Modern Physics

A Second Course in College Physics

BY

G. E. M. JAUNCEY, D.Sc.
Late Professor of Physics, Washington University, St. Louis
Formerly Associate Editor of the Physical Review

THIRD EDITION

1948
D. VAN NOSTRAND COMPANY, INC.
TORONTO NEW YORK LONDON

NEW YORK
D. Van Nostrand Company, Inc., 250 Fourth Avenue, New York 3
TORONTO
D. Van Nostrand Company (Canada), Ltd., 228 Bloor Street, Toronto
LONDON
Macmillan & Company, Ltd., St. Martin's Street, London, W.C. 2

First Published, August 1932
Reprinted June 1933;
February 1935; October 1936

Second Edition, June 1937
Reprinted July 1940; June 1942;
August 1943; May 1944;
January 1946; October 1946

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE TO THE THIRD EDITION

There has been a tremendous upsurge of interest in physics in the past few years. Atomic energy and electronics—to name two of the fields of interest—have captured the popular imagination. Consequently, many more students are taking courses in physics than formerly. The desire of the young student to learn something of the wonders of modern physics has been recognized by other writers, with the result that chapters on modern physics are being included in recent textbooks on college physics, usually at the expense of classical physics. The student must have a foundation of classical physics before he can attempt modern physics, because he cannot appreciate the meaning of the accelerating action of the cyclotron, for example, until he knows the meaning of acceleration.

For the foregoing reasons a course in modern physics to follow the usual course in college physics has been developed at Washington University. A textbook for this course was published in its first edition in 1932 and the second edition of *Modern Physics* appeared in 1937. Startling advances have been made in physics since then, and they are described in this third edition.

The present edition can be covered in a two-semester course at the rate of three hours per week for a total of thirty weeks. In order to keep to this limitation, it has been necessary to omit certain topics which appeared in the second edition so as to make room for more intriguing material in the new edition. In the selection of new subjects due consideration has been given to the relevancy and interest of these subjects in the near future as viewed from 1947.

The level of difficulty of this third edition is intermediate between that of the first-year course in college physics and the advanced undergraduate courses in physics. The prerequisites are one year of college physics and one year of college mathematics. In the new physics curriculum at Washington University, the course based on the third edition of *Modern Physics* is prerequisite to the junior and senior courses in physics.

The thanks of the author are due to Professors A. L. Hughes and

R. N. Varney for their suggestions concerning various topics in the book; to Professors J. H. Manley and F. N. D. Kurie for aid in preparing the chapters on nuclear physics and fission; to Professor R. D. Sard for his interest in the chapter on cosmic rays; and to Mr. D. C. Miller, recently a major serving with the U. S. Signal Corps, for his constructive criticism of the chapter on electronics and radio. Thanks are also due to Mary Beatley and Denis Withers for secretarial aid.

G. E. M. JAUNCEY

Washington University
St. Louis 5, Missouri
February, 1947

The publishers wish to express sincere appreciation to Professors A. L. Hughes and Robert N. Varney, colleagues of Professor Jauncey, and to Mr. Keith B. Mather, for their thoughtful and careful work in proofreading this volume after the sudden and unexpected death of the author.

CONTENTS

CHAPTER I HISTORICAL INTRODUCTION

CHAPTER XI THE PHOTOELECTRIC EFFECT

CHAPTER XII ELECTRONICS AND RADIO

CHAPTER XIII RELATIVITY—SPECIAL THEORY

CHAPTER XIV X-RAYS

CHAPTER XV SOME APPLICATIONS OF X-RAYS

CHAPTER XVI BOHR THEORY OF SPECTRA

CHAPTER XVII THE QUANTUM THEORY

CHAPTER XVIII CRITICAL POTENTIALS

CHAPTER XIX RADIOACTIVITY AND ISOTOPES

CHAPTER XX NUCLEAR PHYSICS

CHAPTER XXI NUCLEAR FISSION

CHAPTER XXII Cosmic Rays

CHAPTER XXIII Astrophysics

CHAPTER XXIV Relativity—General Theory

ARTICLE PAGE

CHAPTER XXV PHILOSOPHICAL IMPLICATIONS

CHAPTER I

HISTORICAL INTRODUCTION

1.1. Relation of Physics to Science.—Just as it is necessary to know something of the history of one's own country in order to understand fully the political questions of one's own time, so it is necessary to know something of the history of physics in order to understand fully the meaning of modern physics. Moreover, just as certain events of one's own time have no particular significance when viewed by themselves but are seen to be a part of a magnificent whole when viewed in the perspective of history, so we obtain a better insight into the true nature of human knowledge when we view the developments of modern physics as a part of the development of physics as a whole. Further, the study of the history of physics is interesting as a study of the development of human thought. We shall find that certain ideas which we are apt in our pride to consider as strictly modern, one such idea being the atomic hypothesis of Dalton, were know to the Ancients. Again, just as in the world of politics of the present day there is a distinct growth of an international consciousness out of our present national consciousness and as this necessitates a knowledge of the present-day problems and of the history of other nations, so there is a growing need for physicists to know something of the problems and history of other sciences. Just as a knowledge of world affairs will make a man a better citizen of his own country, so will a knowledge of other sciences make a student of physics a better physicist.

There has, in the immediate past, been too much of a tendency to enclose knowledge in compartments. Perhaps, in the development of knowledge, it has been necessary to recognize the division of science into the various sciences, such as physics, chemistry, zoology, mathematics and so on; but at the present time we have the advent of such sciences as biophysics and biochemistry, showing the tendency toward a unification of the sciences. Many of the present-day problems in zoology and medicine, for instance, are fundamentally problems of physics and chemistry. When a unifying principle or set of principles which runs through all the sciences has been discovered, a single text on science should be written. At the present time, however, such a text is not practicable and this book will be restricted to physics; but, in order to aid the process of unifying present-day physics with the rest of knowledge, we shall devote this first chapter to a brief outline of the history of physics. It is hoped that this outline will arouse so much interest in the student that he will read some or all of the reference books listed at the end of this chapter.

As we noted above, it is only in recent times that science has been subdivided. Among the Ancients, there was no clear distinction between mathematics, astronomy and physics. Accordingly, we shall find in the following paragraphs that the history of physics in ancient times is intermingled with that of mathematics and astronomy. Further, an attempt has been made not only to give dates, which are mere numbers and have no particular meaning to students, but also to mention the contemporary events of political history in many cases. The author has attempted to put facts in the history of physics in the larger setting of events of world history. It is for this reason that much which does not pertain directly to physics is contained in this chapter. In this way, it is hoped, the student will realize the connection between this course in physics and other college courses, perhaps even so far removed from physics as those in Latin, Greek and political history.

1.2. The Greeks.—Although the Greeks are principally famous for their work in literature, philosophy and art, they are also famous for their scientific knowledge. In fact, Greek science was studied and taught until the fifteenth century A.D. Most histories of physics begin with Thales of Miletus (624–547 B.C.), who was familiar with the attractive power of rubbed amber and of the loadstone. The modern word "electron" is the Greek word for amber. A substance which acts like rubbed amber is said to be electrified. Magnetic iron ore was found near Magnesia in Asia Minor and the word "magnet" is derived from the name of this locality. The fact that successive pieces of iron can be suspended in the form of a chain from a single piece of loadstone was known to the great Greek philosopher Socrates (470–399 B.C.). However, the Greeks were not acquainted with the repulsive properties of electrified bodies and magnets.

Pythagoras (572–497 B.C.) laid the foundations of Greek geometry. His most famous proposition is the theorem that the sum of the squares on the two sides of a right triangle is equal to the square on the hypotenuse. Since modern trigonometry depends on this property of a right triangle, it is seen how much we moderns just on this account owe to the Greeks. About 530 B.C., Pythagoras organized a group of intellectuals at Crotona in Southern Italy. This school maintained its existence for about a century and a half after the death of Pythagoras. The most important contribution of this school to physics is the idea that the relations between physical quantities are best expressed in numbers. This idea is at the basis of modern mathematical physics. The Pythagorean cosmogony supposed the universe to be spherical with the earth at the center and the sun, stars and planets moving in circles about the earth as center. This hypothesis is known as the geocentric hypothesis as distinguished from the heliocentric hypothesis which supposes the earth and planets to revolve around the sun. However, the Pythagoreans also realized the possibility of the earth rotating on its own axis, thus explaining the phenomenon of day and night.

Many of the arguments of the Greeks were based on "mathematico-esthetic" reasoning. The sphere was the most perfect of all figures, and there were such mystic numbers as three, six, and seven. One reason for the mysticism of the number six seems to be that, next to the sphere, the most perfect solid is the cube, which has six faces. The sacredness of the number seven was connected either with the seven heavenly bodies which moved among the fixed stars or with the seven apertures in the head—two ears, two eyes, two nostrils and one mouth. As an example of such mathematico-esthetic reasoning, Aristotle argued in the following manner that the world is perfect: "The bodies of which the world is composed are solids, and therefore have three dimensions. Now, three is the most perfect number,—it is the first of numbers, for of one we do not speak as a number, of two we say both, but three is the first number of which we say all. Moreover, it has a beginning, a middle and an end." Pythagoras had some knowledge of the relation between the musical pitch of a stretched string and its length and tension. However, having speculated on musical intervals, he then extended the idea of harmony to the heavens and stated that the seven planets are the seven strings of the lyre and that this gives us the music of the spheres. This was considered not only beautiful poetry but also good physical theory. We may say of Greek scientific philosophy that it did not distinguish definitely between poetry and scientific theory or fact. However, we, in the fullness of our modern wisdom, should not belittle the intelligence of the Greeks for such lack of discernment.

Some of our modern theories perhaps contain no little of the "music of the spheres." To some people the appeal of the modern relativity theory is its beauty and self-consistency. Only a few years ago, writers on atomic structure were wont to include beautiful diagrams, showing the orbits of electrons in the atom. Speaking after the advent of the New Quantum Theory, one wonders whether these writers were not more entranced with the beauty of their diagrams than they were with the agreement of the experimental evidence with the predictions of the theory. This remark, however, is made in no carping sense, because we ourselves shall make use of such beautiful diagrams when we come to the chapter on atomic structure. Perhaps, the human mind desires a theory which is not only consistent with the facts but which also has some innate beauty of its own. Perhaps, in our present day we innately feel that the circle is the most perfect orbit and that next to the circle the ellipse is the most perfect. Let us, therefore, have some thought of the poetry which exists in all of us when we meet these mystic arguments of the Ancients.

The Greeks knew of seven heavenly bodies which moved among the fixed stars. These bodies were known as wanderers or planets, and consisted of the sun, moon, Mercury, Venus, Mars, Jupiter, and Saturn. Although the geocentric hypothesis was the idea generally held by the Greeks, yet there were not wanting those who supported a hypothesis more like that which we hold today. One of the first to abandon the geocentric hypothesis was Philolaus (about 430 B.C.) who reduced the earth to a planet. Instead, however, of a heliocentric cosmogony, he supposed that the sun, moon, earth, and the other planets revolved about a "central fire." The earth itself was a sphere. The "central fire" was not visible to inhabitants of the earth because the earth's face was always turned away from the "central fire." Anaxagoras (about 450 B.C.) taught that the sun was a hot stone and that the moon shone by reflected light. He correctly explained eclipses of the sun and moon. He was banished from Athens on account of his teachings. Aristarchus (310–230 B.C.) improved on the cosmogony of Philolaus by dispensing with the "central fire." He made the sun the center about which the planets and the earth revolve and thus anticipated by nearly 2000 years the heliocentric cosmogony of Copernicus. This theory, however, gained no headway against Greek thought then current.

Another Greek who anticipated a modern theory by more than 2000 years was Democritus (460–370 B.C.), who stated the hypothesis of empty space and an almost infinite number of indivisible and invisible particles which differ from each other only in form, position and arrangement. It should be remarked here that the atomic theory of Democritus was neither subjected to experimental test nor was it based on any experimental fact. The great deficiency of the Greeks was their lack of experimentation. They believed that knowledge could be extended by argument and discussion alone. To try to verify by experiment a logical deduction seemed to the Greeks to be unnecessary. Either the argument was sound or it was not. If it was sound, there was no need of experimental verification; whereas, if it was not sound, the argument could be refuted by better argument. Hence, although the Greeks were familiar with not a few physical facts and phenomena, they made no attempt at systematic research in physics. Nowadays, we distinguish between a passive observational science, such as astronomy, and an active experimental science, such as physics or chemistry.

No history of physics can be complete without a reference to Aristotle. Socrates is sometimes called the "father of philosophy." Plato (427–347 B.C.), the brilliant pupil of Socrates, was the teacher of Aristotle (384–322 B.C.). Aristotle was born at Stagira in Macedonia. Aristotle shares with Euclid, who developed his geometry at Alexandria about 300 B.C., the distinction of having influenced thought and of having his views taught in the schools for more than twenty centuries. At the age of forty-two, Aristotle was summoned by Philip, King of Macedonia, as tutor for his son Alexander, who afterwards became Alexander the Great. Afterwards, Alexander's empire extended from Greece and Egypt on the west to India on the east, so that this relationship between Aristotle and Alexander has had profound influence upon both the Christian and Mohammedan worlds.

Aristotle's principal fame is in other fields than physics, but nevertheless his prestige was so great that his physical ideas and opinions were accepted until the fifteenth century A.D. In medieval times, to doubt Artistotle's doctrine was considered heresy. Aristotle believed the earth to be spherical, but he believed in a geocentric cosmogony. Aristotle grasped the idea of the parallelogram of forces for the special case of the rectangle; but his views on falling bodies were very far from truth, for he taught that heavy bodies fall faster than do light bodies. This was believed until the time of Galileo. From the apparent change in the altitude of the stars as one travels north and south, Aristotle calculated the circumference of the earth to be 10,000 miles. He under-

Some of our modern theories perhaps contain no little of the "music of the spheres." To some people the appeal of the modern relativity theory is its beauty and self-consistency. Only a few years ago, writers on atomic structure were wont to include beautiful diagrams, showing the orbits of electrons in the atom. Speaking after the advent of the New Quantum Theory, one wonders whether these writers were not more entranced with the beauty of their diagrams than they were with the agreement of the experimental evidence with the predictions of the theory. This remark, however, is made in no carping sense, because we ourselves shall make use of such beautiful diagrams when we come to the chapter on atomic structure. Perhaps, the human mind desires a theory which is not only consistent with the facts but which also has some innate beauty of its own. Perhaps, in our present day we innately feel that the circle is the most perfect orbit and that next to the circle the ellipse is the most perfect. Let us, therefore, have some thought of the poetry which exists in all of us when we meet these mystic arguments of the Ancients.

The Greeks knew of seven heavenly bodies which moved among the fixed stars. These bodies were known as wanderers or planets, and consisted of the sun, moon, Mercury, Venus, Mars, Jupiter, and Saturn. Although the geocentric hypothesis was the idea generally held by the Greeks, yet there were not wanting those who supported a hypothesis more like that which we hold today. One of the first to abandon the geocentric hypothesis was Philolaus (about 430 B.C.) who reduced the earth to a planet. Instead, however, of a heliocentric cosmogony, he supposed that the sun, moon, earth, and the other planets revolved about a "central fire." The earth itself was a sphere. The "central fire" was not visible to inhabitants of the earth because the earth's face was always turned away from the "central fire." Anaxagoras (about 450 B.C.) taught that the sun was a hot stone and that the moon shone by reflected light. He correctly explained eclipses of the sun and moon. He was banished from Athens on account of his teachings. Aristarchus (310–230 B.C.) improved on the cosmogony of Philolaus by dispensing with the "central fire." He made the sun the center about which the planets and the earth revolve and thus anticipated by nearly 2000 years the heliocentric cosmogony of Copernicus. This theory, however, gained no headway against Greek thought then current.

