\mu_0 \frac{\partial}{\partial t} (\nabla \times \boldsymbol{H}) \qquad [14\text{-}5]$$

and substitute equation [14-2] therein. Thus,

$$\nabla \times (\nabla \times \boldsymbol{E}) = -\epsilon_0 \mu_0 \frac{\partial^2 \boldsymbol{E}}{\partial t^2} \qquad [14\text{-}6]$$

In Problem 14.1 the vector identity

$$\nabla \times (\nabla \times \boldsymbol{E}) = \nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E} \qquad [14\text{-}7]$$

is established. The first term on the right-hand side is zero, in the light of equation [14-1]. Hence, equation [14-6] may be rewritten

$$\nabla^2 \boldsymbol{E} = \epsilon_0 \mu_0 \frac{\partial^2 \boldsymbol{E}}{\partial t^2} \qquad [14\text{-}8]$$

Similarly, by eliminating $\boldsymbol{E}$ between equations [14-2] and [14-3], it is found that

$$\nabla^2 \boldsymbol{H} = \epsilon_0 \mu_0 \frac{\partial^2 \boldsymbol{H}}{\partial t^2} \qquad [14\text{-}9]$$

The reader can easily demonstrate to himself that equations similar to [14-8] and [14-9] also exist for the other two field vectors, $\boldsymbol{D}$ and $\boldsymbol{B}$. As mentioned in § 2.4, the written-out form of the Laplacian operator, in Cartesian coordinates, is

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \qquad [14\text{-}10]$$

Equations of the type of [14-8] and [14-9], relating the space and time variations of a quantity, had been encountered in physics long before the time of Maxwell, and were known to represent wave motion. Before pursuing this subject further, we shall digress to recall certain facts concerning harmonic waves.

Digression Concerning Harmonic Waves

The reader may recall from courses in general physics that a transverse, plane, harmonic wave traveling in the positive Y-direction is represented by the function

$$x = A \cos 2\pi \left(\frac{t}{T} - \frac{y}{\lambda} \right) \qquad [14\text{-}11]$$

where T is the period of the motion, λ is the wavelength, A is the amplitude of the wave, and x is the transverse displacement of the wave at time t and position y. In a time equal to one period a point on the wave of constant phase advances a distance equal to one wavelength. The *phase velocity*, v, of the wave is therefore given by

$$v = \frac{\lambda}{T} = \lambda f \qquad [14\text{-}12]$$

where $f = 1/T$ is the frequency of the motion. One who is not familiar with equation [14-11] can readily convince himself that the wave described by it does indeed travel λ meters in T seconds (Problem 14.2).

Solution of the Wave Equation

To solve equation [14-8] let us take a trial solution of the form of equation [14-11]:

$$\boldsymbol{E} = \boldsymbol{E}_0 \cos 2\pi \left(\frac{t}{T} - \frac{y}{\lambda} \right) \qquad [14\text{-}13]$$

This, when substituted in [14-8], yields

$$4\pi^2\left[\frac{1}{\lambda^2}-\frac{\epsilon_0\mu_0}{T^2}\right]\boldsymbol{E}=0 \qquad [14\text{-}14]$$

Evidently, equation [14-13] is a solution of equation [14-8], provided

$$\frac{1}{\lambda^2}=\frac{\epsilon_0\mu_0}{T^2} \qquad [14\text{-}15]$$

or

$$\nu^2=\frac{\lambda^2}{T^2}=\frac{1}{\epsilon_0\mu_0} \qquad [14\text{-}16]$$

We therefore conclude that equation [14-8] describes a transverse, harmonically varying wave moving in the positive Y-direction, with velocity

$$\nu=\frac{1}{\sqrt{\epsilon_0\mu_0}} \qquad [14\text{-}17]$$

Further, since the value of $\boldsymbol{E}$ is independent of the x and z coordinates, the wave is a plane wave. Likewise, the $\boldsymbol{H}$, $\boldsymbol{D}$, and $\boldsymbol{B}$ wave equations are satisfied by solutions similar to equation [14-13], which also represent plane waves moving in the positive Y-direction with phase velocity given by equation [14-17]. On the basis of this type of argument Maxwell concluded that all electromagnetic effects are propagated in a vacuum [1] as a wave and with the uniform velocity $\nu=\sqrt{1/\epsilon_0\mu_0}$.

Since ϵ_0 and μ_0 are numerical constants whose value can be determined experimentally, the velocity of propagation of the wave can be calculated. The value of the velocity thus obtained is

$$\begin{aligned}\nu&=(8.854\times10^{-12}\times4\pi\times10^{-7})^{-\frac{1}{2}}\ \text{m/sec}\\&=2.998\times10^{8}\ \text{m/sec}\end{aligned} \qquad [14\text{-}18]$$

Within the limits of experimental error this velocity is identical with the observed velocity of light in a vacuum. From this Maxwell inferred that the propagation of light is a special case of the propagation of electromagnetic waves.[2]

Maxwell's speculations were confirmed (1889) by Heinrich Rudolph Hertz, professor at the Polytechnic Institute, Karlsruhe. In the course of his experiments, Hertz first produced electromagnetic waves (§ 12.8) and later showed them to be transverse waves that travel with the velocity of light. He further demonstrated that such waves could be reflected, refracted, and polar-

[1] Although the foregoing analysis has been developed for the case of a vacuum, it holds for a dielectric as well, provided ϵ_0 and μ_0 are replaced by ϵ and μ for the dielectric.

[2] From now on we shall write the velocity of electromagnetic waves in a vacuum as c, the velocity of light.

ized in the same manner as visible light. The wavelength of these first "hertzian" waves was about 4 meters.

Relative Directions of E and H

We now discuss the directional relationships that exist in the plane electromagnetic wave between the electric field, $\boldsymbol{E}$, its associated magnetic vector, $\boldsymbol{H}$, and the direction of propagation of the wave.

In the same way that [14-13] is a solution of the $\boldsymbol{E}$ wave equation [14-8], a solution of the $\boldsymbol{H}$ wave equation [14-9] is

$$\boldsymbol{H} = \boldsymbol{H}_0 \cos 2\pi \left(\frac{t}{T} - \frac{y}{\lambda}\right) \qquad [14\text{-}19]$$

This also represents a plane, harmonic wave traveling in the positive Y-direction. With a view to making use of equation [14-2], let us take the curl of equation [14-19].

$$\nabla \times \boldsymbol{H} = \nabla \times \left[\boldsymbol{H}_0 \cos 2\pi \left(\frac{t}{T} - \frac{y}{\lambda}\right)\right]$$

$$= \cos 2\pi \left(\frac{t}{T} - \frac{y}{\lambda}\right) \nabla \times \boldsymbol{H}_0 + \boldsymbol{H}_0 \times \nabla \left[\cos 2\pi \left(\frac{t}{T} - \frac{y}{\lambda}\right)\right] \qquad [14\text{-}20]$$

The first term on the right-hand side is zero since $\boldsymbol{H}_0$ is a constant vector. Equation [14-20] thus becomes

$$\nabla \times \boldsymbol{H} = \boldsymbol{H}_0 \times \boldsymbol{j} \frac{2\pi}{\lambda} \sin 2\pi \left(\frac{t}{T} - \frac{y}{\lambda}\right) \qquad [14\text{-}21]$$

or, making use of equation [14-2],

$$\boldsymbol{H}_0 \times \boldsymbol{j} \frac{2\pi}{\lambda} \sin 2\pi \left(\frac{t}{T} - \frac{y}{\lambda}\right) = \epsilon_0 \frac{\partial \boldsymbol{E}}{\partial t} \qquad [14\text{-}22]$$

From this we see, inasmuch as $\boldsymbol{j}$ points in the direction of propagation of the wave, that $\boldsymbol{E}$ is at right angles to both the magnetic vector and to the direction of wave propagation. We can similarly show, using equation [14-3], that $\boldsymbol{H}$, is normal to both $\boldsymbol{E}$ and $\boldsymbol{j}$. We conclude that for a plane electromagnetic wave traveling in a vacuum, $\boldsymbol{E}$ and $\boldsymbol{H}$ are mutually perpendicular at all times. Further, as implied by [14-13] and [14-19], both $\boldsymbol{E}$ and $\boldsymbol{H}$ lie in the plane normal to the direction of propagation of the wave.

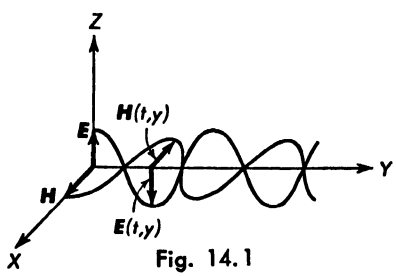


Fig. 14.1

If we choose the Z-axis in such a way that the $\boldsymbol{E}$ vector is parallel to it, Figure 14.1 represents the plane electromagnetic wave that we have been discussing. This type of wave, in which the electric vector is always parallel or antiparallel to a fixed direction, is called a *plane-polarized* wave. The *plane of polarization*, here the ZY-plane, is the one that contains both the electric vector and the direction of propagation. We shall later mention electromagnetic waves that are neither plane nor plane-polarized.

Energy Flow and Poynting's Vector

We may investigate the energy flow associated with the passage of a plane electromagnetic wave by studying the time variation of the density of electromagnetic energy in the region through which the wave is passing. As we have seen in § 9.9, this energy density is given, for free space, by

$$\frac{\boldsymbol{E}\boldsymbol{D} + \boldsymbol{H}\boldsymbol{B}}{2} = \frac{\epsilon_0 E^2 + \mu_0 H^2}{2} \qquad \text{[9-43]}$$

We can obtain an expression for the time variation of this quantity by utilizing the following vector identity (see Problem 14.3)

$$\nabla\cdot(\boldsymbol{E}\times\boldsymbol{H}) = \boldsymbol{H}\cdot(\nabla\times\boldsymbol{E}) - \boldsymbol{E}\cdot(\nabla\times\boldsymbol{H}) \qquad \text{[14-23]}$$

When equations [14-2] and [14-3] are substituted in this equation we obtain

$$\nabla\cdot(\boldsymbol{E}\times\boldsymbol{H}) = -\boldsymbol{H}\cdot\mu_0\frac{\partial \boldsymbol{H}}{\partial t} - \boldsymbol{E}\cdot\epsilon_0\frac{\partial \boldsymbol{E}}{\partial t}$$

$$= -\frac{\partial}{\partial t}(\epsilon_0 E^2 + \mu_0 H^2) \qquad \text{[14-24]}$$

Evidently, the term on the right-hand side is the rate at which the electromagnetic energy per unit volume of field is decreasing. For a volume v the rate of decrease of stored energy is

$$\int_v \nabla\cdot(\boldsymbol{E}\times\boldsymbol{H})\,dv = -\frac{\partial}{\partial t}\int_v (\epsilon_0 E^2 + \mu_0 H^2)\,dv \qquad \text{[14-25]}$$

This decrease can be explained by assuming that energy is flowing out through the surface that bounds the volume under consideration. This assumption is made more plausible by using the divergence theorem to transform the left-hand side of equation [14-25] to a surface integral:

$$\int_s (\boldsymbol{E}\times\boldsymbol{H})\cdot d\boldsymbol{s} = -\frac{\partial}{\partial t}\int_v (\epsilon_0 E^2 + \mu_0 H^2)\,dv \qquad \text{[14-26]}$$

Thus, the vector $\boldsymbol{E} \times \boldsymbol{H}$, usually written as $\boldsymbol{S}$, is a vector representing energy flow. By integrating it over a closed surface, we are able to compute the rate at which energy is outflowing through the surface, or the rate at which it is being lost to the volume enclosed within the surface. Equation [14-26] clearly expresses these facts.

The vector

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H} \qquad [14\text{-}27]$$

is known as *Poynting's vector*. As can be seen from its defining equation, $\boldsymbol{S}$ for an electromagnetic wave points in the direction of propagation. Hence, an electromagnetic wave transports energy in the direction in which the wave is traveling. Since it gives the rate of energy flow, $\boldsymbol{S}$ represents the flux of power per unit area. The total power flowing through any surface is, therefore,

$$\text{Power} = \int_s \boldsymbol{S} \cdot d\boldsymbol{s} \qquad [14\text{-}28]$$

The units of $\boldsymbol{S}$ in the mks system are watts/m^2.

Impedance of Free Space

The dimensions of the ratio $\boldsymbol{E}/\boldsymbol{H}$ are those of impedance. For example, in the mks system,

$$\frac{[E]}{[H]} = \frac{\text{volts/meter}}{\text{amperes/meter}} = \frac{\text{volts}}{\text{amperes}} = \text{ohms} \qquad [14\text{-}29]$$

The value of this ratio is known as the *characteristic impedance* of the medium through which the electromagnetic wave is passing. Let us calculate its value for free space.

Substituting the value of $\boldsymbol{E}$ (equation [14-13]) in equation [14-22] we have

$$\boldsymbol{H}_0 \times \boldsymbol{j} \frac{2\pi}{\lambda} \sin 2\pi \left(\frac{t}{T} - \frac{y}{\lambda} \right) = - \frac{2\pi\epsilon_0}{T} \boldsymbol{E}_0 \sin 2\pi \left(\frac{t}{T} - \frac{y}{\lambda} \right) \qquad [14\text{-}30]$$

from which the relationship between the magnitudes of $\boldsymbol{H}_0$ and $\boldsymbol{E}_0$ is seen to be

$$\frac{H_0}{\lambda} = \frac{\epsilon_0 E_0}{T} \qquad [14\text{-}31]$$

Since $\boldsymbol{E}$ and $\boldsymbol{H}$ are in phase,[8] the value of the ratio E/H is given by the ratio

[8] In choosing equations [14-13] and [14-19] as solutions of the $\boldsymbol{E}$ and $\boldsymbol{H}$ wave equations, we have assumed that $\boldsymbol{E}$ and $\boldsymbol{H}$ are in phase. One can see, however, from equations [14-2] and [14-3] that this phase relationship must exist.

E_0/H_0. Thus, in making use of [14-31], we can write for free space

$$\frac{E}{H} = \frac{E_0}{H_0} = \frac{T}{\epsilon_0 \lambda} = \frac{1}{\epsilon_0 \lambda f} = \frac{1}{\epsilon_0 c} = \sqrt{\frac{\mu_0}{\epsilon_0}} \qquad [14\text{-}32]$$

The numerical value of this characteristic impedance of free space is

$$\sqrt{\frac{\mu_0}{\epsilon_0}} = \sqrt{\frac{4\pi \times 10^{-7}}{8.854 \times 10^{-12}}} = 376.7 \text{ ohms} \qquad [14\text{-}33]$$

14.2 Use of the Exponential Form to Represent Plane Harmonic Waves

Mention has been made, in § 13.22, of the fact that quantities which vary harmonically can be expressed in an exponential form. Thus, for example,

$$\boldsymbol{E} = \boldsymbol{E}_0 e^{j(\omega t + \delta)} \qquad [14\text{-}34]$$

can be written as

$$\boldsymbol{E} = \boldsymbol{E}_0[\cos(\omega t + \delta) + j \sin(\omega t + \delta)] \qquad [14\text{-}35]$$

where both the real component, $\boldsymbol{E}_0 \cos(\omega t + \delta)$, and the imaginary component, $\boldsymbol{E}_0 \sin(\omega t + \delta)$, represent quantities which vary harmonically with time with angular frequency ω and amplitude $\boldsymbol{E}_0$. Apart from a phase difference of $\pi/2$ radians, these two components are identical. Hence, *either* the real *or* the imaginary component of the exponential expression is suitable for describing the harmonic function in question. The value of δ depends upon which component is used.

The exponential expression for a plane, harmonic wave is

$$\boldsymbol{E} = \boldsymbol{E}_0 e^{j(\omega t - \boldsymbol{\kappa} \cdot \boldsymbol{r})} \qquad [14\text{-}36]$$

where

$$\boldsymbol{\kappa} = i\kappa_x + j\kappa_y + k\kappa_z \qquad [14\text{-}37]$$

is a vector (of magnitude to be ascertained below) pointing in the direction of propagation and, as usual, $\boldsymbol{r} = ix + jy + kz$ is the radius vector from the origin to a representative point. The product $\boldsymbol{\kappa} \cdot \boldsymbol{r}$ has the value

$$\boldsymbol{\kappa} \cdot \boldsymbol{r} = \kappa_x x + \kappa_y y + \kappa_z z \qquad [14\text{-}38]$$

We could also write $\boldsymbol{\kappa} \cdot \boldsymbol{r}$ as κr_κ, where r_κ is the component of $\boldsymbol{r}$ in the direction of $\boldsymbol{\kappa}$, that is, in the direction normal to the plane of constant phase

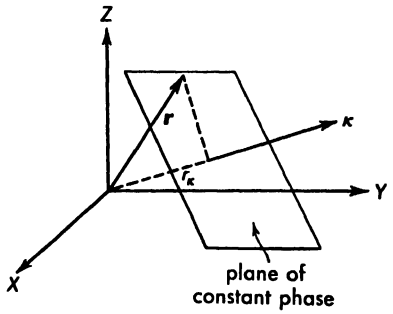


Fig. 14.2

(Fig. 14.2). The value of r_κ is given by the length of the perpendicular from the origin to the plane in question. Hence, as equation [14-36] assumes, the quantity $\boldsymbol{\kappa} \cdot \boldsymbol{r}$ is constant for all points on a plane of constant phase.

By substituting equation [14-36] in the wave equation,

$$\nabla^2 \boldsymbol{E} = \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \qquad [14\text{-}8]$$

which describes the space and time behavior of the electric field in free space, we can easily see that it is a solution, provided κ has a certain value. First, we substitute equation [14-36] in the left side of [14-8]:

$$\nabla^2 \boldsymbol{E} = \nabla^2 \boldsymbol{E}_0 e^{j(\omega t - \boldsymbol{\kappa} \cdot \boldsymbol{r})} \qquad [14\text{-}39]$$

$$= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \boldsymbol{E}_0 e^{j(\omega t - \kappa_x x - \kappa_y y - \kappa_z z)} \qquad [14\text{-}40]$$

$$= -(\kappa_x^2 + \kappa_y^2 + \kappa_z^2)\boldsymbol{E} \qquad [14\text{-}41]$$

$$= -\kappa^2 \boldsymbol{E} \qquad [14\text{-}42]$$

Then, substituting equation [14-36] in the right side of equation [14-8], we obtain

$$\epsilon_0 \mu_0 \frac{\partial^2 \boldsymbol{E}}{\partial t^2} = \epsilon_0 \mu_0 (-\omega^2) \boldsymbol{E} \qquad [14\text{-}43]$$

From [14-42] and [14-43] we see that equation [14-36] is a solution of [14-8] provided

$$\kappa^2 = \epsilon_0 \mu_0 \omega^2$$

or

$$\kappa = \omega \sqrt{\epsilon_0 \mu_0} \qquad [14\text{-}44]$$

The vector $\boldsymbol{\kappa}$ is known as the *propagation vector*.

We therefore conclude that an acceptable solution of the $\boldsymbol{E}$ wave equation is

$$\boldsymbol{E} = \boldsymbol{E}_0 e^{j(\omega t - \boldsymbol{\kappa} \cdot \boldsymbol{r})} \qquad [14\text{-}36]$$

and that the corresponding solution (as can easily be proved) of the $\boldsymbol{H}$ wave equation is

$$\boldsymbol{H} = \boldsymbol{H}_0 e^{j(\omega t - \boldsymbol{\kappa} \cdot \boldsymbol{r})} \qquad [14\text{-}45]$$

In both these solutions κ has the value $\omega \sqrt{\epsilon_0 \mu_0} = 2\pi f/c = 2\pi/\lambda$.

Phase Velocity

In the preceding line we have assumed, as established earlier, that the phase velocity of the electromagnetic wave in free space is given by $c = 1/\sqrt{\epsilon_0\mu_0}$. This, however, we can readily deduce from equation [14-36] (or [14-45]) by calculating the velocity of a plane of constant phase.

As noted earlier, $\boldsymbol{\kappa}\cdot\boldsymbol{r}$ has the magnitude κr_κ, where r_κ is the component of $\boldsymbol{r}$ in the direction normal to the plane of constant phase. r_κ has the same value at all points in the plane. For a plane of constant phase the quantity (cf. equation [14-36])

$$\omega t - \boldsymbol{\kappa}\cdot\boldsymbol{r} = \text{constant} \qquad [14\text{-}46]$$

That is,

$$\omega t_1 - (\boldsymbol{\kappa}\cdot\boldsymbol{r})_{t_1} = \omega(t_1 + T) - (\boldsymbol{\kappa}\cdot\boldsymbol{r})_{t_1+T} \qquad [14\text{-}47]$$

or

$$\omega T = (\boldsymbol{\kappa}\cdot\boldsymbol{r})_{t_1+T} - (\boldsymbol{\kappa}\cdot\boldsymbol{r})_{t_1} \qquad [14\text{-}48]$$

The quantity on the right side of the above equation is κ multiplied by the distance, λ, which the plane of constant phase advances in one period, T. Hence, equation [14-48] may be rewritten

$$\omega T = \kappa\lambda \qquad [14\text{-}49]$$

from which the phase velocity, $c = \lambda/T$, is seen to have the value

$$c = \frac{\lambda}{T} = \frac{\omega}{\kappa} = \frac{\omega}{\omega\sqrt{\epsilon_0\mu_0}} = \frac{1}{\sqrt{\epsilon_0\mu_0}} \qquad [14\text{-}50]$$

as deduced earlier. This, it is emphasized, is the velocity with which a plane of constant phase advances: it applies only to wave trains of infinite length—the type that we are here considering.

14.3 Reflection and Refraction of Plane Electromagnetic Waves

The preceding description of a plane electromagnetic wave traveling in free space applies equally well if the medium is other than a vacuum, provided it is a perfect dielectric. The appropriate values of ϵ and μ must, of course, be used. In this section we shall study the behavior of such a plane wave when it arrives at the interface between two dissimilar dielectrics ($\epsilon_1 \neq \epsilon_2$). We shall limit our attention to nonmagnetic materials, that is, those for which $\mu_1 = \mu_2 = \mu_0$.

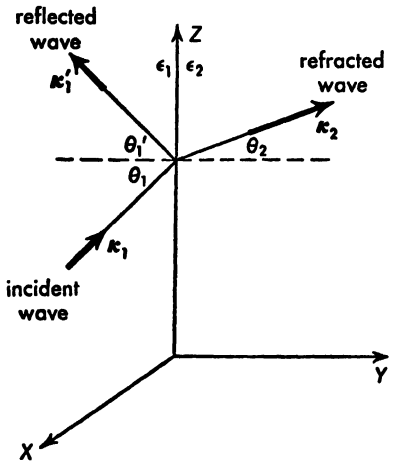


Fig. 14.3

Let us assume, as indicated in Figure 14.3, that the interface between the two dielectrics coincides with the $y = 0$ plane. The direction of propagation of the plane wave that is incident on the interface is given by $\boldsymbol{\kappa}_1$. In general, a portion of the incident wave is reflected at the interface, thereafter having a propagation vector $\boldsymbol{\kappa}_1'$, while the remaining portion of the wave enters the second medium where it proceeds in the direction of $\boldsymbol{\kappa}_2$. The boundary conditions at the interface cannot be satisfied unless both a reflected and refracted wave are postulated.