Another Greek who anticipated a modern theory by more than 2000 years was Democritus (460–370 B.C.), who stated the hypothesis of empty space and an almost infinite number of indivisible and invisible particles which differ from each other only in form, position and arrangement. It should be remarked here that the atomic theory of Democritus was neither subjected to experimental test nor was it based on any experimental fact. The great deficiency of the Greeks was their lack of experimentation. They believed that knowledge could be extended by argument and discussion alone. To try to verify by experiment a logical deduction seemed to the Greeks to be unnecessary. Either the argument was sound or it was not. If it was sound, there was no need of experimental verification; whereas, if it was not sound, the argument could be refuted by better argument. Hence, although the Greeks were familiar with not a few physical facts and phenomena, they made no attempt at systematic research in physics. Nowadays, we distinguish between a passive observational science, such as astronomy, and an active experimental science, such as physics or chemistry.

No history of physics can be complete without a reference to Aristotle. Socrates is sometimes called the "father of philosophy." Plato (427–347 B.C.), the brilliant pupil of Socrates, was the teacher of Aristotle (384–322 B.C.). Aristotle was born at Stagira in Macedonia. Aristotle shares with Euclid, who developed his geometry at Alexandria about 300 B.C., the distinction of having influenced thought and of having his views taught in the schools for more than twenty centuries. At the age of forty-two, Aristotle was summoned by Philip, King of Macedonia, as tutor for his son Alexander, who afterwards became Alexander the Great. Afterwards, Alexander's empire extended from Greece and Egypt on the west to India on the east, so that this relationship between Aristotle and Alexander has had profound influence upon both the Christian and Mohammedan worlds.

Aristotle's principal fame is in other fields than physics, but nevertheless his prestige was so great that his physical ideas and opinions were accepted until the fifteenth century A.D. In medieval times, to doubt Artistotle's doctrine was considered heresy. Aristotle believed the earth to be spherical, but he believed in a geocentric cosmogony. Aristotle grasped the idea of the parallelogram of forces for the special case of the rectangle; but his views on falling bodies were very far from truth, for he taught that heavy bodies fall faster than do light bodies. This was believed until the time of Galileo. From the apparent change in the altitude of the stars as one travels north and south, Aristotle calculated the circumference of the earth to be 10,000 miles. He understood the correct explanation of eclipses. He was familiar with the rectilinear propagation of light and knew of the equality of the angles of incidence and reflection. He denied the atomic hypothesis of Democritus and taught that space is continuously filled. Also he taught that nature abhors a vacuum. The authority of Aristotle was accepted to so great a degree by succeeding philosophers that they tended to settle all scientific questions by reference to the writings of Aristotle and not to bother with experimental tests. However, Aristotle himself recognized the importance of experience, that is, of experiment. He taught that the knowledge of natural laws could be obtained only by gaining knowledge of facts. He also taught that knowledge should be systematically organized. Aristotle established a school at Athens, the Lyceum, in the fifty-third year of his age. Alexander on his conquests instructed his hunters to furnish Aristotle with zoological and botanical specimens of the various lands. These specimens were arranged and catalogued by Aristotle and thus began the organization of science. Aristotle is said by some to be the greatest biologist of the Greeks and is said to have established the science of embryology.

The empire of Alexander the Great was divided at his death, Egypt going to Ptolemy, who had been one of the companions of Alexander in his youth. Ptolemy became Pharaoh with the title of Ptolemy I. He seems to have attempted to carry out the idea of a systematic organization of knowledge which Aristotle had taught Alexander. He set up the famous Museum at Alexandria, which had been founded by Alexander the Great. The word "museum" did not then have its present meaning, but it meant a place dedicated to the nine Muses. Thus, it may be said, the first university of the world was established in Alexandria at about 300 B.C. Ptolemy I also established a library in connection with the Museum. At the close of the Ptolemaic dynasty, which lasted until 43 B.C., it is estimated that this library contained 700,000 papyri. Soon after its foundation the Alexandrian Museum became and, for two or three generations, remained a center of learning, which Athens, even at its best, could not rival. However, the language and culture at the Museum were Greek.

Archimedes (287–212 B.C.), who enunciated the principle of hydrostatics which bears his name, was born at Syracuse in Sicily. He spent considerable time at the Alexandrian Museum. Archimedes correctly understood the idea of center of gravity and the theory of the lever. There are many stories told of Archimedes. The most famous story tells that he was asked by King Hieron to test whether a crown which was supposed to be pure gold contained some baser metal. The story says that the solution came to him while he was in a bath. He immediately jumped from the bath and ran home, shouting, "I have found it." There is a tradition that he set fire to Roman ships by means of concave mirrors. This story is probably an exaggeration but it shows that the Greeks were familiar with some of the properties of concave mirrors. The mathematical ability of Archimedes is shown by his proof that the area of a segment of a parabola is four-thirds the area of a triangle which has the same base as the segment and an equal height. This correct result was of course arrived at without the aid of the integral calculus.

Hipparchus (*fl.* 146–126 B.C.) is reputed to have laid the foundations of modern astronomy. He was born at Nicaea, about one hundred miles southeast of the present city of Istanbul. In spite of the general opinion of the Greeks that experimental tests were unnecessary, Hipparchus, no less than any modern scientist, realized the absolute necessity of the facts of the case. The path of the sun among the stars was well known to the Ancients, and the length of the year was found by two methods. In the first method the interval between two successive summer (or winter) solstices is measured. The length of a shadow cast by a given object when the sun is at its zenith varies with the time of the year and is a minimum at the time of the summer solstice. A second method is to measure the interval between the time when the sun is at a certain position among the stars and the next time when the sun is at the same position. The position of the sun was measured at the vernal equinox. The year, as measured by the first method, is called the *tropical* year, and as measured by the second method, the *sidereal* year. Hipparchus found that the tropical year was about fifteen minutes shorter than the sidereal year. The plane of the apparent path of the sun around the earth is known as the ecliptic. Hipparchus showed that the vernal equinox was gradually moving along the circle of the ecliptic at the rate of 36 seconds of arc a year. This fact is known as the precession of the equinoxes. The modern value of the rate of precession is about 50 seconds. It is now known that the reason for the precession of the equinoxes is that the earth spins like a top and its axis precesses. Although Hipparchus had no idea of the reason for this effect, yet he established the fact and measured its value with a good accuracy, considering the instruments at his disposal.

Hipparchus invented what we now call trigonometry and prepared a table of chords from which can be found the sines of angles. He measured the length of the tropical year and found it to be 365 days 5 hours 54 minutes. To obtain this value, Hipparchus used the observations of Aristarchus made one hundred and fifty years previously. This value of Hipparchus is reasonably near to the modern value, which is 5 minutes 14 seconds less. Hipparchus believed in a geocentric cosmogony. He supposed that the sun, moon and planets revolved about the earth as center. An epicycle is formed by a point which is revolving in one circle, whose center is moving with uniform speed around a second circle.

Egypt was captured by the Romans in 43 B.C., when about one half of the 700,000 volumes of the Alexandrian Library were burned. After the conquest of Egypt by the Romans, the interest in science declined, and it was about a century and a half before another great thinker arose at Alexandria. Ptolemy (A.D. 70–147), who is not to be confused with the pharaohs of the same name, is famous as the man whose cosmogony was taught and believed in Western Europe for fifteen centuries. Ptolemy extended the trigonometry of Hipparchus and arrived at the well-known theorem:

$$\sin (A + B) = \sin A \cos B + \cos A \sin B$$

He divided the circle into 360 equal parts corresponding to our degrees and these he subdivided into smaller parts, from whose Latin names we derive the modern "minutes" and "seconds." Ptolemy followed Hipparchus in believing in a geocentric cosmogony with its system of epicycles. This system is frequently referred to as Ptolemy's system. To doubt the teaching of Ptolemy, as to doubt that of Aristotle, was considered heresy in the Middle Ages.

1.3. The Romans.—At about 300 B.C., Greeks had settled in Southern Italy and Sicily and had established a number of city-states of the Greek type. Greece itself at that time also consisted of a number of independent city-states. Except under Alexander the Great, the Greek civilization was never politically united and as a result was militarily no match for the power of Rome, which began to extend in a southeasterly direction after the defeat of Pyrrhus in 275 B.C. and finally came to envelop the whole of Greece and Egypt, Alexandria being sacked by the Romans in 43 B.C.

Whatever may be said for the genius of the Romans in war, governmental organization and law, they appear to have had no particular genius for the advancement of science and mathematics. The Romans were content to collect and translate into Latin the works of their Greek predecessors. To the hard-headed, practical Romans it appeared to be waste of time to philosophize as the Greeks had done. However, Greek slaves were employed as letter-writers and also as tutors to the children of Roman families of the upper class. In fact, a little, but not too much, knowledge of Greek and of Greek philosophy was considered to be the correct thing. This dilettante attitude of the Romans was expressed by Cicero (106–43 B.C.), who boasted that his countrymen were not as these Greeks but restricted the study of mathematics to what was useful and of practical application.

One Roman who was an exception to the general rule was Lucretius (98–55 B.C.). He was a disciple of Democritus and upheld the atomic hypothesis. Like the Greek philosophers who preceded him and like the modern scientist, his purpose was to unite the branches of science into an intelligible whole. He is quoted as follows: "Bodies are partly first beginnings of things, partly those which are formed of a union of first beginnings." This is remarkably like our modern belief that bodies are made up partly of elements and partly of compounds. Lucretius also supposed his "first beginnings of things" (i.e., atoms) to be in motion, which comes very near to our modern idea.

Although, with such exceptions as that of Lucretius, the Romans did nothing to extend the knowledge handed down to them by the Greeks, yet, by their translation of Greek works into Latin, they preserved for us much of the Greek philosophy and science which would otherwise have been lost.

The Roman emperor Constantine (reigned A.D. 312–337) moved the capital of the Roman Empire from Rome to Constantinople and gave official approval to Christianity. Soon after Constantine's death, the empire was split into a western empire with its capital at Rome and an eastern empire with its capital at Constantinople. The final disintegration of the Roman power occurred in the fifth century. During this period of disintegration the power of the Church increased, and the influence of Greek science came to an end with the closing of the University of Athens in A.D. 529. Also, during this period, Hypatia (A.D. 370–415), a brilliant woman mathematician and astronomer, lost her life at the hands of a Christian mob, and the Alexandrian library was again pillaged as it had been previously in the first century B.C. Finally, in the seventh century A.D., the remnants of the Library were fed to the flames by the Mohammedans under Omar.

1.4. The Arabs.—Mohammed was born at the town of Mecca in Arabia in A.D. 570. Until he was forty years of age there was nothing in particular about him to distinguish him from any other Arab. At that age, however, it is said that he began to receive revelations. He claimed to be a new prophet. He claimed that the prophets before him, especially Jesus and Abraham, had been divine teachers, but that he crowned and completed their teaching. At this time the religion at Mecca was a kind of polytheism, and Mecca was a place of pilgrimage and sanctuary. After some years the new religion gained strength, and an attempt was made by the authorities of Mecca to suppress the new movement. Finally, Mohammed was forced to take flight to the town of Medina in A.D. 622. This flight is referred to in Mohammedan writings as the Hegira. He became the ruler of Medina and then set out to bring all of Arabia under his rule. This was accomplished by 632, the year of Mohammed's death. Thereupon, his followers collected the Prophet's revelations and wrote them into the Koran. Mohammed's successors, who were known as Caliphs, extended their rule over the neighboring countries of Syria, Persia and Egypt. It was during this extension of the Arab power that the books of the Alexandrian library were burned by Omar. In the second decade of the eighth century the Arabs were masters of the whole southern shore of the Mediterranean and also of all Spain. Further extension of the Moslem power in Western Europe was prevented by their defeat at the battle of Tours by Charles Martel in 732. At the beginning of the ninth century the Moslem power was so extensive that along the shores of the Mediterranean only France, Italy and Greece remained outside their control. War and conquest were followed by a period of intellectual activity among the Arabs, and the more enlightened Caliphs began to exhibit an especial interest in medicine and astronomy. They also became interested in the works of Aristotle, Euclid and Ptolemy and had them translated into Arabic. We owe our present-day knowledge of much of the work of Ptolemy to these Arabic translations. The Arabs thus have aided in preserving Greek philosophy and science for us. They established universities at Cordova and Seville in Spain. Meanwhile, Constantinople remained for more than a thousand years a Greek city and the capital of the eastern Christian Church. Here also were collected many of the Greek works on philosophy and science.

The Arabs not only translated works from the Greek into Arabic, but they also acquired much of the philosophy and science of the Hindus. From the Hindus, the Arabs obtained the numerals to which we are now accustomed. The numerals are called Arabic because the Arabs introduced their use into Europe. The distinct advantage of Arabic—really Hindu—arithmetic is that the value of each symbol depends upon its position as well as its form. This aid to clear thinking must not be forgotten. The Greeks and the Romans must have been very limited by their cumbersome method. The operations of multiplication and division must have been very complex in Roman numerals. The Arabs made contributions in algebra and trigonometry, and these branches of knowledge were being introduced into Europe by them at about the time of the Norman Conquest of England in 1066. Albatenius (850–929), an Arabian astronomer, found that Ptolemy's value for the tropical year was in error and gave 365 days 5 hours 46 minutes 22 seconds, which differs from the modern value by only 2 minutes 24 seconds. He also corrected the value for the rate of precession of the equinoxes, giving a value of 55 seconds of arc per year. In physics, as distinguished from astronomy, perhaps the only notable contribution of the Arabs was in the division of optics. Alhazen (died about 1038) was born at Basra, some fifty miles from the Persian Gulf. His life, however, was mostly spent in Spain, which was under Arab rule. Ptolemy had stated that the angle of refraction was proportional to the angle of incidence. Alhazen improved this statement by adding that the proportionality is true only for small angles. He discarded the Euclidean notion that vision is caused by the interaction of something given out from the eye with something given out by the luminous, or illuminated, body.

Finally, the Arab power was brought to an end in Europe, when in 1492 the Moors (a branch of the Arabs) were driven out of Spain by Ferdinand and Isabella, who in the same year sponsored the voyage of Columbus. Just as the Arabs had burned the Alexandrian library in the seventh century, so, when the Moors were driven from Spain, the Christian leaders made a bonfire in the squares of Granada of 80,000 Arabic manuscripts, many of them valuable translations of classical authors. The position of the nations had become strangely reversed. The aforetime foes of learning had become its protectors. The legitimate heirs of the Greek tradition had become its implacable enemies.

1.5. The Middle Ages.—In 395 the Roman emperor Theodosius divided the empire between his two sons. Also, there was a division of the Church into two sections—a western and Latin-speaking section and an eastern and Greek-speaking section. These two sections lived in general amity until 1054, when they split and severed relations on the question of adding the word "filioque" (i.e., and from the Son) to the creed. The two sections are now known as the Roman Catholic Church and the Greek Orthodox Church (also called the Greek Catholic Church). The ecclesiastical authority of the western church was from the beginning concentrated in the Bishop of Rome, who is now known as the Pope, while the authority of the eastern church was divided among four bishops. In 476 the Goths under Odoacer invaded Italy and the last western emperor was defeated. With the fall of the military power of the Western Empire, the Pope took over the ancient title of "pontifex maximus," which the emperors had held, and so became the supreme sacrificial priest of the Roman tradition.

The Popes for several centuries looked with disfavor upon Greek culture, Greek views of a spherical earth and of the motions of the planets being considered as almost heretical. Western Europe now passed through that remarkable phase which is sometimes known as the Dark Ages. Roughly speaking, this period lasted for one thousand years, from 500 to 1500. In the latter half of this period Arabic translations of Greek works began to appear in Western Europe. In particular, the works of Aristotle became well known and began to be accepted by the authorities of the Church. There seemed to be a desire to base the various dogmas of the Church upon logic. The logic which seemed best to fit the case was among the Greek writings. Aristotle was studied in the Church schools, and the system of logic which was developed in these schools is known as Scholasticism. At the end of the fifteenth century the works of Aristotle were considered to be hardly less infallible than the Scriptures. In England and France laws were passed making it a crime to teach or write against the great philosopher. Servility of thought and unquestioning respect for authority characterize this period.

However, in spite of the general darkness, there were individuals who in obscurity were willing to question the teachings of Aristotle. Among such individuals we may mention Roger Bacon (1214–1294). Bacon was a Franciscan monk at Oxford, England. He insisted that the secrets of nature could not be learned by argument and discussion and by reference to authority, but that observations must first be made. As a result of his insistence upon the experimental method, Bacon was persecuted as a necromancer and spent twenty-four of the last thirty-seven years of his life in prison.

1.6. The Development of the Experimental Method.—The main feature of the history of physics so far described in this chapter is the absence of experiment. In the sixteenth century the movement known as the Renaissance occurred. Although this word is usually used to refer to the intense activity in art and literature at this time, yet the same period also marks the beginning of a tremendous intellectual activity in science. We have already noted the teachings of Bacon, who lived three centuries before the Renaissance. We shall now speak of Copernicus (1473–1543), who was born in Polish Prussia. He perceived that the great objection to the Ptolemaic system of epicycles was its great complexity and that a great simplification could be brought about by supposing a heliocentric cosmogony. The Copernican theory was not based on any particular evidence which could not also be used to support the Ptolemaic theory, but, as was remarked in Sec. 1.2, the human being desires theories which have some innate beauty of their own, and perhaps beauty in this connection is synonymous with simplicity. The Copernican theory assumed circular orbits and that the earth not only revolves about the sun but also rotates on its own axis, as had previously been suggested by the Pythagoreans. Copernicus realized the heretical nature of his views and withheld his book "De Revolutionibus Orbium Coelestium" *from publication until he was* approaching seventy years of age, when he knew that, whatever the ecclesiastical authorities might wish to do about it, he did not have long to live anyway. It should be remarked at this place that printing had come into vogue by the time of Copernicus. Just when the first printing press was set up in Europe is not known with certainty, but in 1446 books were being printed from type in Haarlem, Holland. In 1477 Caxton had set up his press in Westminster, England.

Although Copernicus himself was not persecuted for his teachings, due to the fact that he died before these teachings were brought to the attention of the ecclesiastical authorities, some of the protagonists of his ideas were. It has been suggested that the principal objection of the Church to the Copernican ideas was that they argued against the importance of the earth in the universe. In the Ptolemaic system, the earth is the center of the universe and therefore of tremendous importance. Man is the ruler of the earth and, in a sense, the Pope is the ruler of man. If the earth is not the center of the universe and is just one of several similar planets, man's and also the Pope's place in the scheme of things become purely incidental. In view of this argument, one must not judge too harshly those who thought that the Copernican views were heretical. As a result of this view a ban was put upon the ideas of Copernicus. However, in spite of this ban, Guido Bruno (1548–1600), after having read "De Revolutionibus" and become a convert to the new ideas, began to tell openly of the Copernican ideas. He became a sort of evangelist of the new doctrine. He visited England during the reign of Queen Elizabeth (reigned 1558–1603), when Francis Bacon (not to be confused with Roger Bacon), Shakespeare, Ben Jonson, Gilbert and others were making that reign one of the most illustrious of English history. He then journeyed to France and Germany. Finally, he came to Venice, where he came into the clutches of the Inquisition. After languishing for six years in prison, he was sentenced to be burned at the stake in Rome. This atrocity occurred a full century after the discovery of America by Columbus and only three and one-half centuries ago from our present day. However, it must be remembered that this was an age of intolerance and that the intolerance was not confined to any one country or faith. Bruno was burned at the stake about eighty years before the Puritans were burning witches at Salem, Massachusetts.

The first astronomical observer of any consequence since the time of Hipparchus was Tycho Brahe (1546–1601). Tycho was born of a noble family in Sweden, which at that time belonged to Denmark. Although he was being educated at the universities of Copenhagen, Leipzig, and Rostock for a career as a statesman, he became interested in astronomy on the side. In 1575 Frederick II of Denmark appointed him Royal Astrologer and gave him sufficient income to enable him to build instruments and carry on observations. In return for this Tycho made astrological calculations for the royal family. In 1597 Tycho was dismissed because the king complained that he was wasting too much time on astronomy and not giving enough time to astrology. In 1599 he moved to Prague and was made court mathematician to Rudolph II of Germany. His observations were made by means of a seven-foot quadrant and without the aid of a telescope. He introduced the modern astronomical coordinates, right ascension and declination. His observations were of the greatest value to his pupil and successor, John Kepler.

Kepler (1571–1630) was unlike Tycho in almost every respect. Tycho was of noble birth, rich, vigorous, and possessed of great experimental skill, while Kepler was poor, sickly, and possessed of a keen mathematical mind but of almost no experimental skill. Tycho was opposed to, while Kepler favored, the Copernican system. The two met in 1600 at Prague and were able to disagree with one another gracefully in their views of the universe. On Tycho's death in 1601, Kepler became his successor as court mathematician at Prague. Kepler first tried the Copernican system, but his mathematical calculations failed to agree with Tycho's experimental values. He then tried the Ptolemaic system of epicycles and some improvement resulted, but still his calculations failed to agree with Tycho's experimental results—in one case the difference was *8 minutes of arc*. Kepler then returned to the Copernican system but improved the theory of Copernicus by trying orbits of other shapes than circular. He tried an oval orbit and finally an ellipse with the sun at a focus. At last his computations agreed with experiment and the discrepancy of 8 minutes of arc was removed. Kepler's three laws are as follows:

(1) The orbit of each planet is an ellipse with the sun at a focus.
(2) The line joining a planet to the sun carves out equal areas in equal times.
(3) The squares of the periods of the planets are proportional to the cubes of their mean distances from the sun.

The arrangement of the planets from the sun outwards is shown in Table 1.1. The planets whose names are in parentheses were not known in Kepler's time. He knew of only six planets. Kepler speculated on the way the planets were arranged. Even at the present day, the arrangement is of interest. Various physicists have in recent years tried to discover some sort of a quantum relationship corresponding to the quantized electron orbits in the Bohr model of the hydrogen atom. The effect of contemporary thought even on a mind like Kepler's is shown by his rule for the arrangement of the planets. He noted that there were six planets corresponding to the sacred number six and that there were five spaces between their orbits. Now there are five and only five regular geometric solids, and it seemed to Kepler that there must be a connection. Hence Kepler's rule: Begin with a sphere whose great circle represents the orbit of Mercury. Describe about this sphere an octahedron (8 faces), and circumscribe about this

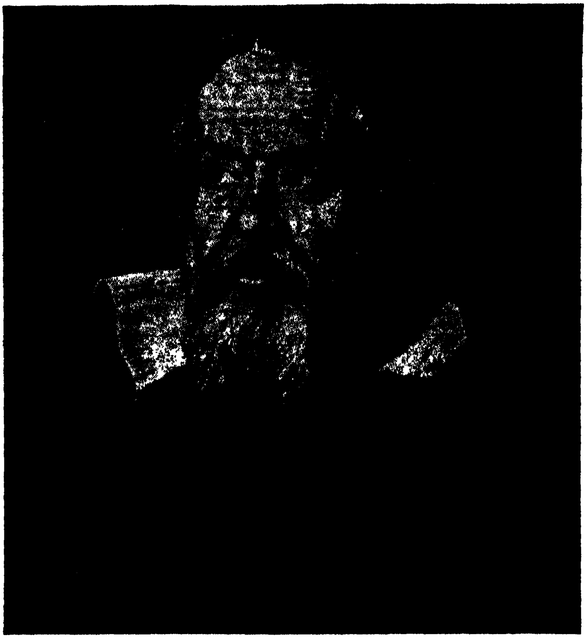
TABLE 1.1

The Solar System

Name	Mean Radius of Orbit mi	Mass (Earth = 1)
Sun	....	330,000
Mercury	3.6×10^7	0.34
Venus	6.7	0.82
Earth	9.3	1.00
Mars	14.2	0.11
(Asteriods)	25.0	
Jupiter	48.3	314.5
Saturn	88.6	94.1
(Uranus)	178.3	14.4
(Neptune)	279.3	16.7
(Pluto)	370.0(?)	

a sphere. A great circle of this sphere represents the orbit of Venus. By a similar use of an icosahedron (20 faces) around the orbit of Venus, we get to the earth; from the earth to Mars by means of a dodecahedron (12 faces); from Mars to Jupiter by a tetrahedron (4 faces); and from Jupiter to Saturn by a cube.* This is an example of the mathematico-esthetic argument, or the argument of perfection, so much used by the Greeks. The human mind seems to delight in beautiful schemes.

Galileo (1564–1642), was born in Pisa, Italy, a few years after the birth of Francis Bacon and a few months before the birth of Shakespeare. His parents intended him to be a physician and sent him to the University of Pisa. However, he became attracted to mathematics, and at the age of twenty-six was appointed professor of mathematics. While a student, he was one day in the cathedral at Pisa when he noticed the swing of a hanging lamp. He timed the swings with his pulse and found that the period was independent of the amplitude. He began experimenting on falling bodies and quickly found that Aristotle's dogmatic assertion that heavy bodies fall faster than light bodies was wrong. In order to convince his colleagues at the University of Pisa of Aristotle's error, Galileo simultaneously let drop a ten pound and a one pound shot from the top of the Leaning Tower of Pisa. Both weights struck the ground simultaneously. This fact did not convince his audience that Aristotle was wrong. Aristotle could not be wrong, and the only effect of Galileo's experiment was in many cases to cause his colleagues to look up Aristotle and find that Galileo was just a foolish young man. He was forced to leave Pisa, and in 1592 he became professor of mathematics at the University of Padua. At Padua he became a supporter of the ideas of Copernicus and corresponded with Kepler. In 1602 he made an air thermometer and in 1609 a telescope. With the telescope he discovered the moons of Jupiter and the mountains of the moon. These discoveries were in direct opposition to Aristotle, who taught that there were seven planets (sun, moon, Mercury, Venus, Mars, Jupiter, Saturn) and that the moon was smooth. In 1610 his fame was so great that he was invited to return to Pisa. By this time, however, the Church had awakened to the heretical nature of Galileo's teachings. In 1615 he was summoned before the Pope and was ordered to relinquish the opinion that the sun was the center of the world and immovable. Later, under the rule of a friendly Pope, Galileo again supported the Copernican theory of a heliocentric universe. In 1632, he published his famous "Dialogues on the Ptolemaic and Copernican Systems." The character in the book who supported the Ptolemaic



Yerkes Observatory

FIG. 1.1.—Galileo Galilei. 1564–1642.

System was made to appear a rather stupid person. Galileo's enemies made it appear to the Pope that this stupid person was a caricature of the Pope himself. Whereupon, the sympathy of the Pope was turned against Galileo. Galileo was brought before the Inquisition and forced to recant. He was imprisoned in his own house. He became blind, but nevertheless he labored on and dictated a treatise on "Motion," in which the way was paved for the later enuciation by Newton of the three laws of motion. During his last years Galileo was attended by Torricelli and was visited by John Milton. Galileo died at the age of seventy-eight, twenty-two years after the voyage of the Mayflower and the landing of the Pilgrims at Plymouth, Massachusetts.

1.7. The Foundations of Contemporary Physics.—The history of physics now merges into the physics taught today. Such names as Newton, Boyle, Young, Fresnel, Faraday, Maxwell and many more immediately recall the experiments and laws which are named for them. In the compass of one chapter it is impossible to tell of all those who have contributed to the advance of physics. We shall therefore restrict ourselves to the biographies of Newton, Rumford, Faraday and Maxwell.

One hundred and fifty years after Columbus made his famous voyage, six years after the foundation of Harvard College in Massachusetts, and in the same year as the death of Galileo, Isaac Newton (1642–1727) was born in the village of Woolsthorpe, Lincolnshire, England. The year of Newton's birth witnessed the beginning of the armed struggle between Charles I and his Parliament. King Charles was beheaded, and Cromwell, backed by Parliament, seized supreme power when Newton was six years old. Newton showed no particular promise of being a genius in his early years at school. In fact, he seemed rather indolent, and it was not until he had received a kick from one of his school-fellows that his dormant genius was aroused. The story goes that Newton responded to the kick in a lusty manner and came off best in the fight. It so happened that the other boy stood above Newton in class and Newton determined to emphasize his victory by also beating the other boy in class. Under this stimulus Newton outstripped all his competitors and finally became top boy in the school when he was fifteen. At this time his mother, who was a widow, needed a man to take care of the farm upon which she lived, and so Newton was removed from school. But Newton had no taste for farming. After he had been discovered under a hedge studying mathematics when he should have been doing farm work, his mother decided that an educational career was more suitable for her son, and he was sent back to school. In 1661, just after the Restoration of the monarchy in the person of Charles II, Newton entered Cambridge University in his nineteenth year, which shows that Newton was no infant prodigy. While he was a student, Newton discovered the binomial theorem, developed the methods of infinite series, and invented fluxions, which is a form of differential calculus. In 1664 he was awarded a scholarship, and in 1665 he obtained the degree of Bachelor of Arts. The Great Plague occurred in 1665 and Cambridge University was closed. Newton returned to the farm in Lincolnshire and stayed there until 1667, when the university was reopened. It was while on the farm that the "falling apple" episode is said to have taken place, and his attention turned to gravity. In 1669 Newton was appointed professor of mathematics at Cambridge. In 1660 the Royal Society was given a

Yerkes Observatory

Fig. 1.2.—Sir Isaac Newton. 1642–1727.

charter by Charles II, thus establishing for the first time in modern history a recognized scientific society. The Royal Society still exists in England and the election to fellowship in this society is still prized as a high honor by all English scientists. Nowadays, there are many such societies all over the world. Newton's famous book, the "Principia," was published in 1687.

From the time of Queen Elizabeth, a coinage based on bimetallism had been in use, both silver and gold coins being used. During the reign of William and Mary (reigned 1688–1702) the silver coinage had become debased by the use of cheap alloys. The situation became so bad that the government (Parliament now had more power) decided to call in all the silver coins and to reissue silver coinage of full value. For this purpose a man who was a chemist and who was incorruptible was needed. The choice fell upon Newton, who was a chemist as well as a physicist and mathematician, and he was offered the Wardenship of the Mint. This he accepted in 1696, and for three years he labored in the nation's cause, when the recoinage was complete. In 1699 Newton was appointed to the high office of Master of the Mint, worth about twelve hundred pounds (about $6,000) per annum. Newton was elected to Parliament on several occasions by Cambridge University. In 1703 Newton was elected President of the Royal Society. Queen Anne (reigned 1702–1714) with her husband, Prince George of Denmark, became interested in the doings of the Royal Society and learned to see in Newton the man of highest genius in her land, so that in 1705 Anne conferred upon Newton the dignity of knighthood. Newton died in his eighty-fifth year and was buried in Westminster Abbey, which is the greatest burial honor which can be bestowed upon an Englishman.