For simplicity, let us suppose that $\boldsymbol{\kappa}_1$ lies in the YZ-plane. We shall also tentatively assume that the reflected and refracted waves are both plane harmonic. Thus,

for the incident wave $$\boldsymbol{E}_1 = \boldsymbol{E}_{0_1} e^{j(\omega_1 t - \boldsymbol{\kappa}_1 \cdot \boldsymbol{r})} \qquad [14\text{-}51]$$

for the reflected wave $$\boldsymbol{E}_1' = \boldsymbol{E}_{0_1}' e^{j(\omega_1' t - \boldsymbol{\kappa}_1' \cdot \boldsymbol{r} + \delta_1')} \qquad [14\text{-}52]$$

and for the refracted wave $$\boldsymbol{E}_2 = \boldsymbol{E}_{0_2} e^{j(\omega_2 t - \boldsymbol{\kappa}_2 \cdot \boldsymbol{r} + \delta_2)} \qquad [14\text{-}53]$$

where δ_1' and δ_2 in equations [14-52] and [14-53] allow for possible phase shifts with respect to the incident wave.

We can now apply the boundary condition derived in § 4.7, namely, that the tangential component of the electric field is continuous across the interface between the two media (that is, at $y = 0$). In this case, at every point in the interface

$$E_{1_{\text{tang}}} + E_{1_{\text{tang}}}' = E_{2_{\text{tang}}} \qquad [14\text{-}54]$$

or $$E_{0_{1_{\text{tang}}}} e^{j(\omega_1 t - \boldsymbol{\kappa}_1 \cdot \boldsymbol{r})} + E_{0_{1_{\text{tang}}}}' e^{j(\omega_1' t - \boldsymbol{\kappa}_1' \cdot \boldsymbol{r} + \delta_1')} = E_{0_{2_{\text{tang}}}} e^{j(\omega_2 t - \boldsymbol{\kappa}_2 \cdot \boldsymbol{r} + \delta_2)} \qquad [14\text{-}55]$$

Rewriting equation [14-55] in somewhat expanded form, and keeping in mind that $y = 0$ at the interface, we have

$$E_{0_{1_{\text{tang}}}} e^{j\omega_1 t} e^{-j(\kappa_{1_x} x + \kappa_{1_z} z)} + E_{0_{1_{\text{tang}}}}' e^{j\omega_1' t} e^{-j(\kappa_{1_x}' x + \kappa_{1_z}' z + \delta_1')} = E_{0_{2_{\text{tang}}}} e^{j\omega_2 t} e^{-j(\kappa_{2_x} x + \kappa_{2_z} z + \delta_2)} \qquad [14\text{-}56]$$

Since this equality is independent of the time, it follows that

$$\omega_1 = \omega_1' = \omega_2 \qquad [14\text{-}57]$$

Thus, the frequency is the same for the incident, reflected, and refracted wave.

For equation [14-56] to hold at all points in the interface, we must have

$$\kappa_{1_x} = \kappa'_{1_x} = \kappa_{2_x} \qquad [14\text{-}58]$$

and

$$\kappa_{1_z} = \kappa'_{1_z} = \kappa_{2_z} \qquad [14\text{-}59]$$

Since $\boldsymbol{\kappa}_1$ lies in the YZ-plane, $\kappa_{1_x} = 0$ and likewise $\kappa'_{1_x} = \kappa_{2_x} = 0$. That is, both $\boldsymbol{\kappa}'_1$ and $\boldsymbol{\kappa}_2$ also lie in the XY-plane.

Law of Reflection

Since $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}'_1$ lie in the same medium, their magnitudes $(= \omega\sqrt{\epsilon_1\mu_0})$ are identical. Hence, rewriting

$$\kappa_{1_z} = \kappa'_{1_z} \qquad [14\text{-}59]$$

in the form

$$\kappa_1 \sin\theta_1 = \kappa'_1 \sin\theta'_1 \qquad [14\text{-}60]$$

$$= \kappa_1 \sin\theta'_1 \qquad [14\text{-}61]$$

we may conclude that

$$\theta_1 = \theta'_1 \qquad [14\text{-}62]$$

This is the familiar law governing the reflection of light—the angle of incidence is equal to the angle of reflection. We see now that it holds not only for light, but for other electromagnetic waves as well.

Law of Refraction

Similarly, rewriting

$$\kappa_{1_z} = \kappa_{2_z} \qquad [14\text{-}59]$$

in the form

$$\kappa_1 \sin\theta_1 = \kappa_2 \sin\theta_2 \qquad [14\text{-}63]$$

or

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{\kappa_2}{\kappa_1} = \frac{\omega\sqrt{\epsilon_2\mu_0}}{\omega\sqrt{\epsilon_1\mu_0}} = \frac{v_1}{v_2} = \frac{n_2}{n_1} \qquad [14\text{-}64]$$

where n_1 and n_2 are the indices of refraction of the two media,[4] we obtain Snell's law governing the refraction of light:

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1} \qquad [14\text{-}65]$$

The correct predictions expressed in equations [14-62] and [14-65], concerning the directions of the reflected and refracted waves, provided further confirmation of the correctness of Maxwell's equations.

[4] The relationship between n and v, implicit in equation [14-64], has been discussed in § 4.9.

14.4 Amplitudes and Phases of Plane Waves Reflected and Refracted at Interfaces Between Different Dielectric Media

Although we shall not work out the full solution of this problem, we shall indicate for a useful special case the manner in which the boundary conditions are applied.

As we have noted earlier, $\boldsymbol{E}$ and $\boldsymbol{H}$ for the plane wave are normal to one another, and both are normal to the propagation vector $\boldsymbol{\kappa}$. We shall consider the special case in which $\boldsymbol{E}$ of the incident wave is perpendicular to the YZ-plane (Fig. 14.4), that is, $\boldsymbol{E}_1 = i E_1$. $\boldsymbol{H}_1$ then lies in the plane of the paper, at right angles to $\boldsymbol{\kappa}_1$. Under these conditions, the tangential component of the electric field is identical with the field itself. Equation [14-54], expressing the boundary condition for the electric field, becomes

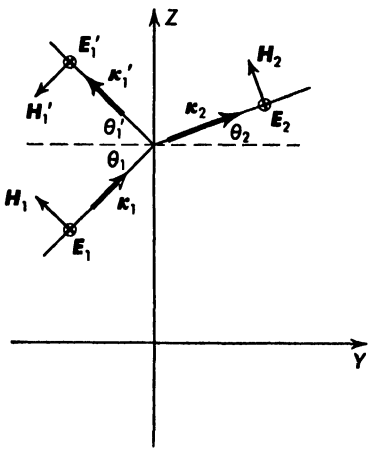


Fig. 14.4

$$E_1 + E_1' = E_2 \qquad [14\text{-}66]$$

We now apply the corresponding boundary condition for the magnetic field (equation [10-62]), namely, that the tangential component of $\boldsymbol{H}$ is continuous across the interface. The tangential component of $\boldsymbol{H}$ is, in this case, its Z-component (Fig. 14.4), and can be found from equation [9-18]. For the incident wave, equation [9-18] has the form

$$\nabla \times [\boldsymbol{E}_{0_1} e^{j(\omega t - \kappa_1 \cdot r)}] = -\mu_0 \frac{\partial}{\partial t} [\boldsymbol{H}_{0_1} e^{j(\omega t - \kappa_1 \cdot r)}] \qquad [14\text{-}67]$$

Before evaluating the left side of this equation, we note that $\boldsymbol{E}_0$ is a constant vector and the exponential factor is a scalar, say ψ. Thus, the left side has the value

$$\nabla \times (\boldsymbol{E}_{0_1} \psi) = \psi(\nabla \times \boldsymbol{E}_{0_1}) + \nabla\psi \times \boldsymbol{E}_{0_1} \qquad [14\text{-}68]$$

$$= \nabla\psi \times \boldsymbol{E}_{0_1} \qquad [14\text{-}69]$$

since $\nabla \times \boldsymbol{E}_0$, the curl of a constant vector, is zero. Working out $\nabla\psi$ we have

$$\nabla\psi = \nabla e^{j(\omega t - \kappa_1 \cdot r)} \qquad [14\text{-}70]$$

$$= j\left[\boldsymbol{i}\frac{\partial}{\partial x}(-\kappa_{1y}y - \kappa_{1z}z) + \boldsymbol{j}\frac{\partial}{\partial y}(-\kappa_{1y}y - \kappa_{1z}z)\right.$$

$$\left. + \boldsymbol{k}\frac{\partial}{\partial z}(-\kappa_{1y}y - \kappa_{1z}z)\right]e^{j(\omega t - \boldsymbol{\kappa}_1 \cdot \boldsymbol{r})} \qquad [14\text{-}71]$$

$$= -j(\boldsymbol{j}\kappa_{1y} + \boldsymbol{k}\kappa_{1z})e^{j(\omega t - \boldsymbol{\kappa}_1 \cdot \boldsymbol{r})} \qquad [14\text{-}72]$$

or $$\nabla\psi = -j\boldsymbol{\kappa}_1 e^{j(\omega t - \boldsymbol{\kappa}_1 \cdot \boldsymbol{r})} \qquad [14\text{-}73]$$

Making use of this result, and carrying out the differentiation indicated on the right-hand side of equation [14-67], we can write equation [14-67] in the form

$$-j\boldsymbol{\kappa}_1 e^{j(\omega t - \boldsymbol{\kappa}_1 \cdot \boldsymbol{r})} \times \boldsymbol{E}_{0_1} = -\mu_0 j\omega \boldsymbol{H}_{0_1} e^{j(\omega t - \boldsymbol{\kappa}_1 \cdot \boldsymbol{r})} \qquad [14\text{-}74]$$

or $$\boldsymbol{\kappa}_1 \times \boldsymbol{E}_1 = \omega\mu_0 \boldsymbol{H}_1 \qquad [14\text{-}75]$$

Solving for $\boldsymbol{H}_1$ we have

$$\boldsymbol{H}_1 = \frac{\boldsymbol{\kappa}_1 \times \boldsymbol{E}_1}{\omega\mu_0} \qquad [14\text{-}76]$$

$$= \frac{(\boldsymbol{j}\kappa_{1y} + \boldsymbol{k}\kappa_{1z}) \times \boldsymbol{i}E_1}{\omega\mu_0}$$

or $$\boldsymbol{H}_1 = \frac{(\boldsymbol{j}\kappa_{1z} - \boldsymbol{k}\kappa_{1y})E_1}{\omega\mu_0} \qquad [14\text{-}77]$$

Hence, the tangential component of $\boldsymbol{H}_1$, H_{1z}, is

$$H_{1z} = -\frac{\kappa_{1y}E_1}{\omega\mu_0} \qquad [14\text{-}78]$$

The tangential components of $\boldsymbol{H}_1'$ and $\boldsymbol{H}_2$ are obtained in a similar manner, and have the same form as equation [14-78].