How different the end of Newton's life was from that of Galileo! Newton died with all the honor and respect which an admiring nation could bestow upon him. Galileo died in poverty in a miserable cottage. Disgrace and public condemnation had been heaped upon him. With Newton, science came into its own, and, from his time on, science and scientists have been held in esteem.

Perhaps Newton is most famous for his work on gravitation and optics. He put the science of mechanics on a firm basis with his three laws of motion, which are as follows:

(1) Every body continues in its state of rest, or of uniform motion in a straight line, except in so far as it is compelled by external impressed force to change that state.

(2) The rate of change of momentum of a body is proportional to the impressed force and occurs in the direction of the force.

(3) Action and reaction are equal and opposite.

It was during the Great Plague when Newton was forced back to his Lincolnshire farm that his mind was occupied with the problem of gravitation. Newton first found that if the gravitational force between two bodies varies as the product of the masses of the bodies and inversely as the square of the distance between them, then, for circular orbits of the planets around the sun, Kepler's third law becomes a consequence of this law of gravitation. However, Newton was not satisfied with this agreement of theory with fact, but desired also to correlate the period of the revolution of the moon about the earth with the acceleration of gravity as observed at the surface of the earth. Newton's argument was as follows: The gravitational force on the moon is

$$\frac{GMm}{d^2}$$

where M and m are the masses of the earth and moon respectively, d is the distance of the moon from the center of the earth and G is a constant of proportionality. Newton was familiar with the formula for centrifugal force, which, for the moon about the earth, is

$$\frac{4\pi^2 md}{P^2}$$

where P is the period of the moon's revolution. For the moon to continue in circular motion

$$\frac{GMm}{d^2} = \frac{4\pi^2 md}{P^2}$$

so that

$$G = \frac{4\pi^2 d^3}{MP^2} \tag{1-1}$$

Now a mass μ at the surface of the earth is attracted with a force $GM\mu/a^2$ where a is the radius of the earth. By Newton's second law the acceleration g of the mass μ is given by $g = GM/a^2$. From (1-1), we then have

$$g = \frac{4\pi^2 d^3}{a^2 P^2} \tag{1-2}$$

But astronomical observations give $d = 60\ a$, so that

$$g = 8{,}550{,}000 \times \frac{a}{P^2} \tag{1-3}$$

The value of P is $27\frac{1}{3}$ days, and in 1666 the best value of a, the radius of the earth, was 3820 miles. These values gave $g = 31.0$ ft/sec^2, the experimental value of g being 32.2 ft/sec^2. The agreement between the theoretical and experimental values of g was thus fairly close. However, it seems that Newton was dissatisfied with the argument presented because it is assumed that the mass of a sphere such as the earth may be considered as being concentrated at the center of the sphere. This assumption may be valid when two attracting spheres are at a distance apart which is large compared with the radius of either sphere as is the case for the moon and the earth but may not be valid when a small body at the surface of the earth is attracted to the earth. Because of this lack of rigor in his argument, Newton delayed publication of his inverse square law of gravitation for twenty years. However, in 1685 Newton proved the theorem that, although a small body near a large massive sphere is attracted to each of the particles of matter of which the sphere consists, yet the vector sum of all these forces of attraction is the same as that obtained by supposing the whole mass of the sphere to be concentrated at the center of the sphere. This theorem removed all objections to the above argument. Meanwhile Picard, a Frenchman, had obtained the value 3960 miles for the radius of the earth. *Substituting this value in* (1–3), $g = 32.1$ *ft/sec^2 in excellent agreement* with the experimental value of g. It was only after this agreement of theory with experiment that Newton announced his inverse square law of gravitation. Unlike the Greeks, Newton required not only that a theory must be a beautiful or simple theory, but that its consequences must agree with all the facts. It is not sufficient that the consequences agree with some but not all of the facts.

Newton's work in optics is of particular interest because for nearly a century after the death of Newton, his prestige was so great that further work was retarded by his ideas. A century before Newton it had been considered heresy to question Aristotle, and for about a century following Newton it became not heresy but bad taste to question Newton. It happened, however, that in two instances Newton was wrong. In Newton's time the phenomenon of spherical aberration of lenses was known. Sometime previously Descartes (1596–1650) had shown that lenses with hyperbolic surfaces should be free from spherical aberration. After his experiments on the dispersion of light by means of a prism, Newton was led to announce that what we now call chromatic aberration was a natural property of lenses, and that it could not be eliminated. Newton said that lens telescopes could not be further improved and that the only hope was to develop reflecting telescopes. The development of lens telescopes was thus retarded for many years. Further, Newton supported the corpuscular theory of light as against the wave theory of Huygens (1629–1695) and Hooke (1635–1703), his contemporaries. Newton was familiar with the colors of thin films and certain other interference phenomena, but instead of explaining these by means of a wave theory he endowed his corpuscules with "fits" of easy reflection and easy transmission. Newton's views in this regard were held by physicists until the beginning of the nineteenth century. Even after the interference experiments of Young in 1800 the prestige of Newton was so great that Young was ruthlessly attacked by his scientific contemporaries. It was not until 1850 that the final vindication of the wave theory came with the determination of the velocity of light in water. According to Newton's corpuscular theory, the refractive index of water is v_1/v_2 where v_1 and v_2 are the velocities of light in water and air respectively; whereas, according to the wave theory, the refractive index is v_2/v_1. In 1850 both Foucault and Fizeau found that the velocity of light was less in water than in air, as is required by the wave theory, since the refractive index of water is greater than unity.

At the end of the eighteenth century the prevalent theory of heat was the caloric theory. Heat was supposed to be an imponderable (*i.e.* weightless) and indestructible substance called caloric. The particles of caloric were supposed to be self-repellent and to occupy space. Also the particles of caloric possessed an affinity for matter. The caloric theory explained very nicely why bodies expand on heating. When a body was being heated, the amount of caloric in the body was being increased, and the particles of caloric had to find room between the atoms (Dalton's atomic theory of chemistry was announced in 1811) of the body, so that the atoms were pushed apart and the body expanded.

The foundations of the modern kinetic theory of heat were laid by Benjamin Thompson, better known as Count Rumford (1753–1814). Thompson was born in a humble New England home within two miles of the birthplace of Benjamin Franklin. At the age of nineteen he taught school near Boston. When the War of Independence broke out he took sides with the mother country and fled to England, leaving behind him a wife and daughter. His scientific attainments were such that in 1778 he was elected a Fellow of the Royal Society. He seems to have been of a roving and adventurous disposition, for when, in 1783, there was a possibility of war between Austria and Turkey, he left England to serve with the Austrians. The war, however, did not break out, and he became military engineer to the Bavarian government. In 1790 the Elector of Bavaria honored him with the title of Count Rumford. It was in 1798 that he noticed the large amount of heat which was evolved when cannon were bored. According to the caloric theory, one body could become hotter only if another body became colder. In the case of the cannon, however, all bodies, the gun, the chips and the tool, all became hot. Rumford then set up an experiment in which water was boiled by means of friction. He argued that the heat must have been produced by motion. In 1799 Rumford returned to London, where in 1800 he founded the Royal Institution for the diffusion of the knowledge of applied science. Forty-six years later, by means of funds willed by an Englishman, James Smithson, the Smithsonian Institution was founded in Washington, D. C. for similar purposes. Rumford read a paper on the results of his researches in heat before the Royal Society. As a result, Sir Humphry Davy was influenced to try experiments on the production of heat by friction. In 1812 Davy announced that heat was produced by motion. Thomas Young, who is famous for his interference experiment, espoused the kinetic theory of heat in 1807. However, the great majority of physicists were unconvinced. It was not until 1847, when Joule announced that 778 foot-pounds of work would raise one pound of water one degree Fahrenheit, and when Helmholtz in Berlin announced the principle of the conservation of energy, that the caloric theory of heat was finally abandoned. It is interesting to note that Helmholtz's paper on "Die Erhaltung der Kraft" was rejected for publication in the justly famous *Annalen der Chemie und Physik*. In 1803 Rumford left England and went to France, where he married the widow of Lavoisier, the chemist who discovered oxygen. Lavoisier had been guillotined in 1894 by fanatical French revolutionists on the ground that "the Republic had no need of savants."

Current electricity, as distinct from static electricity, was discovered about 1790 by Galvani and Volta. In 1800 Volta described the first battery. Electrolysis was also discovered in 1800, and in 1802 Sir

Humphry Davy showed that the electrolysis of water yields two parts of hydrogen to one part of oxygen. Davy isolated sodium and potassium by means of electrolysis in 1807. In 1820 Oersted of Copenhagen announced the discovery of the effect of an electric current on a magnet. The importance of Oersted's discovery was immediately recognized by Ampère, who thereupon showed that two currents exert forces upon each other. Ampère also discovered the law of force between the two currents and invented the electromagnet. Just at this time there appeared a man whose researches in electromagnetism have had a profound effect on every-day life. That man was Michael Faraday.

Michael Faraday (1791–1867) was born near London of poor parents. In 1804, in order to assist in providing for the family, he was engaged as errand boy to a bookseller, and, later, was apprenticed to learn the art of bookbinding. He read some of the books and became interested in science. He even performed such of the simpler experiments as he could afford out of his meager wages. He attended four lectures on chemistry by Sir Humphry Davy in 1812. Disliking his trade as a bookbinder, he applied to Davy for a position, however menial, at the Royal Institution. As proof of his earnestness Faraday sent along with his application the notes which he had taken of Davy's four lectures. Davy was so interested that he engaged Faraday in 1813 as apparatus assistant at 25 shillings ($6.25) per week. In 1814–15 Faraday traveled with Davy on the continent and saw many of the continental scientists. Even though he was only an assistant he left an impression because of his modesty and intelligence. On returning to England in 1815, Faraday was reengaged at the Royal Institution at 30 shillings per week. He was encouraged by Davy to begin original research, and in 1816 he published his first paper. In 1817 he published six papers; in 1818, eleven; and in 1819 nineteen papers. This remarkable output was attained in spite of the fact that Faraday had had no university training and was almost entirely self-taught. About 1820 he began his electrical researches. In 1825 he succeeded Davy as Director of the Royal Institution, and, in spite of very lucrative offers elsewhere, he remained at the Insitution for the remainder of his active life.

In 1821, Faraday constructed the first electric motor. Faraday felt that to every experiment in physics there must be a converse experiment. As an example, he felt that there should be a converse of Ampère's experiment where two parallel wires, each carrying an electric current, attract or repel each other according as the currents are flowing in the same direction or opposite directions. The expected converse was that a steady current flowing in a wire would induce a steady current in a neighboring parallel wire, which could be shown by the deflection of a galvanometer to which the two ends of the second wire were connected. Faraday tried the experiment and obtained no result. As a second example, Faraday felt that there should be a converse of Ampère's electromagnet where a current flowing in coils around unmagnetized iron causes the iron to become magnetized. In this case the expected converse was that a permanent magnet placed in a coil of wire whose two ends were connected to a galvanometer would produce a current in the coil. Faraday tried this experiment and as before obtained no result. Perhaps, although there must be a converse to every experiment, the proper converse may not be that which is most obvious, and so Faraday continued at this problem for nine years. Finally, in 1831 he found the proper method of attack. He took a soft iron ring and wound two separate coils A and B upon it. The coil A was connected to a battery and coil B to a galvanometer. Upon starting or stopping the current in

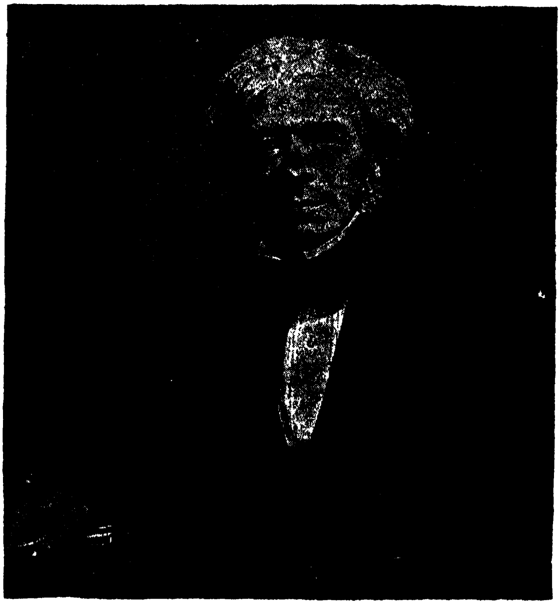


FIG. 1.3.—Michael Faraday. 1791–1867.

coil *A* by means of a switch, Faraday found that the galvanometer showed a transitory current. This is the proper converse in the first example mentioned above. Two months later he found the proper converse of the electromagnet. He inserted the magnet into the coil and found that the galvanometer showed a transitory current during the time of the insertion. A transitory current in the reverse direction was shown upon removal of the magnet from the coil. Following up the consequences of this experiment, Faraday made the first direct-current dynamo by introducing a commutator.

Faraday's practical and concrete mind disliked the idea of action at a distance. He thought that all electromagnetic effects must be transmitted by some kind of medium. He gave us the concept of a field and that of lines of force. "Cutting" of lines of force is a concept due to Faraday. Faraday's was the type of mind which needs a model to explain physical phenomena. He endowed the medium with certain mechanical properties and so was able to explain in a concrete way the experimental observations. He supposed that there must be a tension along the lines of force and a pressure at right angles to these lines in an electric or magnetic field. According to a school of modern physicists, models of the Faraday type are both unnecessary and harmful. However, it must be said that without models many experimental physicists would not have been led to make the discoveries which they have made.

Faraday was not a great mathematician, and did not put his ideas of lines of force and strain in the medium into mathematical form. It required another intellect to do that. Such an intellect was forthcoming in the person of James Clerk Maxwell, who was the best mathematical physicist since Newton. Whereas Faraday came of humble origin, Maxwell (1831–1879) came from a long line of distinguished ancestors. Faraday was poor, Maxwell was rich. Faraday was self-taught, Maxwell received a university education. Faraday was an experimentalist, Maxwell a theoretical and mathematical physicist. Faraday and Maxwell constituted a team somewhat like Tycho Brahe and Kepler some two and a half centuries previously. Maxwell was born in Edinburgh, Scotland, and to the end of his life retained his Scottish brogue. When he was a schoolboy in his fifteenth year, Maxwell presented a paper to the Royal Society of Edinburgh on "The Description of Oval Curves and Those having a Plurality of Foci." Unlike Newton, Maxwell was more of the prodigy type. He was somewhat queer and eccentric, and was given to writing verses of the jingle type. In 1847 he entered the University of Edinburgh. While a student and before he was nineteen years old, Maxwell published two papers—"The Theory of Rolling Curves" and "The Equilibrium of Elastic Solids." It had been his father's intention to educate Maxwell for the legal profession. It soon became evident, however, that Maxwell's abilities were more mathematical than legal and his father allowed him to proceed to Cambridge University in 1850. From 1856 to 1860 he was professor of natural philosophy at Aberdeen, Scotland. In 1860 he became professor at King's College at London, and in 1870 he was elected to the newly founded Professorship of Experimental Physics at Cambridge. In this capacity, he superintended the planning and equipment of the now famous Cavendish Laboratory at Cambridge. He died at the early age of forty-seven. In his comparatively short life, Maxwell published over one hundred papers.

In 1857, in competing for a prize at Cambridge University, Maxwell wrote an essay on Saturn's rings. He showed that the rings must consist of swarms of small particles each particle revolving in its own orbit around the planet. This led Maxwell to consider the problem of the swarm of particles of which a gas is composed, and, as a result, he founded the modern kinetic theory of gases, which we shall consider in a later chapter.

Maxwell took Faraday's ideas on the electric and magnetic fields and put them into mathematical form. Beginning in 1861, he published papers in which was developed the modern electromagnetic theory of light. This theory was completed in 1864. This theory predicted the existence of radio waves, which were discovered by Hertz in 1887, eight years after Maxwell's death. Maxwell's electromagnetic theory seemed to be the last word in optical theory—all known optical effects could be explained by the theory. The idea that everything had now been explained took such a hold upon the minds of physicists that not a few of them believed that all the important laws of physics had been discovered and that the only thing left to do was to measure everything more accurately, or, in other words, "to investigate the next decimal place." However, such complacency was very ill-founded. Beginning in 1895 three discoveries flashed upon the world of physics and created a revolution in physical thought: the discovery of x-rays in 1895 by Roentgen in Germany, the discovery of radioactivity in 1896 by Becquerel in France, and the discovery of the electron in 1897 by J. J. Thomson in England. These discoveries set off a train of further discoveries and new physical concepts which culminated in man's controlled release of atomic energy in 1945, just fifty years after the discovery of x-rays. Even so, the end of the revolution is not in sight. Much of this book will be devoted to the advances made during the period 1895 through 1945. These advances are known collectively as "modern physics."

In closing this chapter the author wishes to acknowledge his indebtedness for much of the material which appears in this chapter to F. Cajori's *A History of Physics* (Macmillan, New York), H. Crew's *The Rise of Modern Physics* (Williams and Wilkins, Baltimore), H. Buckley's *A Short History of Physics* (Methuen, London), W. C. Dampier's *A History of Science* (Macmillan, New York), and to C. T. Chase's *A History of Experimental Physics* (Van Nostrand, New York). An attempt at a scientific interpretation of world history has been made by H. G. Wells in *The Outline of History* (Macmillan, New York). Perhaps Chapters XXII and XXIV of the *Outline* are the most interesting to physics students.

CHAPTER II

SOME USEFUL MATHEMATICS

2.1. Necessary Mathematics.—Although the reader is assumed to possess an elementary knowledge of calculus, yet there will be some readers who do not possess this knowledge. For the benefit of the reader who has had no course in calculus, we shall, in this present chapter, describe sufficient of the methods of the calculus to enable him to understand the mathematics which is used in the succeeding chapters. It is hoped that even to those readers who had had a course in calculus the method of approach to the subject will prove of interest. Further, we shall discuss the properties and use of complex quantities and of the exponential function. For this present chapter, a knowledge of algebra, trigonometry, and geometry is assumed.

2.2. Differential Calculus.—In his first year course in physics, the reader has become familiar with such terms as velocity and acceleration. The velocity of a point is the rate of change of its position. If the point is moving with uniform velocity along the axis of x and is at a distance x_1 from the origin at time t_1 and at a distance x_2 at time t_2, the velocity is obtained by dividing the distance moved, $x_2 - x_1$, by the elapsed time, $t_2 - t_1$. The formula for the velocity v is therefore

$$v = \frac{(x_2 - x_1)}{(t_2 - t_1)} \tag{2-1}$$

Let us apply this formula to the case of a body falling under gravity. In his first physics course, the reader has learned that the formula for the distance s through which a body starting from rest falls in time t is

$$s = \left(\frac{g}{2}\right)t^2 \tag{2-2}$$

where g is the acceleration of gravity. For convenience, we shall take $g = 32$ ft/sec^2. In 5 sec the body has fallen 400 ft, and in 6 sec it has fallen 576 ft. Now, apply (2-1) and we obtain $v = (576 - 400)/(6 - 5) = 176$ ft/sec. This, however, is the average velocity during the 6th second, and is greater than the instantaneous velocity at 5 sec.

Similarly, the average velocity during the fifth second is (400 − 256)/(5 − 4) = 144 ft/sec and is less than the instantaneous velocity at 5 sec. All we can say is that the instantaneous velocity at 5 sec is somewhere between 144 ft/sec and 176 ft/sec. The intervals of time taken in the preceding two cases have each been 1 sec. Now let us take an interval of 0.5 sec after the 5 seconds and an equal interval before. Again applying (2–1) and (2–2), we find the average velocity between 5 and 5.5 sec to be (484 − 400)/(5.5 − 5.0) = 168 ft/sec and the average velocity between 4.5 and 5 sec to be (400 − 324)/(5.0 − 4.5) = 152 ft/sec. We have now narrowed down our range for the value of the instantaneous velocity at 5 sec, so that this velocity is between 152 and 168 ft/sec. Suppose we take the intervals before and after 5 sec to be 0.1 sec. The instantaneous velocity is then between 16.16/0.1 and 15.84/0.1 ft/sec, that is, between 161.6 and 158.4 ft/sec. Taking the intervals to be .0001 sec, we find that the instantaneous velocity at 5 sec is between .01600016/.0001 and .01599984/.0001 ft/sec, that is, between 160.0016 and 159.9984 ft/sec. It is seen that, as the intervals are made shorter, the range of possible values for the instantaneous velocity at 5 sec grows smaller. For exceedingly small equal intervals, the instantaneous velocity is, within an exceedingly small range, about 160 ft/sec. The value of 160 ft/sec is therefore taken as the true value of the instantaneous velocity at 5 sec. The reader will note that this value of 160 ft/sec could also have been obtained by taking only one of the equal intervals. If the interval is taken after 5 sec, then the velocity is slightly greater than the instantaneous velocity, but as the interval is taken smaller and smaller the value of 160 ft/sec is approached more and more closely. When the interval is infinitesimal, the distance moved is also infinitesimal, but the quotient of the infinitesimal distance moved by the infinitesimal interval is finite, and in the case discussed the limiting value of this quotient is 160 ft/sec.

Let us consider a very small interval of time which we shall represent by dt. In (2–2), s ft is the distance through which a body falls in t sec, starting from rest. Now let the distance a body falls in $(t + dt)$sec be represented by $s + ds$. Then we have from (2–2)

$$\begin{aligned} s + ds &= \left(\frac{g}{2}\right)(t + dt)^2 \\ &= \left(\frac{g}{2}\right)\{t^2 + 2t\cdot dt + (dt)^2\} \qquad (2\text{–}3) \end{aligned}$$

Now dt is very small, so that the square of dt can be neglected and hence

$$s + ds = \left(\frac{g}{2}\right)(t^2 + 2t \cdot dt) \tag{2–4}$$

Subtracting (2–2) from (2–4), we obtain

$$ds = \left(\frac{g}{2}\right) \cdot 2t \cdot dt$$

$$= gt \cdot dt \tag{2–5}$$

Dividing both sides of (2–5) by dt, we obtain

$$\frac{ds}{dt} = gt. \tag{2–6}$$

The quantity on the left is known as the rate of change of the distance s and in this case is the velocity. In mathematical language ds/dt is called the derivative of s with respect to t. The process of finding ds/dt when s is given as a function of t as in (2–2) is called differentiation. The differential calculus has to do with rates of change or derivatives.

In the preceding, we have spoken of the rate of change of a function. In differentiating (2–2) and thereby obtaining (2–6) we have determined what is more correctly known as the time rate of change. There are other kinds of rate of change besides the time rate of change, one such being the space rate of change. If an electric current flows along a wire which has resistance, the potential at a point on the wire falls as the point is moved in the direction of the current. If the potential at a point P on a wire is V at a distance s from the place where the current enters the wire and is $V - dV$ at a distance $s + ds$, the space rate of change of potential along the wire is $-dV/ds$. Space rate of change is also called the gradient. The negative sign appears in this case because the potential decreases as the distance along the wire increases. In general, if y is a function of x, such as x^n, sin x, cos x and so on, y changes to $y + dy$ when x changes to $x + dx$, and dy/dx is called the rate of change of y with respect to x or the derivative of y with respect to x. We shall now determine a few of the more common derivatives.

Let

$$y = x^n \tag{2–7}$$

then y becomes $y + dy$ when x becomes $x + dx$, so that

$$y + dy = (x + dx)^n \tag{2–8}$$

Applying the binomial theorem of algebra, we expand the right side of (2–8) and obtain

$$y + dy = x^n + nx^{n-1}dx + \frac{n(n-1)}{1\cdot 2}x^{n-2}(dx)^2 + \cdots \tag{2–9}$$

Remembering that dx is very small and therefore neglecting the second and higher powers of dx, we have

$$y + dy = x^n + nx^{n-1}dx \tag{2–10}$$

Subtracting (2–7) from (2–10), we have

$$dy = nx^{n-1}dx \tag{2–11}$$

whence

$$\frac{dy}{dx} = nx^{n-1} \tag{2–12}$$

In this case the rate of change of y with respect to x or the derivative of y with respect to x is nx^{n-1}. Another way of stating the same thing is to say that the derivative of x^n is nx^{n-1}. This formula is also true for fractional and negative values of n.

If x^n in (2–7) is multiplied by a constant a, so that $y = ax^n$,

$$\frac{dy}{dx} = anx^{n-1} \tag{2–13}$$

The effect of a multiplying constant on the right side of (2–7) is therefore to multiply the derivative by the same constant. This rule is valid no matter what the function of x may be. Applying (2–13) to (2–2), which is

$$s = \left(\frac{g}{2}\right)t^2$$

we obtain

$$\frac{ds}{dt} = \left(\frac{g}{2}\right)\cdot 2t = gt \tag{2–14}$$

and this is the rate of change of the distance s with respect to the time t, or, in other words, the velocity of a body after it has fallen from rest for a time t.

If a constant b is added to the right side of (2–7), we have

$$y = x^n + b$$

so that, proceeding as in the case of $y = x^n$, we obtain

$$\frac{dy}{dx} = nx^{n-1}$$

as in (2–12). An additive constant therefore has no effect on the derivative. This rule is valid no matter what the function of x may be.

Let

$$y = \sin x \tag{2–15}$$

then y becomes $y + dy$ when x becomes $x + dx$, so that

$$y + dy = \sin (x + dx). \tag{2–16}$$

In trigonometry, the reader has learned that

$$\sin (A + B) = \sin A \cos B + \cos A \sin B.$$

Applying this to (2–16), we have

$$y + dy = \sin x \cos (dx) + \cos x \sin (dx) \tag{2–17}$$

Remembering that dx is small, and that therefore $\sin dx = dx$ and $\cos dx = 1$, we have

$$y + dy = \sin x + \cos x \cdot dx \tag{2–18}$$

Subtracting (2–15) from (2–18), we obtain

$$dy = \cos x \cdot dx$$

whence

$$\frac{dy}{dx} = \cos x \tag{2–19}$$

In this case the rate of change of y with respect to x is $\cos x$; or we may say that the derivative of $\sin x$ is $\cos x$.

Let

$$y = \cos x \tag{2–20}$$

Then, proceeding as in the case of $\sin x$, we have

$$\begin{aligned} y + dy &= \cos (x + dx) \\ &= \cos x \cos (dx) - \sin x \sin (dx) \end{aligned} \tag{2–21}$$

Remembering that dx is very small, we obtain

$$y + dy = \cos x - \sin x \cdot dx \tag{2–22}$$

Subtracting (2–20) from (2–22), we obtain

$$dy = -\sin x \cdot dx$$

whence

$$\frac{dy}{dx} = -\sin x \tag{2–23}$$

and the derivative of cos x with respect to x is $-\sin x$. The meaning of the negative sign is that for acute angles the cosine of x decreases as x increases.

The derivatives of various functions are given in the following table:

TABLE 2.1

DERIVATIVES

y	$\frac{dy}{dx}$
x^n	nx^{n-1}
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\sec^2 x$
e^x	e^x
$\log_e x$	$\frac{1}{x}$

A very useful rule in the differential calculus is that, if y is a function of u and u is a function of x, then

$$\frac{dy}{dx} = \left(\frac{dy}{du}\right) \cdot \left(\frac{du}{dx}\right) \tag{2–24}$$

An example of the application of this rule is as follows: Let

$$y = \cos^n x \tag{2–25}$$

and let it be required to find dy/dx. Put

$$u = \cos x \tag{2–26}$$

so that

$$y = u^n$$

Then, from (2–7) and (2–12),

$$\frac{dy}{du} = nu^{n-1} \tag{2–27}$$

and, differentiating (2–26), we have from (2–20) and (2–23)

$$\frac{du}{dx} = -\sin x \tag{2–28}$$

Hence, from (2–24),

$$\frac{dy}{dx} = -nu^{n-1} \sin x \tag{2–29}$$

or, substituting from (2–26), we have

$$\frac{dy}{dx} = -n \cos^{n-1} x \sin x \tag{2–30}$$

A second useful formula is that for the derivative of a product of two functions. Suppose that u and v are each functions of x, and that

$$y = uv \tag{2–31}$$

Then

$$\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{du}{dx} \tag{2–32}$$

As an example, let us consider the derivative with respect to x of $y = x^2 \sin x$. Put $u = x^2$ and $v = \sin x$, then $du/dx = 2x$ and $dv/dx = \cos x$, so that, applying (2–32),

$$\frac{dy}{dx} = x^2 \cos x + 2x \sin x$$

A third useful rule is that the derivative of a sum of functions of x is the sum of the derivatives of the separate functions. As an example, let

$$y = x^4 + \sin x$$

Then

$$\begin{aligned}\frac{dy}{dx} &= \frac{d}{dx}(x^4) + \frac{d}{dx}(\sin x)\\ &= 4x^3 + \cos x\end{aligned}$$

The notation $\frac{d}{dx}(x^4)$ means the derivative of x^4 with respect to x, and similarly $\frac{d}{dx}(\sin x)$ means the derivative of $\sin x$.

The derivative of a function has a very useful geometrical interpretation. When a graph is made of $y = (\frac{1}{3})x^3$ by plotting y against x a curve as in Fig. 2.1 is obtained. The slope of a straight line is the tangent of the angle between the line and the horizontal. The slope of a curve at a point is defined as the slope of the line which is tangent to the curve at the point. For instance AB is the tangent to the curve in Fig. 2.1 at the point P. The slope of the line AB is $\tan\theta$ where θ is the angle between AB and OX. The slope of the curve at P is therefore measured by BC/AC. However, if we move forward along the curve to Q and backward to R and then draw a horizontal line through R and a vertical line through Q, we obtain the figure RSQ. The closer Q and R are made to approach P the more nearly does the arc RPQ become

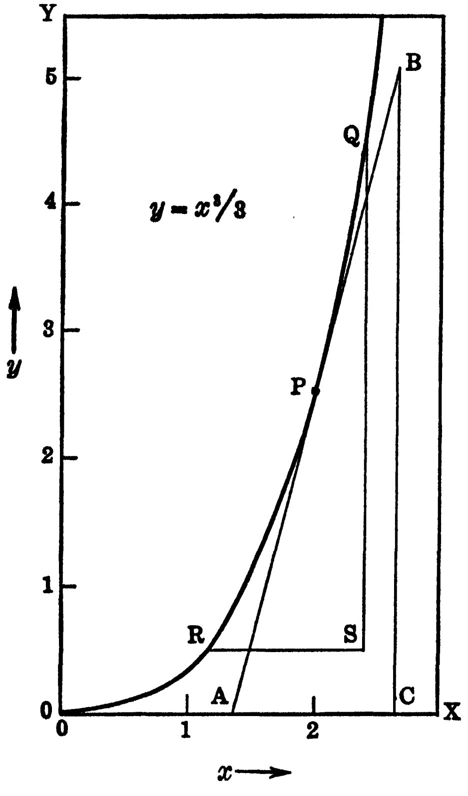


Fig. 2.1.

straight and the more nearly does its slope approach that of AB, the tangent. When RP and PQ are very small, RQS becomes a triangle similar to ABC, so that the slope of the curve at P is given by QS/RS or by dy/dx. The slope of the curve $y = (1/3)x^3$ is therefore $dy/dx = (1/3) \cdot 3x^2 = x^2$. At $x = 2$ the slope is 2^2 or 4 and $\theta = \tan^{-1} 4 = 75° 58'$.