We are now in a position to apply the magnetic field boundary conditions:

$$H_{1_{\text{tang}}} + H'_{1_{\text{tang}}} = H_{2_{\text{tang}}} \qquad [14\text{-}79]$$

We refer to Figure 14.4 for the proper signs of the different terms. Hence, we have

$$-\frac{\kappa_{1y}E_1}{\omega\mu_0} + \frac{\kappa'_{1y}E'_1}{\omega\mu_0} = -\frac{\kappa_{2y}E_2}{\omega\mu_0} \qquad [14\text{-}80]$$

or $$E_1\kappa_1 \cos\theta_1 - E'_1\kappa_1 \cos\theta_1 = E_2\kappa_2 \cos\theta_2 \qquad [14\text{-}81]$$

or $$E_1\sqrt{\epsilon_1}\cos\theta_1 - E'_1\sqrt{\epsilon_1}\cos\theta_1 = E_2\sqrt{\epsilon_2}\cos\theta_2 \qquad [14\text{-}82]$$

This equation can now be solved simultaneously with equation [14-66] to obtain relationships among E_1, E_1', and E_2.

Making use of the fact that $n \propto \sqrt{\epsilon}$, we may rewrite equation [14-82]:

$$E_1 n_1 \cos\theta_1 - E_1' n_1 \cos\theta_1 = E_2 n_2 \cos\theta_2 \qquad [14\text{-}83]$$

Multiplying equation [14-66] by $n_1 \cos\theta_1$ and adding the result to the above, we have:

$$2E_1 n_1 \cos\theta_1 = E_2(n_1 \cos\theta_1 + n_2 \cos\theta_2) \qquad [14\text{-}84]$$

whence

$$E_2 = E_1 \frac{2\cos\theta_1}{\left(\cos\theta_1 + \frac{n_2}{n_1}\cos\theta_2\right)} \qquad [14\text{-}85]$$

or, by eliminating E_2 between [14-66] and [14-83], we obtain

$$E_1' = E_1 \frac{\left(\cos\theta_1 - \frac{n_2}{n_1}\cos\theta_2\right)}{\left(\cos\theta_1 + \frac{n_2}{n_1}\cos\theta_2\right)} \qquad [14\text{-}86]$$

These two equations may be simplified somewhat by substituting

$$\frac{n_1}{n_2} = \frac{\sin\theta_2}{\sin\theta_1} \qquad [14\text{-}65]$$

When this is done, we obtain

$$E_2 = E_1 \frac{2\cos\theta_1 \sin\theta_2}{\sin(\theta_2 + \theta_1)} \qquad [14\text{-}87]$$

and

$$E_1' = E_1 \frac{\sin(\theta_2 - \theta_1)}{\sin(\theta_2 + \theta_1)} \qquad [14\text{-}88]$$

We shall discuss some of the implications of these equations.

Reflected Wave

We see from equation [14-88] that the sign of E_1' relative to E_1 depends upon the sign of the quantity $(\theta_2 - \theta_1)$. When $\theta_2 > \theta_1$, as is the case when $n_1 > n_2$, E_1' and E_1 have the same sign, that is, the electric field of the reflected wave is in phase with that of the incident wave. In this event, δ_1' in equation [14-52] is zero. On the other hand, if $n_1 < n_2$, E_1' is negative relative to E_1, and $\delta_1' = \pi$. That is, there is a phase reversal during reflection.

The case of total reflection will be separately dealt with below. The phase relationships described above do not apply to it.

Refracted Wave

It is evident from equation [14-87] (inasmuch as both θ_1 and θ_2 lie between 0 and $\pi/2$ radians) that E_2, the electric field associated with the refracted wave, is always in phase with the electric field, E_1, belonging to the incident wave. That is, δ_2 in equation [14-53] is zero.

Total Reflection

We shall work out an expression, involving κ_1 and κ_{2_y} from which the condition for total reflection can be deduced.

Since $\boldsymbol{\kappa}_2$ lies in the YZ-plane,

$$\kappa_2^2 = \kappa_{2_y}^2 + \kappa_{2_z}^2 \qquad [14\text{-}89]$$

or

$$\kappa_{2_y}^2 = \kappa_2^2 - \kappa_{2_z}^2 \qquad [14\text{-}90]$$

Upon substitution of $\kappa_2 = \kappa_1(n_2/n_1)$, from equation [14-64], and $\kappa_{2_z} = \kappa_{1_z} = \kappa_1 \sin\theta_1$, from equation [14-59], equation [14-90] becomes

$$\kappa_{2_y}^2 = \kappa_1^2\left[\left(\frac{n_2}{n_1}\right)^2 - \sin^2\theta_1\right] \qquad [14\text{-}91]$$

whence

$$\kappa_{2_y} = \pm\kappa_1\sqrt{\left(\frac{n_2}{n_1}\right)^2 - \sin^2\theta_1} \qquad [14\text{-}92]$$

When θ_1 has such a value that

$$\frac{n_2}{n_1} < \sin\theta_1 \qquad [14\text{-}93]$$

the quantity under the square root sign in equation [14-92] is negative, indicating that κ_{2_y} is an imaginary quantity, say,

$$\kappa_{2_y} = \pm j\alpha \qquad [14\text{-}94]$$

Only one of these two values is physically acceptable, as we shall see below.

Using $\boldsymbol{\kappa}_2$ in the form

$$\boldsymbol{\kappa}_2 = \boldsymbol{j}\kappa_{2_y} + \boldsymbol{k}\kappa_{2_z} \qquad [14\text{-}95]$$

$$= \boldsymbol{j}(\pm j\alpha) + \boldsymbol{k}\kappa_{1_z} \qquad [14\text{-}96]$$

the expression for $\boldsymbol{E}_2$ (equation [14-53]) becomes

$$\boldsymbol{E}_2 = \boldsymbol{E}_{0_2}e^{j(\omega t \mp j\alpha y - \kappa_{1_z}z)} \qquad [14\text{-}97]$$

or

$$\boldsymbol{E}_2 = \boldsymbol{E}_{0_2}e^{j(\omega t - \kappa_{1_z}z)}e^{\pm\alpha y} \qquad [14\text{-}98]$$

Clearly, the value $\kappa_{2_y} = +j\alpha$ leads to an infinite value of $\boldsymbol{E}_2$ at $y = \infty$. This is physically unacceptable. On the other hand, the value $\kappa_{2_y} = -j\alpha$ leads to a factor, $e^{-\alpha y}$, which represents an exponential decay of $\boldsymbol{E}_2$ in the Y-direction. This is the state of affairs when total reflection occurs. Although the incident wave is totally reflected, the wave actually penetrates the second medium, where its amplitude is quickly damped according to the factor $e^{-\alpha y}$. The existence of this damped wave within the second medium has been established in experiments employing electromagnetic waves of a few centimeters wavelength (microwaves).

To summarize, then, Maxwell's equations predict that total reflection of electromagnetic waves will occur at an interface between two dielectric media when the condition represented by equation [14-93] is met.

The preceding results, describing the phases of the reflected and refracted waves relative to the incident wave and setting forth the condition for total reflection, have been derived on the assumption that the incident wave is plane polarized with its electric vector parallel to the X-axis. It can be shown that these results also hold for other directions of the electric vector.

In a general treatment of this type (as in Stratton, *Electromagnetic Theory*, McGraw-Hill, 1941), the electric vector of the incident wave is assigned an arbitrary direction (or course, always normal to $\boldsymbol{\kappa}_1$). It is then resolved into two components, one lying in the YZ-plane and the other perpendicular to it. The behavior at the interface of both components can be investigated by applying the boundary conditions. It is found that the two components are reflected to differing extents, depending upon the angle of incidence. In particular, when

$$\theta_1 + \theta_2 = \pi/2 \qquad [14\text{-}99]$$

the electric vector of the reflected wave is always perpendicular to the YZ-plane. At the same time, in the refracted wave, the electric vector lies in the YZ-plane. It can be shown (Problem 14.6) that under these circumstances

$$\tan\theta_1 = \frac{n_2}{n_1} \qquad [14\text{-}100]$$

The angle of incidence, θ_1, that satisfies this equation is known as *Brewster's angle*. For a wave entering glass ($n_2 = 1.55$) from air ($n_1 = 1$), θ_1 is approximately 57°. Thus, a simple way of obtaining a plane-polarized beam of light is to allow unpolarized light to strike glass at this angle of incidence and make use of either the reflected or transmitted wave. Here again Maxwell's equations predict an effect long known in optics.

14.5 Electromagnetic Waves in a Conducting Medium

Let us consider a conductor that obeys Ohm's law. Then the current density, $\boldsymbol{J}$, is proportional to the electric field, $\boldsymbol{E}$:

$$\boldsymbol{J} = \sigma \boldsymbol{E} \qquad [5\text{-}13]$$

The second of Maxwell's equations [8-49] then becomes

$$\nabla \times \boldsymbol{H} = \sigma \boldsymbol{E} + \epsilon \frac{\partial \boldsymbol{E}}{\partial t} \qquad [14\text{-}101]$$

This can be combined with the third of Maxwell's equations [9-18]:

$$\nabla \times \boldsymbol{E} = -\mu \frac{\partial \boldsymbol{H}}{\partial t} \qquad [9\text{-}18]$$

to obtain (using the same method of eliminating $\boldsymbol{H}$ as in § 14.1)

$$\nabla^2 \boldsymbol{E} - \mu\epsilon \frac{\partial^2 \boldsymbol{E}}{\partial t^2} - \mu\sigma \frac{\partial \boldsymbol{E}}{\partial t} = 0 \qquad [14\text{-}102]$$

This is the wave equation for the electric field in a conducting medium. It will be noticed that equation [14-8], for the electric field in free space, is simply a special case of this ($\mu = \mu_0$, $\epsilon = \epsilon_0$, and $\sigma = 0$).

In solving [14-102] we shall assume, for simplicity, that the wave is moving in the positive Y-direction. We shall use the trial solution

$$\boldsymbol{E} = \boldsymbol{E}_0 e^{j(\omega t - \boldsymbol{\kappa} \cdot \boldsymbol{r})} = \boldsymbol{E}_0 e^{j(\omega t - \kappa_y y)} \qquad [14\text{-}103]$$

When this trial solution is substituted in equation [14-102] we obtain

$$(-j\kappa_y)^2 \boldsymbol{E} - \mu\epsilon(j\omega)^2 \boldsymbol{E} - \mu\sigma(j\omega)\boldsymbol{E} = 0 \qquad [14\text{-}104]$$

Hence, the trial solution is a proper solution provided

$$(-j\kappa_y)^2 - \mu\epsilon(j\omega)^2 - \mu\sigma(j\omega) = 0 \qquad [14\text{-}105]$$

When this equation is solved for κ_y^2 we obtain

$$\kappa_y^2 = \mu\epsilon\omega^2 \left[1 - \frac{j\sigma}{\epsilon\omega}\right] \qquad [14\text{-}106]$$

For good conductors,

$$\frac{\sigma}{\epsilon\omega} \gg 1 \qquad [14\text{-}107]$$

Therefore,

$$\kappa_y \doteq \sqrt{-j\mu\sigma\omega} = (1-j)\sqrt{\frac{\mu\sigma\omega}{2}} = \frac{(1-j)}{\delta} \qquad [14\text{-}108]^5$$

where

$$\delta = \sqrt{\frac{2}{\mu\sigma\omega}} \qquad [14\text{-}109]$$

is known as the *skin depth.* Substituting equation [14-108] in equation [14-103], we obtain the solution of the wave equation for the electric field in a good conductor:

$$\boldsymbol{E} = \boldsymbol{E}_0 e^{j\left(\omega t - \frac{y}{\delta}\right)} e^{-\frac{y}{\delta}} \qquad [14\text{-}110]$$

The first portion of the solution simply represents a harmonic wave moving along the Y-axis. The final factor, however, represents an attenuation of this wave. Indeed, the amplitude of the wave drops to $1/e$ of its value as the wave progresses a distance equal to the skin depth, δ. As is seen from equation [14-109], the skin depth decreases with increasing frequency. For copper, at 60 cycles, δ is 0.86 cm but, at 1 megacycle, it has dropped to .0067 cm.

Skin Effect

One conclusion to be drawn from the above result is that conducting sheets which are used as electromagnetic shields must be thicker than the skin depth. One can calculate the thickness necessary to secure the desired degree of shielding at any given frequency.

It can also be shown that a consequence of the rapid attenuation of high-frequency fields as they penetrate a conductor is that high-frequency currents in wires are concentrated in a shallow surface layer. This effect becomes more pronounced as the frequency increases, and has the result that the resistance of a wire increases with frequency. The effective cross-sectional area of the wire is simply decreasing. For high-frequency use, therefore, it is preferable to use a wire comprised of many fine strands, rather than a single large-diameter wire.

14.6 Conclusion

In the preceding sections we have proceeded from Maxwell's equations to solve certain problems dealing with the propagation of plane electromagnetic waves in unbounded media or in two adjacent semi-infinite media. Presumably

[5] $\sqrt{-j} = \sqrt{\frac{-2j}{2}} = \sqrt{\frac{1-2j-1}{2}} = \sqrt{\frac{(1-j)^2}{2}} = \frac{(1-j)}{\sqrt{2}}$

these examples will have given the reader some feeling for the method of applying Maxwell's equations and for their general usefulness. The examples chosen for discussion are simple enough that the methods of procedure are not masked by mathematical complexities.

The examples leading to results that are applicable to optics were selected for two reasons. First, these examples were important historically in hastening the acceptance of Maxwell's equations. Second, it was felt that the reader would be interested to discover that certain results known to him in optics are the logical consequence of electromagnetic theory.

The important applications of Maxwell's equations to the radiation, transmission (by bounded media), and reception of electromagnetic energy have not been dealt with. The reader interested in these subjects (e.g., antennas, transmission lines, and wave guides) will have the opportunity of studying them in advanced courses or in advanced texts. There he will discover, among other things, that electromagnetic waves are not always transverse, that electromagnetic energy is not always transmitted with the phase velocity, and that impedance matching is always an important aspect of the transmission of electromagnetic waves. In the investigation of each of these matters, Maxwell's equations provide a useful, often the only, starting point.

14.7 Problems

E (14.1) Establish the vector identity given in equation [14-7].

E (14.2) Equation [14-11] is alleged to represent a plane harmonic wave moving in the positive Y-direction. Show that it correctly predicts that a point of constant phase advances a distance λ in T seconds. How should equation [14-11] be modified so as to represent a harmonic wave moving in the negative Y-direction?

E (14-3) Establish the vector identity given in equation [14-23].

E (14.4) Show that the quantity $\boldsymbol{E} \times \boldsymbol{H}$ has units of power/unit area.

M (14.5) Consider a current, I, flowing in a cylindrical wire of uniform cross-section and resistivity. Show that the rate of energy flow into unit length of the wire, calculated from equation [14-28], is equal to the rate of electrical power loss within the wire.

E (14.6) Establish equation [14-100].

Appendix

Some Useful Vector Relationships

∇V and $\nabla^2 V$ in Other Systems of Coordinates

We give below the expressions for ∇V and $\nabla^2 V$ in spherical polar and cylindrical coordinates.

Spherical polar:

$$\nabla V = \hat{\mathbf{r}} \frac{\partial V}{\partial r} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial V}{\partial \theta} + \hat{\boldsymbol{\phi}} \frac{1}{r \sin \theta} \frac{\partial V}{\partial \phi} \qquad \text{[A-1]}$$

$$\nabla^2 V = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial V}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial V}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 V}{\partial \phi^2} \qquad \text{[A-2]}$$

Cylindrical:

$$\nabla V = \hat{\mathbf{r}} \frac{\partial V}{\partial r} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial V}{\partial \theta} + \mathbf{k} \frac{\partial V}{\partial z} \qquad \text{[A-3]}$$

$$\nabla^2 V = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial V}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 V}{\partial \theta^2} + \frac{\partial^2 V}{\partial z^2} \qquad \text{[A-4]}$$

The unit vectors appearing in equations [A-1] and [A-3] are defined in Problem 2.1. The expressions themselves are derived in any standard text on vector analysis (see, for example, H. B. Phillips, *Vector Analysis*, Wiley, 1933).

The Divergence Theorem

At several places in the text the divergence theorem has been invoked in order to transform a surface integral into a volume integral. Symbolically the theorem is stated as follows:

$$\int_s A \cdot ds = \int_v \nabla \cdot A \, dv \qquad \text{[A-5]}$$

The left hand side represents the surface integral of the vector A over a *closed* surface, while the right hand side represents the volume integral of the divergence of A throughout the volume enclosed by the surface.

To establish the theorem let us consider the element of volume $dv = dx\,dy\,dz$. Writing out the right hand side of equation [A-5] we obtain

$$\int_v \nabla \cdot A \, dv = \iiint \left(i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \right) \cdot (iA_x + jA_y + kA_z)\, dx\, dy\, dz \qquad \text{[A-6]}$$

$$= \iiint \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) dx\, dy\, dz \qquad \text{[A-7]}$$

$$= \iiint \frac{\partial A_x}{\partial x} dx\, dy\, dz + \iiint \frac{\partial A_y}{\partial y} dx\, dy\, dz + \iiint \frac{\partial A_z}{\partial z} dx\, dy\, dz \qquad \text{[A-8]}$$

$$= \iint (A_x'' - A_x')\, dy\, dz + \iint (A_y'' - A_y')\, dx\, dz + \iint (A_z'' - A_z')\, dx\, dy \qquad \text{[A-9]}$$

where the double-primed quantities refer to values (of A_x, A_y or A_z) at the upper limits of the variable concerned and the single-primed quantities refer to the corresponding lower-limit values.

In looking in detail at the first term on the right hand side of equation [A-9], let us visualize in the volume under consideration a rectangular prism of cross-sectional area $dy\,dz$ stretching, parallel to the X-axis, from one surface of the volume to the other. This rectangular prism cuts out surface elements ds'' and ds' at its two ends. Evidently,

$$A_x'' i \cdot ds'' = A_x'' \, dy\, dz \qquad \text{[A-10]}$$

and

$$A_x' i \cdot ds' = -A_x' \, dy\, dz$$

Thus, the integration of $(A_x'' - A_x')\, dy\, dz$ with respect to y and z has the value

$$\iint (A_x'' - A_x')\, dy\, dz = \int_s A_x i \cdot ds \qquad \text{[A-11]}$$

The second and third terms on the right hand side of equation [A-9] may be similarly evaluated with the result that equation [A-9] becomes

$$\int_v \nabla \cdot A \, dv = \int_s (A_x i + A_y j + A_z k) \cdot ds \qquad \text{[A-12]}$$

$$= \int_s A \cdot ds \qquad \text{Q.E.D} \qquad \text{[A-5]}$$

Stokes' Theorem

Stokes' theorem has been used occasionally in the text to transform a line integral to a surface integral. The symbolic statement of the theorem is

$$\int_s (\nabla \times A) \cdot ds = \oint A \cdot dl \qquad \text{[A-13]}$$

The left hand side represents the surface integral of the curl of the vector A over any surface, while the right hand side represents the line integral of A around the boundary of the surface.

To establish the theorem let us consider a square surface element, $ds = k\, dx\, dy$. We shall assume that the origin of coordinates is located at the center of this element, and that the value of A at the origin is $A = iA_x + jA_y + kA_z$. Then the value of the line integral of A around this element is

$$A \cdot dl = \left(A_x - \frac{1}{2}\frac{\partial A_x}{\partial y} dy\right) dx + \left(A_y + \frac{1}{2}\frac{\partial A_y}{\partial x} dx\right) dy - \left(A_x + \frac{1}{2}\frac{\partial A_x}{\partial y} dy\right) dx$$

$$- \left(A_y - \frac{1}{2}\frac{\partial A_y}{\partial x} dx\right) dy \qquad \text{[A-14]}$$

$$= \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) dx\, dy \qquad \text{[A-15]}$$

Using the expression for $\nabla \times A$ (equation [2-15]) and the value $ds = k\, dx\, dy$, the reader can readily demonstrate that

$$(\nabla \times A) \cdot ds = \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) dx\, dy \qquad \text{[A-16]}$$

Hence, by comparing equations [A-15] and [A-16], we conclude that for the surface element under study

$$(\nabla \times A) \cdot ds = A \cdot dl \qquad \text{[A-17]}$$

Obviously, this relation is independent of the system of coordinates used to describe the surface element. Further, since any surface can be approximated by square elements, and since the line integral about any surface is the sum of the line integrals around its constituent superficial parts (at all interior boundaries the contributions to the line integral cancel one another), we conclude that

$$\int_s (\nabla \times A) \cdot ds = \oint A \cdot dl \qquad \text{Q.E.D.} \qquad \text{[A-18]}$$

Index

abampere, 201, 258
abcoulomb, 201
acceleration of charged particles, 49 ff.
acceptor levels, 151
admittance, 372 f.
Alnico 5, 291
alpha particles, 63
accelerators
  betatron, 250 f.
  cyclotron, 221 ff.
  synchrocyclotron, 223
  Van de Graaff generator, 51 f., 223
American Institute of Physics Handbook, 138, 173, 188
ammeter, 160 f., 321 ff.
  calibration of, 176
  electrodynamometer, 322 ff.
  hot-wire, 327
  moving-coil, 217, 321
  moving-iron, 325
  rectifier type, 325 f.
  resistance of, 160 f.
  thermocouple, 190, 326 f.
  use of in determination of resistance, 162 f.
Ampère, André Marie, 194
ampere, the, 7, 116, 143, 218 f.
  in alternating currents, 346
ampere-meter, the, 257
Ampère's law, 195, 198 ff., 202, 235
  in circuital form, 207, 210
  direction of induced current from, 235
  use in calculating magnetic induction, 208 f.
Ångstrom unit, 53
anisotopic elements, 4
anode, 140
Armco iron, 288
atom
  chemical properties of, 3
  excited state of, 268 f.
  ground state of, 268 f.
  magnetic moment of, 268 f.
atomic mass
  conversion factor from physical to chemical scale, 4
  determination of, 221
  unit of, 4
atomic number, 3
atomic weight, chemical, 4
  relation to gram-equivalent weight, 141
average values, calculation of, 107 f., 346 f., 359
Avogadro's number, 105, 142
Ayrton shunt, 143