So far we have been discussing what are known as first derivatives. Velocity is a rate of change of position; but the velocity of a falling body is not constant and is itself changing. The rate of change of velocity with respect to time is known in physics as acceleration. We may say then that acceleration is the rate of change of the rate of change of the position of a body. This double rate of change is called the second derivative in mathematics and is written in the form d^2s/dt^2. This is read as "the second derivative of s with respect to t." To obtain the second derivative of a function of x, the function is differentiated and the result differentiated again. Thus, if $y = x^n$, $dy/dx = nx^{n-1}$ and $d^2y/dx^2 = n(n - 1)x^{n-2}$. In addition to the first and second derivatives with respect to x, there are, in general, third, fourth, and still higher derivatives. These are written d^3y/dx^3, d^4y/dx^4 and so on.

The electric field strength F at a distance r from a point charge q is

$$F = \frac{q}{r^2}$$

The rate of change of F with respect to r is, applying (2–12),

$$\frac{dF}{dr} = -\frac{2q}{r^3} \tag{2–33}$$

and the rate of change with respect to r of this rate of change is

$$\frac{d^2F}{dr^2} = +2 \times \frac{3q}{r^4} = \frac{6q}{r^4} \tag{2–34}$$

The meaning of the negative sign in (2–33) is that F decreases as the distance r increases. In this case d^2F/dr^2 is a rate of change of a rate of change with respect to distance and not with respect to the time.

2.3. Integral Calculus.—Let us suppose that the instantaneous velocity of a body is given by

$$v = gt \tag{2–35}$$

and let us consider how we may obtain the distance which a body falls between the time $t = 10$ sec and the time $t = 15$ sec. At the beginning of this interval, the velocity is 32 × 10 = 320 ft/sec and at the end is 32 × 15 = 480 ft/sec. The distance traveled in this interval of 5 sec is greater than 320 × 5 = 1600 ft and less than 480 × 5 = 2400 ft. All we know is that the distance traveled is between 1600 and 2400 feet. Now let us consider the distance traveled in the 1st second after 10 sec, that is, from $t = 10$ to $t = 11$ sec. The velocity at $t = 10$ sec is 320 ft/sec and at $t = 11$ sec is 352 ft/sec, so that the distance traveled in the 1st second after 10 sec is between 320 × 1 and 352 × 1 ft or between 320 and 352 ft. The results for later seconds are shown in Table 2.2.

TABLE 2.2

Number of second after $t = 10$ sec	Distance traveled (ft)	
	Lower limit	Upper limit
1st	320	352
2nd	352	384
3rd	384	416
4th	416	448
5th	448	480
Total	1920	2080

Hence by using the smaller intervals of 1 sec and adding we find that the distance traveled is between 1920 and 2080 ft instead of between 1600 and 2400 ft. Now suppose we use intervals of one-tenth of a second. At the beginning of the first tenth of a second after 10 sec, the velocity will be 320 ft/sec and at the end it will be 323.2 ft/sec. The distance traveled in this tenth of a second will therefore be between 320 × 0.1 and 323.2 × 0.1 ft/sec or between 32.00 and 32.32 ft. Now construct a table as before and find the two totals. It will be found that the distance traveled from $t = 10$ sec to $t = 15$ sec is between 1992 and 2008 ft. The limits are now much closer. If we take still smaller intervals of, say, 0.001 sec and again find the two totals, we find the distance is between 1999.92 and 2000.08 ft. The smaller we take the intervals, the more nearly do both limits approach 2000 ft. In the notation of the calculus we divide the time from $t = 10$ sec to $t = 15$ sec into small intervals dt. We take the velocity at either the beginning or end of each interval and multiply by dt, thus obtaining the distance moved in each interval. Then we add all of these distances. If dt is small enough, it matters not whether the velocity is taken at the beginning or end of each interval.

In the notation of the integral calculus, the distance traveled is

$$s = \int v dt \tag{2–36}$$

This means that the velocity at the beginning or end of each interval dt is multiplied by dt and the results summed. The integral sign is an elongated S, representing a summation. If v is a function of t, the value of the integral can be found. In our case $v = gt$ and we write

$$s = \int gtdt \tag{2–37}$$

In the numerical problem above, it makes a difference whether we wish to know the distance traveled from $t = 10$ to $t = 15$ sec or from $t = 25$ to $t = 30$ sec, although the time of travel, 5 sec, is the same in each case. We therefore put at the top and bottom of the integral sign the limits of the integral thus

$$s = \int_{10}^{15} gtdt \tag{2–38}$$

If we wish to state (2–38) in words, we say that s equals the integral of gt with respect to t between the limits of $t = 10$ and $t = 15$. It is shown in textbooks on calculus that the integral of a function of x with respect to x is a second function of x whose derivative with respect to x is the first function of x. Integration is thus the reverse process of differentiation, and, being a reverse process, cannot always be accomplished in terms of known functions. Suppose we wish to find the integral of x^n with respect to x. It is then necessary to find that function of x whose derivative with respect to x is x^n. Such a function is $x^{n+1}/(n + 1)$, so that

$$\int x^n dx = \frac{x^{n+1}}{(n + 1)} \tag{2–39}$$

In this case the limits of the integral are not shown. The limits can only be inserted in a particular problem.

As in differentiation, a multiplying constant simply multiplies the result of integration, so that

$$\int ax^n dx = \frac{ax^{n+1}}{(n+1)} \tag{2–40}$$

Applying this to (2–38) and forgetting the limits for a moment, we have

$$s = \frac{gt^{1+1}}{(1+1)} = \frac{gt^2}{2} \tag{2–41}$$

This is correct because, if we find ds/dt, we obtain gt, which is the integrand of (2–38). The integrand is the function which appears between the integral sign and the dt in (2–38) and between the integral sign and the dx in (2–39).

Returning to the matter of the limits, the upper limit is substituted for t in the right side of (2–41) and a numerical value is obtained. A second numerical value is obtained by substituting the lower limit. The second numerical value is subtracted from the first and the result is the answer to the problem. In the above numerical problem, putting $g = 32$ ft/sec^2, we have, in the usual notation,

$$\begin{aligned} s &= \int_{10}^{15} 32t\,dt = \left[\frac{32t^2}{2}\right]_{10}^{15} \\ &= \left[16t^2\right]_{10}^{15} = 16 \times (15)^2 - 16 \times (10)^2 \\ &= 3600 - 1600 = 2000 \text{ ft} \end{aligned}$$

This is the value we obtained previously.

The derivative of a constant is zero, but the integral of a constant is not zero. Thus, if a is a constant,

$$\int a\,dx = ax$$

because, if we differentiate ax, we obtain a. Because the derivative of $x^n + c$ is nx^{n-1}, where c is a constant, we have

$$\int nx^{n-1}\,dx = x^n + c \tag{2–42}$$

c is known as a constant of integration. It is never given in tables of integrals but its presence should always be remembered. In the case of an integral with limits, the constant of integration cancels out and does not appear in the answer. *An integral such as (2–38) in which the limits are given is called a definite integral.* An integral such as (2–42) in which no limits are given is called an indefinite integral.

Just as we have second and higher derivatives in the differential calculus, so we have double and multiple integrals. An example of a double integral is

$$\int_{r=0}^{r=a}\int_{\theta=0}^{\theta=\pi} 2\pi r^2 \sin\theta dr d\theta \tag{2–43}$$

First, we integrate with respect to θ, treating r as though it were a constant, and we obtain

$$\int_{r=0}^{r=a} 2\pi r^2 \Big[-\cos\theta\Big]_{\theta=0}^{\theta=\pi} dr$$

$$= 2\pi\int_{r=0}^{r=a} r^2(-\cos\pi + \cos 0)dr = 4\pi\int_0^a r^2 dr \tag{2–44}$$

Next, we integrate with respect to r and obtain

$$4\pi\left[\frac{r^3}{3}\right]_0^a = \frac{4\pi a^3}{3} \tag{2–45}$$

A useful rule in integration is that the integral of a sum of functions of x is the sum of the integrals of the separate functions of x.

A useful geometrical interpretation of an integral is as follows: The graph of $y = x^2$ made by plotting y against x is shown in Fig. 2.2. Now consider the definite integral

$$\int_1^2 x^2\, dx \tag{2–46}$$

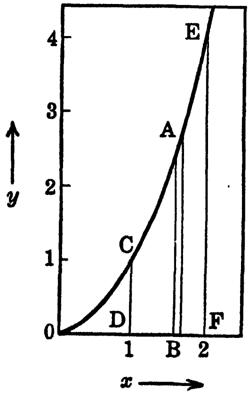


FIG. 2.2.

The width of the strip AB represents the infinitesimal dx while the height of AB represents the value of the integrand x^2 at a value of x represented by the position of B along the axis of x. The area of the strip AB represents the quantity x^2dx and the area of the sum of such strips lying between the ordinates CD and EF represents the integral (2–46). The positions of the ordinates CD and EF are determined by the limits of the integral. Thus, CD is drawn at $x = 1$ and EF at $x = 2$. Finally, the integral

(2–46) is represented by the area *DCEF* which is under the curve *OCAE* and between the ordinates *CD* and *EF*.

The above geometrical representation of an integral suggests a graphical method of obtaining the value of an integral which cannot be expressed in terms of a known function. Suppose that it is desired to determine the value of

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

graphically. The function $f(x)$ is plotted against x and two ordinates at $x = a$ and $x = b$ are drawn. The area under the curve and between the ordinates is measured by counting the squares of the graph paper. This area gives the value of the integral.

2.4. The Exponential Function.—A very important function in mathematics and physics is the exponential function e^x. The sine and cosine of x may be written in series form as follows:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \tag{2–47}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \tag{2–48}$$

The exponential function is also written in series form thus,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \tag{2–49}$$

The exponential function of x is sometimes written exp x instead of e^x.

If each term of the right side of (2–47) is differentiated, we obtain

$$\frac{d}{dx}(\sin x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

$$= \cos x \tag{2–50}$$

from (2–48). We obtained this result previously. If we differentiate each term on the right of (2–49), we obtain

$$\frac{d}{dx}(e^x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

$$= e^x \tag{2–51}$$

from (2–49). We therefore obtain the important result that e^x is its own derivative. When $x = 1$, we have

$$\begin{aligned} e &= 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots \\ &= 2.71828 \cdots \end{aligned} \tag{2-52}$$

This number is the base of natural logarithms. If

$$a = e^b$$

then

$$b = \log_e a$$

In ordinary logarithms, the base is 10. The relation between ordinary and natural logarithms is very nearly

$$\log_e a = 2.3 \log_{10} a \tag{2-53}$$

In virtue of (2-51), we also have

$$\int e^x dx = e^x + \text{const} \tag{2-54}$$

We shall have occasion to use several other integrals involving the exponential function. In the kinetic theory of gases we shall use the intergrals

$$\int e^{-x^2}\,dx \quad \text{and} \quad \int x^2\,e^{-x^2}\,dx$$

These integrals cannot be expressed in terms of known functions. However, if the limits are given, they can be evaluated graphically by the method described in Sec. 2.3. For the special case of the limits 0 and ∞ , we have the definite integrals

$$\int_0^\infty e^{-x^2}\,dx = \sqrt{\pi}/2 \quad \text{and} \quad \int_0^\infty x^2 e^{-x^2}\,dx = \sqrt{\pi}/4 \tag{2-55}$$

In the theory of specific heats we shall make use of the definite integral

$$\int_0^\infty \frac{x^3\,dx}{e^x - 1} = \pi^4/15 \tag{2-56}$$

2.5. Maxima and Minima.—A very useful application of the differential calculus is in the finding of the positions of maxima and minima. We have seen that the slope of curve $y = f(x)$ at any point is given by dy/dx. For instance, consider the curve $y = 2x - x^2$ as shown in

Fig. 2.3. This curve rises to a maximum at C. It is required to find the coordinates of C. The slope of the tangent at C is zero and therefore at the maximum $dy/dx = 0$. But in the present problem $dy/dx = 2 - 2x$, so that at the maximum

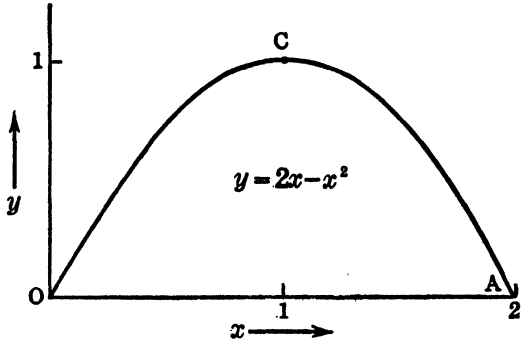


FIG. 2.3.

$$2 - 2x = 0 \tag{2-57}$$

whence $x = 1$. The maximum therefore occurs at that point whose abscissa is 1. At this value of x the ordinate $y = 2 \times 1 - 1^2 = 2 - 1 = 1$, so that the maximum occurs at the point whose abscissa is 1 and whose ordinate is 1.

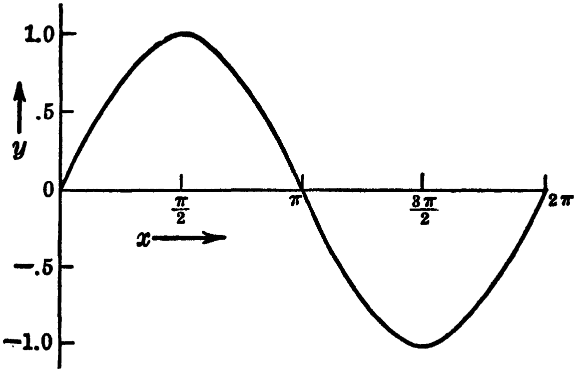


FIG. 2.4.

Consider the curve $y = \sin x$ as shown in Fig. 2.4. This curve has both maxima and minima. The derivative dy/dx is zero at both maxima and minima. In this case $dy/dx = \cos x$ and the maxima or minima occur at values of x which make

$$\cos x = 0 \tag{2-58}$$

The solution of this equation is $x = \pi/2$, $3\pi/2$, $5\pi/2$ and so on. The maxima occur at $\pi/2$, $5\pi/2$, etc., while the minima occur at $3\pi/2$, $7\pi/2$, etc. The maxima and minima are not distinguished from each other by putting $dy/dx = 0$.