back emf, *see* electromotive force, back
ballast lamp, resistance of, 158
band theory of solids, 146 ff.
barium titanate, 75, 111 f.
Barkhausen effect, 286
battery, 130 ff.
  cell connections in, 132 ff.
  matching resistance of to load, 131 ff.
  open-circuit voltage of, 131
  terminal voltage of, 131
beta rays, 227
betatron, 250 f.
Biot and Savart, law of, 206
Bohr, Niels, 3, 269
Bohr angular-momentum principle, *see* Bohr's First Postulate
Bohr magneton, 265
Bohr's First Postulate, 3, 264
Bohr's Second Postulate, 269
Boltzmann, Ludwig, 105
Boltzmann constant, 105
Boltzmann probability factor, 105 f.
boundary conditions
  in current flow, 122 ff.
  in electrostatic problems, 59 ff., 96 ff.
  at insulator-conductor boundary, 123
  at interface between magnetically different media, 273 ff.
  in motion of galvanometer coil, 315 f., 319
  in reflection and refraction of electromagnetic waves, 396
  in transient currents, 332 ff.
branch of a circuit, 135
Brewster's angle, 402
bridge
  AC, 247, 374 ff.
    capacitance, 376
    inductance, 375 f.
  Kelvin double, 168 f.
  slide wire, 167 f.
  Wheatstone, 163 ff., 170
    sensitivity of, 164 ff.
Busch, 227

calorie, 126
calutron, *see* isotope separator
Canton, John, 2
capacitance, 66 ff.
  method of calculation of, 75
  rules for combining, 70 f.
capacitivity, *see* dielectric constant
capacitor, 68 ff.
  in AC circuits, 354 ff.
  charging through a resistor, 333 f.
  cylindrical, 60, 78 f.
  discharging through a resistor, 331 ff.
  miniature, 112
  parallel plate, 25, 27 f., 41, 52, 74 ff., 90 ff.
    with compound dielectric, 75
    with guard ring, 76
    paper, 77
    variable, 77
  spherical, 77 f., 84
  standard, 76, 78
  in transient currents, 331 ff.
  use of, in selection of radio signal, 73 f.
    in smoothing fluctuating voltages, 72 f.
    in transmission of AC signals, 73
carbon, as standard of atomic mass, 4
carnauba wax, 91
cathode, 140
cathode ray oscilloscope, 85, 335
cathode rays, 27
Cavendish, Henry, 5, 21
cavity, electric field and displacement within, 97 f., 103 f.
  magnetic effects inside, 200
cells, electrolytic
    batteries of, 130, 132 ff.
    internal resistance of, 131 ff.
    open-circuit voltage of, 131
    primary, 130
    secondary, 130
    standard, 172 f.
    storage, *see* secondary
    terminal voltages of, 131
    Weston, 172
cgse, *see* electrostatic units
cgsm, *see* electromagnetic units
charge, elementary, *see* elementary charge
  moving, magnetic field of, 194

charge density, 12, 91, 95
induced on grounded sheet by point charge, 55 f., 64
induced on grounded sphere by point charge, 57 f.
on irregularly shaped conductor, 46 f.
of polarization charges, 91 f., 102 f.
choke coil, 360, 383
circuits
AC, parallel, 363 ff., 370
AC, series, 348 ff., 368 ff.
coupled, 238, 248
DC, 134 ff.
quality of, 338, 362 f., 372
tuned, 73, 363
use of complex algebra in solution of, 368 ff.
circuit theorems, 138 ff.
circular mil, 151
Clausius-Mossotti relation, 104 f., 109
coaxial cable, 78 f.
self-inductance of, 244 f.
coefficient of coupling, 248
coercive force
in ferroelectricity, 112
in ferromagnetism, 289, 291
complex algebra, 365 ff.
Kirchhoff's laws expressed in, 373 f.
use of, in parallel AC circuits, 370 ff.
use of, in series AC circuits, 368 ff.
use of, in study of electromagnetic waves, 393 ff.
complex conjugate, 366
conductance
in AC circuits, 372 f.
in DC circuits, 151 f.
conduction band, 149 f., 154
conductivity, 152
relation between electrical and thermal, 176 f., 191
thermal, 176 f., 191
conductor, 86
capacitance of, 67
current-carriers in, 117
distribution of charge on irregularly shaped, 46 f.
electric field inside, 21
electromagnetic waves in, 403 f.
conductor (Cont.)
potential of, 42, 45 f.
resistance of, 120 f., 149
calculation of, 152 f.
variation with temperature, 153 ff.
stress on surface of charged, 24 f.
conservation of charge, 118, 134
conservation of energy, law of, 39 f.
conservative field, 40 f.
continuity, equation of, 118, 123
coordinates, Cartesian, 33 ff.
coordinates, cylindrical
Laplacian operator in, 60, 406
unit vectors in, 61
coordinates, spherical polar
electric field components in, 62
Laplacian operator in, 59, 406
unit vectors in, 61
corona discharge, 47
Coulomb, Charles, 5, 255
Coulomb barrier, 51, 63
coulomb, the, 6 f., 218
Coulomb's law
for electrostatics, 5 ff., 196
in electrostatic units, 6
for magnets, 255 ff.
in electromagnetic units, 257 f.
in mks units, 256 f.
in mks units, 7 f.
Coulomb's Law Committee, 11, 200
counter emf, *see* electromotive force, back
coupled circuits, 238, 248
covalent bond, 87
crystal lattice, 125, 149
defects in, 149, 191
thermal vibrations of, 149
Curie, Pierre, 113, 280
Curie law, 279 ff.
Curie temperature
of ferroelectric materials, 112
of ferromagnetic materials, 283 f.
of paramagnetic materials, 279 ff.
Curie-Weiss law, 281
curl, 36
current, 95, 116 ff.
alternating, 116, 342 ff.
effective value of, 324, 346 ff.
graphical representation of, 344 f.

current, alternating (Cont.)
in parallel circuit, 364, 370 f.
phase of, relative to voltage in series LR circuit, 348 ff.
phase of, relative to voltage in series LCR circuit, 356 ff.
phase of, relative to voltage in series RC circuit, 354 ff.
RMS value of, *see* effective value of
as a rotating vector, 381 f.
Ampèrian, 262
closed, dipole equivalence of, 261 ff.
closed, magnetic effects of, 255, 261 ff., 272
closed, magnetic shell equivalence of, 262 ff.
determination of, by electrolytic deposition, 142 f.
determination of a small, by means of electrometer, 327 ff.
direct, 116 ff.
direction of, 117
direction of induced, 231, 235
displacement, 94 ff., 210
eddy, 248 ff.
refraction of lines of current flow, 122 f.
transient, 331 ff.
in LR circuit, 335 f.
in LCR circuit, 336 ff.
in RC circuit, 331 ff.
current balance, 219, 323
current density, 117 f., 150
behavior at interface between two conductors, 122 f.
cyclic currents, method of, 136 ff.
cyclotron, 221 ff., 251
frequency modulated, 223
cyclotron frequency, 222, 226

Dalton, John, 3
damped oscillation
critically damped, 337
logarithmic decrement, 338
overdamped, 337
underdamped, 337 f.
damping of galvanometer, 307 ff.
critical damping resistance, 309, 311
damping of galvanometer (Cont.)
critically damped, 309 ff., 315
during initial deflection, 314 ff.
electromagnetic, 308 f.
logarithmic decrement, 313
overdamped, 309 ff., 315
underdamped, 310 ff., 315 f.
Davisson, C. J., 54
de Broglie, Louis, 54
de Broglie wave length, 54, 64
Debye, Peter, 109
decibel, 383
del (∇)
in Cartesian coordinates, 36
in cylindrical coordinates, 406
in spherical polar coordinates, 406
demagnetization, 291 f.
curve, 291
deuterium, 3 f.
diamagnetism, 277 f.
dielectric coefficient, 11, 75, 93, 105, 112
determination of, 79
relation to index of refraction, 100 f.
table of, 70
dielectric constant, 11, 93, 111
relation to index of refraction, 100 f.
variation of, with frequency, 79 f., 101
dielectrics, 11, 86 ff.
anisotropic, 91, 93
boundary conditions at interface of two, 96 ff.
isotropic, 91
values of **D** and **E** in cavities in, 97 f.
dielectric strength, 69 f.
table of, 70
diode, 159
dip, angle of, 254
dipole, *see* electric dipole *and* magnetic dipole
dispersion, of light, 101
displacement current, 94 ff., 210
divergence theorem, 37, 118, 407 f.
domain structure
in ferroelectrics, 112
in ferromagnetics, 282 ff.
donor levels, 151
Dufay, 2

Earnshaw's theorem, 61
eddy currents, 248 ff.
electret, 91 f.
electric charge
  density of, 12
    on irregularly shaped conductors, 46 f.
    at sharp points, 47
  elementary, 8, 27 ff.
  induced, 16
  test, 8
electric dipole, 86 ff., 95 f., 105 ff.
  field of, 88 f.
  induced, 87, 105, 108 f.
  moment of, 87
  permanent, 86 ff., 105 f.
  potential energy associated with orientation of, 105 f., 114
  potential of, 88 f.
  table of values of, 110
electric displacement, 12, 92 ff., 111 f.
  as flux density, 94
  behavior at interface between two dielectrics, 96 f., 99, 122
  independent of polarization charges, 93
  operational definition of, 97 f.
electric field, 8 ff., 38 ff.
  behavior at interface between two conductors, 122 f.
  behavior at interface between two dielectrics, 96 f., 99, 122
  between oppositely charged conducting sheets, 15 f.
  close to a charged conductor, 23 f.
  of electromagnetic wave, 388 f.
  external to spherical charge distribution, 20 f.
  the inner, 101 ff.
  inside a charged conductor, 21
  of large charged sheet, 14 f., 22 f.
  of long charged wire, 12 ff., 21 f.
  normal to equipotential surfaces, 42
  operational definition of, 97 f.
  of a point charge, 9 ff.
    dependence upon medium, 11 f.
  relation to electric potential, 38
  units of, 41
electric flux, 94
electric flux density, 94
electric intensity, *see* electric field
electric monopole, 90
electric motor, 217
electric polarization, 90 ff., 104 f., 111 f.
electric potential, 33 ff.
  of conductors, 42, 45
  of continuous charge distribution, 44
  of cylindrically symmetric charge distribution, 45 f.
  difference in, 38, 46, 49
    determination of absolute value of, 76
    relation to electromotive force, 128 f.
  gradient of, 40 ff.
  independent of path of integration, 39 f.
  inside conductors, 45
  of point charge, 43 f.
  relation to electric field, 38 ff.
  as single-valued function of position, 39 f.
  of spherically symmetric charge distribution, 45 f.
  of system of point charges, 43 f.
  units of, 39
  zero of, 38, 46
electric quadrupole moment, 89 f.
electric susceptibility, 90 f.
electrical image, of point charge near grounded conducting sheet, 55 f.
electrical image, of point charge near grounded conducting sphere, 57 f.
electrical images, method of, 54 ff., 64
electrical stress, on surface of charged conductor, 24 f.
electrochemical equivalent, 141
electrodynamometer action, 218, 322 ff.
electrolysis, 141
electrolytes, 117 f.
electromagnetic induction, 230 ff.
  Faraday's law of, *see* Faraday's law of electromagnetic induction
electromagnetic theory of light, 37, 96, 100 f., 202, 387 ff.
electromagnetic units, 7, 201 ff., 257 f.
electromagnetic waves, 386 ff.
  in conducting medium, 403 f.
  energy transported by, 391 f.

electromagnetic waves (Cont.)
exponential representation of, 393 ff.
plane, in free space, 387 ff.
plane-polarized, 391
reflection of, 395 ff.
refraction of, 395 ff.
total reflection of, 401 f.
velocity of, 100, 202, 389, 395, 405
electrometers, 327 ff.
limit to value of voltage-dropping resistor, 329 f.
quadrant, 328
vacuum-tube, 328 f.
vibrating-reed, 329
electromotive force, 1, 22, 127 ff.
alternating, effective or RMS value of, 347
back, 238, 240, 289, 309, 350 f.
electrolytic cell, 130 ff., 319
Hall, 224 f.
induced, 130, 231 ff., 342 f.
direction of, 231, 235
in galvanometer damping, 308 f.
magnitude of, 231 f., 235
junction, 182
thermal, 130, 184 ff.
method of tabulating, 187 f.
source of, 192 f.
table of, 188
variation of, with temperature, 185
Van de Graaff generator, 129 f.
versus difference in potential, 128 f.
electron, 3, 27 ff., 119, 250 ff.
conduction, 119 f., 149 ff., 192, 224 f.
in heat flow, 177
motion of, under applied field, 119 f.
wave belonging to, 153, 157, 191
holes, 150 f., 225
mass of, 3, 51
orbit of, 3, 30, 54, 64, 278
orbital angular momentum of, 147
magnetic moment associated with, 264 ff.
specific charge of, 27
spin of, 147
magnetic moment associated with, 147, 266, 284 f.
valence, 146, 148 f.
electrostatic field, 8 ff., 40 f.
curl of, 41
as gradient of potential, 40 f.
line integral of, 39 f., 97, 122, 128, 207
electrostatic lenses, *see* lenses, electrostatic
electrostatic units, 6 f., 202
element, 3
elementary charge, 8, 27 ff., 116
energy
absorption and liberation of, in thermoelectricity, 181 ff.
electric
of a charged conductor, 67 f.
density of, 25 f., 93 f.
dissipated in current flow, 125 ff.
units of, 126
electromagnetic
density of, 242
radiation or absorption of, 3, 251, 269, 405
transmission of, 244, 391 f., 405
heat, produced during current flow, 125 ff., 249 f., 346 f., 359 f.
kinetic, relativistic expression for, 50, 63
loss in hysteresis cycle, 289 f.
magnetic, 242 ff.
in current-carrying coil, 240 f.
density of, 242
potential, 38
associated with orientation of electric dipole, 105 f., 114
associated with orientation of magnetic dipole, 267 ff.
of gas molecules, 105
solar, utilization of, 192
source of, in deflecting electrostatic voltmeter, 81 f.
energy bands in a solid, 146 ff.
energy-level diagram, 148, 269
energy product, maximum, 291
equipotential surfaces, 42, 46 ff., 52 f.
experimental plotting of, 49, 124
relation to lines of force, 47 ff.
exact differential, 40

farad, 67
Faraday, Michael, 10, 67, 141, 215, 230 ff.

faraday, the, 141 f.
Faraday disc generator, 253
Faraday's laws of electrolysis, 140 f.
Faraday's law of electromagnetic induction, 232 ff., 248, 250, 343
  derived from Ampère's law, 233 f.
  sign convention in, 232
Fermi-Dirac distribution, 146 ff.
Fermi energy, 148
Fermi gas, 147 f.
fermions, 147
ferrites, 290 f.
ferroelectricity, 92, 111 f.
ferromagnetism, 278, 281 ff.
  Weiss' theory of, 281 ff.
Fleming's generator rule, 235
Fleming's left-hand rule, 216, 235
flux, *see* electric flux *and* magnetic flux
flux density, *see* electric displacement *and* magnetic induction
fluxmeter, 319 ff.
force, central, 265
  electric
    between point charges, 5 ff.
    on dielectric in electric field, 82 f.
    inverse square law of, 5 ff., 21
    on uncharged conductor in electric field, 26 f.
  exchange, 285
  Lorentz, 197
  magnetic
    between electric currents, 194 ff., 217 ff.
    on charge in motion, 196 f., 215, 220 ff.
    on current, 194 ff., 215 ff.
    on dipole in nonuniform magnetic field, 270 f.
Franklin, Benjamin, 2, 5
Fröhlich, H., 104, 111

galvanometer, 168, 302 ff.
  ballistic, 293, 316 ff., 333
    calibration of, 295, 318 f.
    as a fluxmeter, 319 ff.
    motion of coil of, 316 ff.
    sensitivity of, 293, 318
  d'Arsonval, 217, 302 ff.
    ballistic sensitivity of, 293, 318
galvanometer, d'Arsonal (Cont.)
    critical damping resistance, 309, 311
    current sensitivity of, 304 f., 311
    damping, 307 ff.
    megger, 306 f.
    megohm sensitivity of, 304 ff.
    relation of current sensitivity to resistance of, 166, 305
    voltage sensitivity of, 304 f., 311
  moving coil, *see* galvanometer, d'Arsonval
  tangent, 330
gas molecules, variation of concentration with height, 105 f.
gas constant, 147
Gauss, Karl Friedrich, 16
gauss, the, 201, 257
Gauss' theorem
  for electrostatics, 16 ff., 208
    applications of, 20 ff., 97
    for irregular surfaces, 17 ff.
    in terms of electric displacement, 93, 272
    in vector form, 36 f.
  for magnetism, 272
Gaussian surface, 20, 22 f.
Geiger counter, 63
generators, 130
  AC, 253, 342 f.
Gerlach, 271
Gilbert, William, 1, 254
gradient, 36
gram equivalent weight, 141
gravitational potential, 38
Grover, F. W., 247
guard ring, for capacitor, 76
Gutman, F., 92

Hall, Edwin, 224
Hall coefficient, 225
Hall effect, 224 f.
Harnwell, G. P., 138
Harris, F. K., 173
Heisenberg, 281
Helmholtz coils, 203 f., 228
Henry, Joseph, 230
henry, the, 238
Hertz, Heinrich Rudolph, 340, 389 f.

hole, electron, 150, 225
horsepower, 127
hydrodynamics, divergence and curl in, 36
hydrogen, atomic mass, 4
hydrogen, heavy, *see* deuterium
hydrogen, isotopes, 3 f.
hysteresis
  in ferroelectric materials, 112
  in ferromagnetic materials, 289 ff.
    experimental tracing of loop, 292 ff.
    significance of area of hysteresis loop, 289 f.

images, method of, *see* electrical images
impedance, 349
  characteristic, 392 f.
  in complex form, 368 ff.
  diagram of, 349 f., 355, 357
  in parallel resonant circuit, 371 f.
  rules for combining complex, 369, 373
  in series LR circuit, 349
  in series LCR circuit, 357
  in series RC circuit, 355
  in series resonant circuit, 360 f.
index of refraction, 100 f., 397, 400 ff.
  table of values of, 101
inductance
  calculation of mutual and self, 243 ff.
    by flux linkage, 245 f.
    using Neumann's formula, 246 f.
    using notion of magnetic energy, 243 ff.
  determination of, 247, 375
  mutual, 237 ff.
    of coaxial toroidal solenoids, 238
    coefficient of coupling, 248
  rules for combining, 239, 385
  self, 238 ff.
  standard, 247
induction heating, 249
inductometer, 240
inductors
  in AC circuits, 348 ff., 356 ff., 360
  quality of, 363
  in transient currents, 335 ff., 353 f.
impurity atoms, 149 ff.
incandescent lamp, resistance of, 156 ff.
insulators, 86
  band theory of, 148 f.
insulators (Cont.)
  dependence of resistance upon temperature, 154
intermediate metals, law of, 186 ff.
intermediate temperatures, law of, 184 f.
International Electrotechnical Commission, 7 f.
International Union of Pure and Applied Physics, 11, 67
ionization, 5
ions, 117, 130, 140 f., 220, 222 ff.
isotope, 3
isotope separator, 229

Jeans, James, 58
Joule, James Prescott, 126
joule, the, 7, 126
joule heat, 125 f., 180, 183
junction point, 135

Kelvin, Lord, 54, 76, 168, 182
Kerst, D. W., 251
kilowatt-hour, 127
Kirchhoff, Gustav Robert, 134
Kirchhoff's first law
  in complex form, 373 f.
  for DC circuits, 134 ff.
  for transient currents, 332
Kirchhoff's second law
  for AC circuits, 348, 351 f.
  in complex form, 373 f.
  for DC circuits, 129, 134 ff., 160, 162
  for transient currents, 332 ff., 353 f.
Kohlrausch, 202
KVA rating, 360

Langevin, Pierre, 109, 279
Langevin function, 282 f.
Laplace, Pierre Simon de, 21
Laplace's equation, 42 f.
  for current flow, 123 f.
  use of, in electrostatic problems, 59 ff., 64 f.
  use of, to obtain potential outside charged conducting sphere, 59
  use of, to obtain potential between coaxial conducting cylinders, 60

Laplacian operator
in Cartesian coordinates, 43, 388
in cylindrical coordinates, 60, 406
in spherical polar coordinates, 59, 406
Larmor precession, 278
lattice defects, 149
Lawrence, E. O., 221
lenses
electrostatic, 49, 52 f.
magnetic, 325 ff.
Lenz, H. F. E., 230
Lenz's law, 231 ff.
in diamagnetism, 278
in galvanometer damping, 308 f.
lines of force
electric, 10, 13 ff.
and equipotential surfaces, 47 f.
refraction of, 98 f.
relation to electric flux and displacement, 94
magnetic
of current-carrying wire, 199, 215 f.
refraction of, 274 f.
Lorentz force equation, 197
Lucretius, 254

magnetic circuit, 296 ff.
magnetic coefficient, *see* relative permeability
magnetic dipole
equivalence of a current loop to, 261 f.
field of, 259 f.
orientation in domains in ferromagnetic materials, 282 ff.
potential energy associated with orientation of, 267 ff.
torque experienced by in field of uniform induction, 260 f.
magnetic energy, *see* energy, magnetic
magnetic fields, 196, 198 ff.
magnetic field strength, *see* magnetic intensity
magnetic flux, 199, 231
magnetic flux density, *see* magnetic induction
magnetic induction, 196 ff.
on axis of circular current, 202 f.
behavior at interface between magnetically different media, 273
magnetic induction (Cont.)
at center of circular current, 200 f., 256 ff.
of current element, 198 ff.
line integral of, *see* Ampère's law in circuital form
of long solenoid, 204 f., 208 ff.
operational definition of, 275
of straight wire, 205 f.
of toroidal solenoid, 208 ff.
magnetic intensity, 198, 258 f.
behavior at interface between magnetically different media, 273 f.
in electromagnetic wave, 390 ff.
the "inner field," 279 f.
line integral of, 207, 210, 273 f.
as magnetizing force, 276
operational definition of, 275
magnetic moment, 147, 259
associated with particle spin, 266, 268
of atom, 268 f., 271
of closed current, 261 f.
of dipole, 259
of magnetic shell, 263
of molecules of paramagnetic material, 279
of orbital electron, 264 f., 268
magnetic polarization, 275 ff.
spontaneous, in ferromagnetic materials, 283 ff.
magnetic potential
scalar, 208, 211 f., 259 f., 263
vector, 208, 211 ff., 273
magnetic properties of matter, 254 ff.
magnetic shell, 262 ff.
indifference of shape, 263 f.
strength of, 263
magnetic susceptibility, 275 ff.
magnetism, history of, 254 f.
magnetization, *see* magnetic polarization
magnetization curve, 287 ff.
experimental tracing of, 295
magnetomotive force, 296 ff.
magnetostatics, 194, 198, 255 ff.
magnetostriction, 286
magnets
bar, 256
electro-, 222, 298 f.
permanent, 291 f., 321