The function $V^2e^{-V^2/\alpha^2}$ occurs in the kinetic theory of gases. V is the independent variable while α is a constant. The maximum value of the function $y = V^2e^{-V^2/\alpha^2}$ occurs at that value of V for which $dy/dV = 0$. Put $u = V^2$, $v = e^w$ where $w = -V^2/\alpha^2$. From (2-24), $dv/dV = (dv/dw) \cdot (dw/dV) = e^w \times (-2V/\alpha^2) = -2(V/\alpha^2)e^{-V^2/\alpha^2}$. Also $du/dV = 2V$, so that, applying (2-31) and (2-32),

$$\frac{dy}{dV} = \frac{d(uv)}{dV} = V^2 \times -2\left(\frac{V}{\alpha^2}\right)e^{-V^2/\alpha^2} + e^{-V^2/\alpha^2} \times 2V \tag{2-59}$$

Hence, at a maximum or minimum

$$\left(-\frac{2V^3}{\alpha^2} + 2V\right)e^{-V^2/\alpha^2} = 0$$

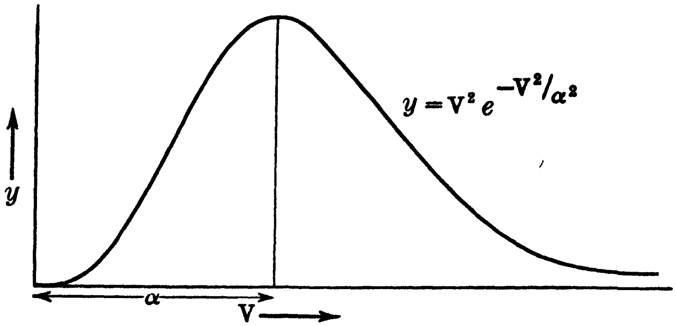


FIG. 2.5.

Either factor may be put equal to zero but the maximum occurs at that value of V which makes

$$-\frac{2V^3}{\alpha^2} + 2V = 0$$

or at $V = \alpha$. The curve $y = V^2e^{-V^2/\alpha^2}$ is shown in Fig. 2.5.

2.6. Average or Mean Values.—When we speak of the average of several quantities a_1, a_2, $a_3 \cdots a_n$, we usually imply the arithmetic average or mean of these quantities. If there are n of these quantities, then

$$\text{Arithmetic Mean} = \frac{a_1 + a_2 + a_3 + \cdots + a_n}{n} \tag{2-60}$$

However, the arithmetic mean is not the only mean which can be taken of the quantities $a_1, a_2 \cdots a_n$. Two other means of interest to physicists are the geometric mean and the root mean square. These are defined by relations

$$\text{Geometric Mean} = \sqrt[n]{a_1 a_2 a_3 \cdots a_n} \tag{2-61}$$

and

$$\text{Root Mean Square} = \sqrt{\frac{(a_1{}^2 + a_2{}^2 + \cdots a_n{}^2)}{n}} \tag{2-62}$$

We shall find considerable use for the root mean square in the theory of alternating currents and in the kinetic theory of gases. On account of its importance, we shall describe the root mean square in words: Each of the n quantities is squared and the sum of the squares is divided by the number of quantities n. The result gives the arithmetical mean of squares, or, briefly, the mean square. The square root is then taken of this mean square and the result is called the root mean square.

As an example of the above three types of mean, let us consider the three numbers 6, 8 and 10. The arithmetic mean is $(6 + 8 + 10)/3 = 8$; the geometric mean is $\sqrt[3]{6 \times 8 \times 10} = 7.830$; and the root mean square is $\sqrt{(6^2 + 8^2 + 10^2)/3} = \sqrt{66.67} = 8.165$.

Sometimes we wish to find the arithmetic mean of a quantity which is continually varying. For instance, we may desire to find the average value of $\sin\theta$ between $\theta = 0$ and $\theta = 2\pi$. In this case, we know that $\sin\theta$ is as often negative as positive, so that the average value of $\sin\theta$ is zero. Similarly the average value of $\cos\theta$ for the range $\theta = 0$ to $\theta = 2\pi$ is zero. Now consider the arithmetic mean of $\sin^2\theta$ for the range $\theta = 0$ to $\theta = 2\pi$. Although $\sin\theta$ is negative between $\theta = \pi$ and $\theta = 2\pi$, yet $\sin^2\theta$ is always positive. In order to find the average of $\sin^2\theta$ we might find $\sin^2 5°$, $\sin^2 10°$, $\sin^2 15°$, $\cdots \sin^2 355°$, $\sin^2 360°$, add all these together and divide by the total number 72. Obviously the average is the more correct the smaller the steps taken. If we take very small steps $d\theta$, then the average of $\sin^2 \theta$ is the sum

$$\sin^2 d\theta + \sin^2 2d\theta + \sin^2 3d\theta + \cdots \sin^2 2\pi$$

divided by the very large number of steps from 0 to 2π. Since the sum of a large number of quantities is involved, the formula for the mean value contains an integral.

In text books on calculus, it is shown that the average of a function $f(x)$ for the range $x = a$ to $x = b$ is

$$\text{Average} f(x) = \frac{1}{(b-a)} \int_a^b f(x)dx \tag{2-63}$$

Applying this formula for finding the average value of $\sin^2 \theta$, we obtain

$$\text{Average } \sin^2 \theta = \frac{1}{\pi 2} \int_0^{2\pi} \sin^2 \theta d\theta$$

Now $\int \sin^2 \theta d\theta = \theta/2 - (\sin 2\theta)/4$, since, on differentiation of the right side of this equation, we obtain $\sin^2 \theta$. Hence

$$\text{Average } \sin^2 \theta = \left(\frac{1}{2\pi}\right) \cdot \left[\frac{\theta}{2} - \frac{(\sin 2\theta)}{4}\right]_0^{2\pi}$$

$$= \left(\frac{1}{2\pi}\right) \cdot \pi = \frac{1}{2} \tag{2-64}$$

Hence the root mean square of $\sin \theta$ between 0 and 2π is $1/\sqrt{2}$. Similarly the root mean square of $\cos \theta$ between 0 and 2π is $1/\sqrt{2}$.

An average is sometimes expressed in terms of probability. Suppose that a quantity x can have one of the set of values $x_1, x_2, x_3, \cdots x_n$ and no other value. Then, if the probabilities of x having these various values are respectively proportional to $p_1, p_2, p_3, \cdots p_n$, the average value of x is

$$\bar{x} = \frac{p_1 x_1 + p_2 x_2 + p_3 x_3 + \cdots + p_n x_n}{p_1 + p_2 + p_3 + \cdots + p_n} \tag{2-65}$$

If p_1, p_2, etc., are the probabilities and not numbers proportional to the probabilities

$$\bar{x} = p_1 x_1 + p_2 x_2 + \cdots + p_n x_n \tag{2-66}$$

since in this case $p_1 + p_2 + \cdots + p_n = 1$.

2.7. Complex Quantities.—Consider the quadratic equation

$$x^2 - 2x + 5 = 0. \tag{2-67}$$

The roots of this equation are $1 \pm \sqrt{-4}$. The square root of a negative number is an imaginary quantity. In this case each root of the equation consists of a real part and an imaginary part. A quantity which consists of both a real and an imaginary part is called a complex quantity. For $\sqrt{-4}$ we may write $\sqrt{4 \times -1}$ or $2\sqrt{-1}$, so that the roots of the equation are $1 \pm 2\sqrt{-1}$. Becasue $\sqrt{-1}$ enters so frequently into mathematical formulas, the letter i is used for $\sqrt{-1}$, and a complex quantity is written in the style $A + iB$, where both A and B are real numbers. However, in electrical engineering and in physics the letter i is used for electric current and the letter j has come to be used for $\sqrt{-1}$. It will be the practice to use j for $\sqrt{-1}$ in this book. The two roots of the above equation will then be written $1 \pm 2j$. Two complex quantities such as $1 + 2j$ and $1 - 2j$, where one is obtained from the other by substituting $-j$ for $+j$, are called the complex conjugates of each other.

Any complex quantity can always be written in the form $A + Bj$, where A and B are real numbers. For instance, suppose we have the expression

$$\frac{1}{C + Dj}$$

Multiplying numerator and denominator by $C - Dj$, we have

$$\frac{1}{C + Dj} \times \frac{C - Dj}{C - Dj} = \frac{C - Dj}{C^2 - D^2j^2} \tag{2-68}$$

But $j^2 = -1$, so that the right side of (2–68) may be written

$$\frac{C}{C^2 + D^2} - \frac{D}{C^2 + D^2}j$$

which is in the required form. The process we have just used is called rationalizing the denominator.

A geometrical way of representing a complex quantity is as follows: Consider a complex quantity $A + Bj$, where A and B are real numbers, then A and B are taken as being the abscissa and ordinate of a point P in a rectangular system of coordinates, as shown in Fig. 2.6, where O is the origin, $OM = A$ and $MP = B$. If A is negative, OM is drawn to the left of O, and, if B is negative, MP is drawn downwards. The line OP then represents the complex quantity $A + Bj$. It is frequently convenient to represent a complex quantity by a single letter. We shall adopt the convention that small letters such as a, p, v, z will represent complex numbers, while capital letters such as A, V, R, X and Z will represent real numbers. The line OP then represents the complex number z if z is written for $A + Bj$. Now OP has both magnitude and direction and may be used to represent a vector. Hence a vector in the xy plane may be represented by a complex number. A force whose components are 4 lb to the right and 3 lb upward may be represented by the complex number $4 + 3j$. The magnitude of the force is then $\sqrt{4^2 + 3^2} = 5$ lb and its direction with respect to the horizontal is $\theta = \tan^{-1} (3/4) = 36° 52'$. Just as we can speak of the magnitude of a vector such as a force, so we can speak of the magnitude of a complex number. If we have a complex number $z = A + Bj$, we shall represent the magnitude of z by Z (i.e., the small letter representing the complex number is changed to the capital letter). The magnitude of z is then

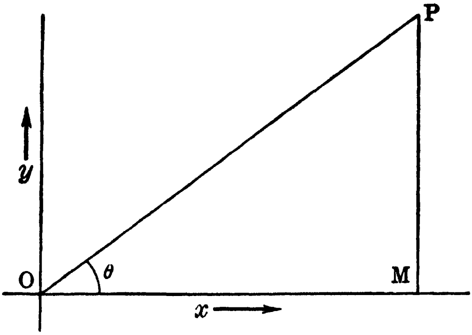


FIG. 2.6.—Geometric representation of a complex quantity.

$$Z = \sqrt{A^2 + B^2} \tag{2–69}$$

Instead of speaking of the direction of z we speak of the phase of z, the phase being given by the angle θ in Fig. 2.6. Hence the angle θ is called the phase angle, and is given by

$$\theta = \tan^{-1} (B/A) \tag{2–70}$$

The phase angle may be in any quadrant. If both A and B are positive, θ is in the first quadrant; if both are negative, θ is in the third quadrant;

if A is positive and B is negative, θ is in the fourth quadrant; and, if A is negative and B is positive, θ is in the second quadrant. Since a complex quantity may be represented by a vector and a vector by a complex quantity, we shall use the terms complex quantity and vector interchangeably.

If we multiply the two complex quantities $z_1 = A + Bj$ and $z_2 = C + Dj$, we obtain the complex quantity

$$z_1z_2 = (AC - BD) + j(AD + BC) \tag{2-71}$$

Now the phase of z_1 is $\theta_1 = \tan^{-1} (B/A)$ and that of z_2 is $\theta_2 = \tan^{-1} (D/C)$. Also the phase of z_1z_2 is θ, where

$$\tan \theta = \frac{AD + BC}{AC - BD} \tag{2-72}$$

But from trigonometry

$$\tan (\theta_1 + \theta_2) = \frac{\tan \theta_1 + \tan \theta_2}{1 - \tan \theta_1 \tan \theta_2} \tag{2-73}$$

Hence the phase θ of z_1z_2 is related to the phase of z_1 and z_2 by

$$\theta = \theta_1 + \theta_2 \tag{2-74}$$

Further, the magnitude of z_1z_2 is

$$\sqrt{(AC - BD)^2 + (AD + BC)^2}$$

and this reduces to

$$\sqrt{(A^2 + B^2)} \cdot \sqrt{(D^2 + C^2)} = Z_1Z_2 \tag{2-75}$$

where Z_1 and Z_2 are the magnitudes of z_1 and z_2, respectively. So from (2-74) and (2-75) the result of multiplying a complex quantity z_1 by a second complex quantity z_2 is to turn the vector represented by the first complex quantity through an angle equal to the phase of the second complex quantity and also to multiply the magnitude of the first complex quantity by that of the second.

There is a very important special case of the above when a vector is multiplied by j. The magnitude of j is 1 and its phase angle is 90°. So multiplication by j turns the vector through a right angle in the counter-clockwise direction without changing the magnitude of the vector. Multiplication by $-j$ turns the vector through a right angle in the clockwise direction without changing its magnitude.

We now come to the very important complex quantity p which is defined by

$$p = \sqrt{2} \, (\sin 2\pi ft - j \cos 2\pi ft) \tag{2-76}$$

where t is the time (a real variable which is not complex) and f (a real constant which is not complex) is to be determined. At $t = 0$, $p = -j\sqrt{2}$. This is a vector of magnitude $\sqrt{2}$ drawn vertically downward in Fig. 2.6. At $t = 1/(8f)$, $p = \sqrt{2}(1/\sqrt{2} - j/\sqrt{2}) = 1 - j$. This is a vector of magnitude $\sqrt{2}$ in the fourth quadrant making an angle of 45° with the horizontal in Fig. 2.6. Continuing on to $t = 1/(4f)$, $3/(8f)$, $1/(2f) \cdots 1/f$, we find that the vector rotates through one revolution in $1/f$ seconds, the magnitude of the vector remaining constant. Hence p is represented by a rotating vector which revolves at a constant angular speed of f revolutions per second, the magnitude of the vector being $\sqrt{2}$.

The real part of p at any instant of time is $\sqrt{2}\sin 2\pi ft$. From Sec. 2.6 we have seen that the root mean square of $\sin\theta$ is $1/\sqrt{2}$ and so the root mean square of the real part of the complex quantity or vector p is unity. Multiplication of p by a real quantity V means that the magnitude of the vector pV is $V\sqrt{2}$ and that the root mean square of the real part of pV is V. This multiplication does not change the phase of the rotating vector p. Multiplication of pV by a complex quantity $A + Bj$ gives us

$$u = (A + Bj)Vp \tag{2-77}$$

This is a rotating vector which is out of phase with the rotating vector p by an angle given by

$$\tan\phi = B/A \tag{2-78}$$

If B is negative and A is positive, ϕ is known as the angle of lag of u with respect to p; while, if both A and B are positive, ϕ is the angle of lead of u with respect to p. The root mean square of the real part of u is the magnitude of $(AV + BVj)$, which is

$$U = V\sqrt{A^2 + B^2} \tag{2-79}$$

Differentiation of the complex quantity p with respect to the time yields

$$dp/dt = 2\pi f\sqrt{2}\,(\cos 2\pi ft + j\sin 2\pi ft) \tag{2-80}$$

This may be written in the form

$$dp/dt = j\,2\pi f\sqrt{2}\,(\sin 2\pi ft - j\cos 2\pi ft)$$

and this in virtue of (2–76) may be written in the form

$$dp/dt = j\,2\pi f p \tag{2–81}$$

Differentiation thus has the effect of increasing the phase angle of p by a right angle in the counterclockwise direction and multiplying its magnitude by $2\pi f$.