Malus, law of, 184
manganin, 155, 170 f.
Marconi, 340
mass number, 4
mass spectrograph, 221
mass spectrometer, 221
mass spectroscope, 219 ff.
Maxwell, James Clerk, 21, 37, 95 f., 202, 386 ff.
Maxwell's equation, first, 37, 93, 386
Maxwell's equation, fourth, 272 f., 386
Maxwell's equation, second, 210 f., 386
Maxwell's equation, third, 236, 386
Maxwell-Boltzmann distribution function, 119, 147
McMillan, E. M., 223
megger, 306 f.
mho, the, 151
mica, 75
microscope
  electron, 53 f.
  optical, 53
Millikan, R. A., 27, 29
mks units, 6 ff., 39, 41, 67, 94, 116, 120 f., 126 f., 151, 199 f., 218 f., 232, 238, 256 f.
modulus of a complex number, 367
molecular field coefficient, 280
molecular polarizability, 104 ff.
  effect of temperature on, 105 ff.
  and molecular structure, 109 f.
molecular structure, from molecular polarizability, 109 ff.
moment of force, *see* torque
motor, 217
motor rule, *see* Fleming's left-hand rule
multiplier, 161, 322

Neumann's formula, 246 f.
neutron, 3
  magnetic moment of, 266
  spin of, 147
Newton, Isaac, 6
Newton's law, second, 220, 250, 307, 317
Newton's law, third, 195, 257
newton, the, 7 f.
nichrome, 155
nonlinear resistances, *see* nonohmic materials
nonohmic materials, 121 f., 157 ff.
  bilateral nonlinearity, 157 f.
  unilateral nonlinearity, 159 f.
nuclear forces, 51
nucleon, 4
nucleus of an atom, 3, 51, 63

Oersted, Hans Christian, 194, 199, 230
oersted, the, 257 f., 301
Ohm, Georg Simon, 121
ohm, the, 121
Ohm's law, 121, 124 f., 403
  for heat flow, 176
ohms per volt, 161, 322
oil drop experiment, 27 ff.
Onnes, Kamerlingh, 156
oscillator
  relaxation, 334 f., 341
  Tesla, 339 f.
  underdamped LCR, 336 ff.
oxygen, as standard of atomic mass, 4

paramagnetism, 277 ff.
  variation of, with temperature, 279
Pauli exclusion principle, 147, 284
Peck, E. R., 213
Peltier, Jean, 181
Peltier coefficient, 182
Peltier effect, 180 ff.
periodic table, 3
permalloy, 288
permeability, 100, 198
  in electromagnetic units, 201, 258
  of ferromagnetic materials, 274, 288
  of free space, 200
  initial, 288
  maximum, 288
  relative, *see* relative permeability
permittivity
  of a dielectric, *see* dielectric constant
  of free space, 8, 93
phanotron, 159
phase angle, 345
  between current and voltage in serie LR circuit, 348 ff.

phase angle (Cont.)
between current and voltage in series LCR circuit, 356 ff.
between current and voltage in series RC circuit, 354 ff.
initial, 344
relative, 345
phase velocity, 388, 395, 405
Phillips, H. B., 406
photoelectric effect, 119
piezoelectricity, 113
pinch effect, 218
Planck's constant, 3, 54, 147
platinum resistance thermometer, *see* thermometer, resistance
points, electrical discharge from, 47
Poisson's equation, 42 f.
polar molecules, 12, 26, 79, 86 f., 91 f., 101
polarization, of dielectric, 12
see also electric polarization
poles, magnetic, 6, 254 ff.
Coulomb's law for, 255 ff.
demagnetizing effect of, 291 f.
in electromagnetic system, 257 f.
in mks system, 256 f.
potassium dihydrogen phosphate, 111 f.
potentiometer, 171 ff.
applications of a, 175 f.
direct reading, 173 ff.
standardization of, 174
power, electrical, 126, 359 f.
conditions for maximum transfer of
in AC circuits, 376 f., 380 f.
in DC circuits, 131 ff., 140
loss of
in AC circuits, 126, 359 f., 369 f.
in DC circuits, 126 f.
wattless, 360
power factor, 359
Poynting's vector, 391 f.
practical units, 6
prefixes, representing powers of 10, 67
probability, *see* Boltzmann probability factor
propagation vector, 394
proton, 3, 223
magnetic moment of, 266
proton (Cont.)
spin of, 147
pyroelectricity, 113

quality, of a circuit, 338, 362 f., 372
quantum theory, 3, 281
quartz, 93, 113

rationalization, 366
reactance, 349
capacitative, 355
graphical representation of, 361
inductive, 349
reciprocity theorem, 138
rectifier
metallic, 72, 159, 325 f.
tungar, 159
vacuum tube, 159
relative capacitivity, 11
relative permeability, 200, 288
relativistic increase in mass, 50 f.
relativity, theory of, 50 f.
reluctance, 297 ff.
rules for calculating total, 298
remanence, 289, 291
resinous electricity, 2
resistance, 120 ff., 146 ff.
calculation of, 151 ff.
cause of in a conductor, 149
critical damping, 309, 311
determination of by voltmeter-ammeter method, 162 f.
determination of by Wheatstone bridge, 163 ff.
determination of low, 168 f.
determination of very high, 168, 319, 333, 340
grid, 328
non-ohmic, 121, 157 ff.
of ohmic materials, 121 f., 151
potentiometric determination of, 175
rules for combining, 124 f.
standards of, 170 f.
temperature coefficient of, 154
variation of with frequency, 404
variation of with temperature, 153 ff.

resistivity, 120 f., 151
  table of values of, 152
resonance, in AC circuits
  antiresonance, *see* resonance, parallel
  frequency of, 361, 365
  parallel, 365, 371 f., 373
  potential differences at, 361 f.
  series, 360 f., 373
  sharpness of, 362 f., 372
right-hand rule, for determining direction of induction, 199
Rochelle salt, 111 ff.
Rowland ring, 292 ff.
Rutherford, Ernest, 51, 63

Scott, W. T., 138, 215
Seebeck, Thomas Johann, 180
Seebeck effect, 180 f.
  relation to Peltier and Thomson effects, 184
Seignette, 111
Seignette electricity, *see* ferroelectricity
semiconductors, 86, 150 ff.
  ferromagnetic, 290 f.
  as thermoelectric generators, 191 f.
  variation of resistance with temperature, 151, 154
Siemens, William, 169
shielding
  electromagnetic, 249 f., 404
  magnetic, 274 f., 321
simple harmonic motion, 307, 312, 343 ff., 381 f.
skin depth, 404
skin effect, 404
Smyth, C. P., 111
solenoid
  long, 204 f., 208 ff.
    focusing action of, 225 f.
    inductance of, 243, 245 f.
  short, focusing action of, 226
  toroidal, 208 ff., 238, 241 f., 245, 289 f., 292, 294
solid angle, 18
space quantization, 268, 271, 281
specific heat, of a monatomic gas, 147
statampere, 202
statcoulomb, 6 f., 202
steradian, 18
Stern, 271
Stern-Gerlach experiment, 271
Stokes, G. G., 28
Stokes' law, 28 f.
Stokes' theorem, 210, 236, 246, 408 f.
Stratton, J. A., 402
stress, *see* electrical stress
striking potential, 159
superconductivity, 156
supermalloy, 288
superposition theorem, 138
susceptance, 372 f.
susceptibility, *see* electric susceptibility *and* magnetic susceptibility
synchrocyclotron, 223, 251

temperature
  critical, for superconductivity, 156
  effect of, on energy distribution of fermions, 148
  effect of, on energy distribution of gas molecules, 147
  effect of, on molecular polarizability, 105 ff.
  effect of, on susceptibility of paramagnetic materials, 279 ff.
  neutral, 185
  reference, 185
  variation of resistance with, 153
temperature coefficient of resistance, 154 f.
  table of values of, 155
  of thermistors, 178
Temperature Scale, International, 170, 189
tensor, 91, 277
terminal velocity, 28
Tesla, Nicolai, 339
test charge, 8 f., 38 f.
Thales, 1
thermal equilibrium, 105 ff.
thermal emf, *see* electromotive force, thermal
thermionic emission, 119
thermistors, 158, 178

thermal conductivity, *see* conductivity, thermal
thermocouple, 180 f., 184 ff.
thermocouple meters, 190, 326
thermoelectric effect, *see* Peltier effect, Seebeck effect, *and* Thomson effect
thermoelectric generators, 191 f.
thermoelectric power, 186, 191
thermoelectric refrigerators, 192
thermoelectric series, 181
thermoelectricity, 180 ff.
thermometer
  resistance, 155, 169 f.
  thermocouple, 188 f.
thermopile, 190
Thévenin's theorem, 138 ff.
Thomson coefficient, 183
Thomson effect, 180, 182 ff.
Thomson, W., *see* Kelvin, Lord
Thomson, G. P., 54
Thomson, J. J., 27
threshold energy, 51
thyrite, 156, 158
time constant
  in LR circuit, 336
  in RC circuit, 332 f.
titanium dioxide, 75
torque, 35
  in electrodynamometer action, 323
  in electrostatic voltmeter, 80 ff.
  experienced by current loop in field of uniform induction, 261 f.
  experienced by magnetic dipole in field of uniform induction, 260 f., 270
  on galvanometer coil, 303, 307 ff.
  in moving-iron meters, 325
torsion balance, 5, 255
transducer, 113
transformer, 248 f., 377 ff.
  auto-, 380
  current transformation in, 378 f.
  energy loss in, as result of eddy currents, 249, 291, 380
  energy loss in, as result of hysteresis, 290, 380
  impedance matching with, 380 f.
  voltage transformation in, 378 f.
triboelectric series, 2
tritium, 4
tuned circuit, 73, 363

ultrasonics, 113
units, *see* electromagnetic units, electrostatic units, mks units, *and* practical units

valence shell, 146
Van de Graaf generator, 51, 129 f., 223
vector
  complex quantity as a, 366 ff.
  curl of a, 36
  divergence of a, 36
  infinitesimal, 43
  multiplication, 34 f.
  Poynting's, 391 f.
  propagation, 394
  unit, 33, 61
vector analysis, 33 ff.
Veksler, V., 223
Vigoureux, P., 166, 219
viscosity of air, 29 f.
vitreous electricity, 2
volt, the, 39, 128
Volta, Alessandro, 84, 100, 130
voltage, alternating
  in complex form, 370 f.
  as a rotating vector, 381 f.
  in series LR circuit, 353
  in series LCR circuit, 357
  in series RC circuit, 356
  of simple AC generator, 253, 342 ff.
voltage diagram
  in series LR circuit, 353
  in series LCR circuit, 357
  in series RC circuit, 356
voltage regulator tube, 159 f.
voltameter, 142 f.
voltmeter, 142, 161 f.
  calibration of, 175
  electrodynamometer, 322 ff.
  electrostatic, 80 ff., 131, 161
  moving-coil, 217, 321 f.
  moving-iron, 325
  ohms-per-volt rating, 161, 322
  rectifier type, 325 f.

voltmeter (Cont.)
resistance of, 161 f.
thermocouple, 326 f.
use of, in determination of resistance, 162 f.
vacuum tube, 80, 161 f.

Walton, E. T. S., 251
Watt, James, 127
watt, the, 127
wattmeter, 324
watt-second, 127
wave equation, 387 ff.
Webb, C. E., 166, 219
Weber, Wilhelm Eduard, 202, 232
weber, the, 199, 232
weber per m^2, the, 199
Weiss, 280
Wheatstone, Charles, 163
Wideröe, R., 251
Wiedemann-Franz law, 177, 191
Wien velocity filter, 228
work, 25 f., 34, 38 f., 49, 68, 82, 125, 267, 290

x-rays, 29, 53, 251

Zeeman, Pieter, 268
Zeeman effect, 268 f.