CHAPTER II

PROBLEMS

1. Find the first and second derivatives with respect to x of the following functions of x:

$$x^5,\ \sqrt{x^3},\ \sqrt{x},\ \frac{1}{\sqrt{x}},\ \frac{1}{x},\ \frac{1}{\sqrt{x^3}}$$

2. Find the first derivatives with respect to x of the following functions of x: $\cos 3x$, $\sin^4 x$, $\log (\sin x)$, $(x - 4)^3$, $(x^3 + 9)^4$, $\sin (3x + 4)$, e^{-3x}, $1/\cos 2x$.

3. Using the relation $\tan x = \sin x/\cos x$ and the rule for the derivative of a product of functions of x, show that the derivative of $\tan x$ is $\sec^2 x$.

4. Find the derivative with respect to t of $e^{-at} \sin (bt + c)$, where a, b, and c are constants.

5. Consider the function $i = AT^2 e^{-b/T}$ in which A and b are constants. Find di/dT.

6. Show that the answers given to the following integrals are correct:

(a) $\int x^5\,dx = (x^6 + C)/6$

(b) $\int \sin^3 x\,dx = -\cos x + (\cos^3 x)/3 + C$

(c) $\int \log x\,dx = x \log x - x + C$

(d) $\int e^{-at} \sin bt\,dt = C - e^{-at} (a \sin bt + b \cos bt)/(a^2 + b^2)$

where in each case C is a constant of integration and in (d) a and b are constants:

7. Show that the maximum of the curve $y = 6x^3 e^{-2x}$ occurs at $x = 3/2$.

8. Find both the magnitude and phase of the following complex quantities.

$$\frac{25}{3 + 4j},\ \frac{169(4 + 3j)}{5 - 12j},\ \frac{1 + j}{1 - j} - \frac{1 - j}{1 + j}$$

9. Determine the arithmetic mean, the geometric mean, and the root mean square of the numbers 1, 3, 6, 10.

10. Find the average value of xe^x between $x = 1.5$ and $x = 3$, given that

$$\int xe^x\,dx = xe^x - e^x$$

11. The quantity x can have the values 5, 7, 8, 9 with probabilities respectively proportional to 1, 2, 3, 4. Find the average of x.

CHAPTER III

WAVE MOTION

3.1. Simple Harmonic Motion.—Consider a point P_1 which is revolving with uniform speed in a circle as in Fig. 3.1*a*. At each instant of time Q_1 is the projection of P_1 on the fixed straight line E_1D_1. The point Q_1 will then oscillate about O_1 between the extreme limits E_1 and D_1. The motion of Q_1 is then said to be simple harmonic. For convenience, the circle around which P_1 revolves is called the circle of reference and P_1 is called the reference point. The phase of the motion of Q_1 at a given instant of time is measured by the angle between C_1P_1 in the circle of reference and some standard direction. In Fig. 3.1*a*, the standard direction is the direction of C_1O_1, which is perpendicular to the line E_1D_1 along which the point Q_1 oscillates. The angle θ_1 is then the phase of the motion of Q_1 at the instant shown. The length O_1D_1 is called the amplitude of the simple harmonic motion and the number of to and fro oscillations of Q_1 per second is called the frequency

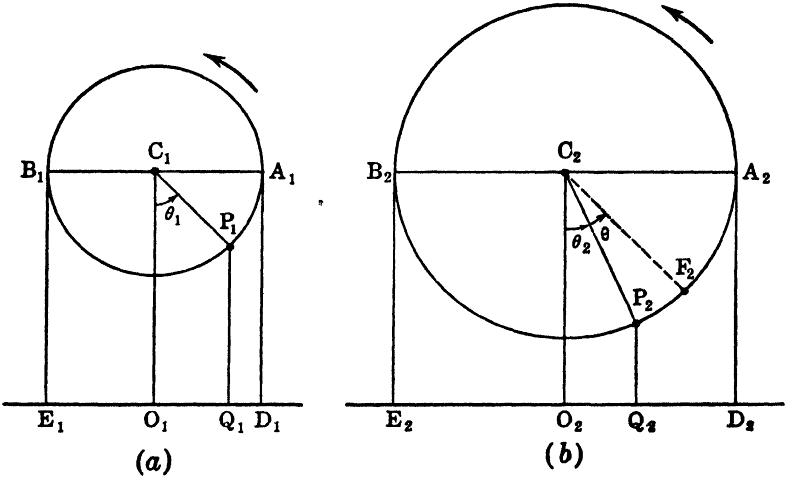


FIG. 3.1.—Relation of simple harmonic motion to uniform circular motion.

of the motion. Obviously, the frequency is the same as the number of revolutions per second of the reference point P_1.

Suppose now we have a second point Q_2 in Fig. 3.1*b* moving with simple harmonic motion of the same frequency as that of Q_1 in Fig. 3.1*a*, but of different amplitude and phase. Suppose that the reference point for Q_2 is at P_2 when the reference point for Q_1 is at P_1. The phase difference between the two motions is $\theta_1 - \theta_2$. Although θ_1 and θ_2 are both varying with the time, yet if the frequencies of the two motions are the same the phase difference $\theta_1 - \theta_2$ is a constant. In Fig. 3.1*b*, C_2F_2 is drawn parallel to C_1P_1 of Fig. 3.1*a*, so that $\theta_1 - \theta_2$ equals the angle $F_2C_2P_2$. This angle we shall represent by ϕ. The angle ϕ is called either the angle of lag or the angle of lead, depending on the point of view. If we are considering the motion of Q_2 relative to that of Q_1, the phase of Q_2 is said to be behind that of Q_1 and ϕ is called the angle of lag. For the motion of Q_1 relative to Q_2, ϕ is called the angle of lead. If time is counted from the instant when θ_1 is zero, we have $\theta_1 = 2\pi ft$ and $\theta_2 = 2\pi ft - \phi$, where f is the frequency. The distance O_1Q_1 is called the displacement and this we shall designate by x_1. Then at any instant t

$$x_1 = a_1 \sin 2\pi ft \tag{3-1}$$

where a_1 is the amplitude of the motion of Q_1. Likewise the displacement of Q_2 at any instant t is

$$x_2 = a_2 \sin (2\pi ft - \phi) \tag{3-2}$$

where a_2 is the amplitude of the motion of Q_2.

If x_1 of (3-1) is plotted as ordinate against t as abscissa and also x_2 of (3-2) is plotted against t, the sine curves shown in Fig. 3.2 are obtained. Graph 1 represents the motion of Q_1, Fig. 3.1*a*, and graph 2 represents the motion of Q_2, Fig. 3.1*b*. The graphs show that the motion of Q_2 lags behind the motion of Q_1.

A point which is moving with uniform speed in a circle has an acceleration v^2/a towards the center of the circle, where v is the linear speed of the point and a the radius of the circle. In terms of the number of revolutions per second, the acceleration is $4\pi^2af^2$. When the reference point is at P_1, Fig. 3.1*a*, its acceleration is $4\pi^2af^2$ along P_1C_1. Since Q_1 is the projection of P_1 on the line E_1D_1, the acceleration of Q_1 must be in the direction Q_1O_1 and the magnitude of the acceleration of Q_1 must equal the component of the acceleration of P_1 in the direction

D_1O_1. Hence

$$\text{Acceleration of } Q_1 = 4\pi^2 a_1 f^2 \sin\theta_1$$
$$= 4\pi^2 f^2 x_1 \qquad (3\text{–}3)$$

Thus a characteristic of simple harmonic motion is that the acceleration is proportional to the displacement and in a direction towards the undisplaced position.

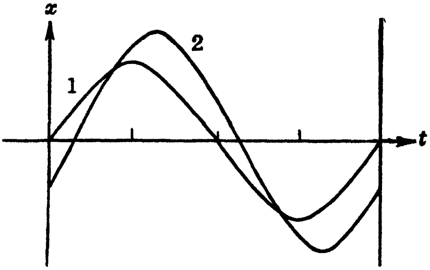


FIG. 3.2.—Representation of simple harmonic motions by sine curves.

Obviously the displacement in simple harmonic motion may be represented by the component of a rotating vector. Hence from Sec. 2.7 the displacements given by (3–1) and (3–2) may be represented respectively by the real parts of the complex quantities

$$A_1\sqrt{2}(\sin 2\pi ft - j\cos 2\pi ft) \qquad (3\text{–}4)$$

and

$$A_2\sqrt{2}\{\sin(2\pi ft - \phi) - j\cos(2\pi ft - \phi)\} \qquad (3\text{–}5)$$

where, from Secs. 2.6 and 2.7, A_1 and A_2 are the root mean square values of the displacements. Using the complex quantity p as defined by (2–76), we may replace (3–4) and (3–5) by A_1p and $A_2(\cos\phi - j\sin\phi)p$, respectively.

Let a particle of mass m be acted upon by a restoring force which is proportional to the displacement so that the force equals kx, where k is a proportionality constant and is the force for unit displacement. Then the acceleration of the particle is $(k/m)x$ towards the undisplaced position. Comparing with (3–3), we have $k/m = 4\pi^2 f^2$ or

$$f = (1/2\pi)\sqrt{k/m} \qquad (3\text{–}6)$$

Putting $f = 1/P$, where P is the period, we may write (3–6) in the form

$$P = 2\pi\sqrt{m/k} \tag{3–7}$$

In the case of rotational vibrations

$$P = 2\pi\sqrt{I/K} \tag{3–8}$$

where I is the moment of inertia and K is the restoring torque per radian of angular displacement.

✓**3.2. Energy of a Vibrating Particle.**—From the previous section a particle vibrating with simple harmonic motion has an acceleration of $4\pi^2f^2x$. If the particle has a mass m the force producing this acceleration must be

$$F = -\,4\pi^2f^2mx \tag{3–9}$$

where the negative sign indicates that the force is in the opposite direction to the displacement x. If now the particle moves from x to $x + dx$ the work done against the restoring force and therefore the increase in potential energy is $4\pi^2f^2mxdx$. If the potential energy is taken as zero when the displacement is zero, the potential energy at a displacement x is

$$\text{P.E.} = \int_0^x 4\pi^2f^2mxdx$$

$$= 2\pi^2f^2mx^2 \tag{3–10}$$

But from (3–1)

$$x = a \sin 2\pi ft \tag{3–11}$$

where a is the amplitude of the simple harmonic motion. Thus the velocity of the particle when at a displacement x is

$$v = dx/dt = 2\pi fa \cos 2\pi ft.$$

In virtue of (3–11), this may be written in the form

$$v^2 = 4\pi^2f^2(a^2 - x^2) \tag{3–12}$$

and the kinetic energy of the particle is

$$\text{K.E.} = \tfrac{1}{2}mv^2 = 2\pi^2f^2m(a^2 - x^2) \tag{3–13}$$

Hence total energy = P.E. + K.E., or

$$\text{Total energy} = 2\pi^2f^2ma^2 \tag{3–14}$$

We see from (3–14) that the total energy of a vibrating particle is proportional to the square of the amplitude and to the square of the frequency.

From Secs. 2.6 and 3.1, we see that the mean square displacement which we shall represent by $\overline{x^2}$ is one-half the square of the amplitude. Hence the relation between the total energy and the mean square displacement of the vibrating particle is

$$\text{Total Energy} = 4\pi^2 f^2 m \overline{x^2} \tag{3–15}$$

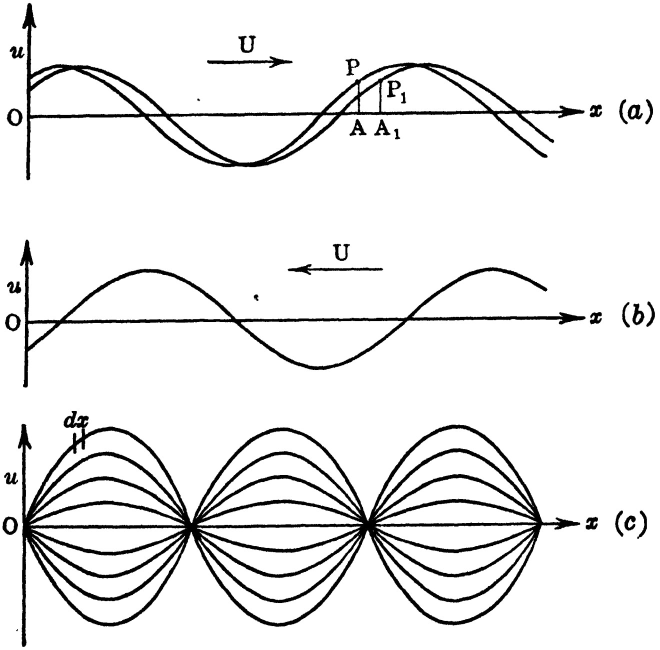


FIG. 3.3.—Progressive and standing waves.

3.3. Progressive Waves.—In these the wave form moves forward as in the case of the waves seen on the ocean. The displacement of a particle depends upon the relation of the particle to the wave form at a given instant of time. In the case of waves traveling along a rope the displacement of a particle of the rope may be written in the form

$$u = a \cos 2\pi f(t - x/U) \tag{3–16}$$

where u is the displacement AP of a particle P on the rope represented in Fig. 3.3a, x is the distance OA, and t is the time at which the displacement AP occurs. At a time $t + dt$ the wave has moved to the right so that A_1P_1, the displacement of the particle P_1, is now equal to AP. For this to be so, the new value of $(t - x/U)$ in (3–16) must be the same as before. Hence

$$t + dt - (x + dx)/U = t - x/U$$

or

$$dx/dt = U \qquad (3\text{–}17)$$

where dx is the distance AA_1. The displacement thus travels with velocity U to the right—in other words, the wave form travels with velocity U to the right. However, the particle at P does not move with this velocity U, but vibrates up and down with simple harmonic motion of frequency f as the waves pass over it.

A wave traveling along a rope to the left with velocity U may similarly be represented by

$$u = -a \cos 2\pi f(t + x/U) \qquad (3\text{–}18)$$

Such a wave is shown in Fig. 3.3b.

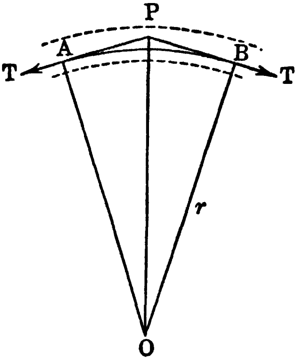


FIG. 3.4.—Rope under tension.

3.4. Wave Velocity Along a Rope.—First, let us consider a rope which passes through a small length of tube AB, as shown in Fig. 3.4, and which is under a tension T. The tube is bent into an arc of a circle of radius r. The resultant of the two forces T is along PO and equals $2T \cos \angle APO$ or $2T \sin \angle AOP$ since $\angle OAP$ is a right angle. But $\angle AOP = ds/2r$ where ds is the length of the arc AB. Hence the force inwards along an arc of length ds is $2Tds/2r = Tds/r$ since $\sin \theta = \theta$ when θ is small. Now suppose that the rope is pulled through the tube with a velocity U toward the left. The part of the rope between A and B has a mass mds where m is the mass of the rope per unit length. This mass mds is moving in a circle of radius r with velocity U and gives rise to a centrifugal force $mdsU^2/r$. The total inward pressure on the tube is therefore

$$Tds/r - mdsU^2/r$$

When this is zero we have

$$U = \sqrt{T/m} \qquad (3\text{–}19)$$

If the rope is moving with this velocity there is no pressure on the tube either inwards or outwards. The radius of curvature of the tube does not enter into (3–19) and so the curvature of the tube can change from point to point along the tube as in Fig. 3.5. Hence the rope having once been started through the tube with velocity U as given by (3–19) will retain its form or shape even though the tube is removed. Now the velocity U is simply the velocity of the wave form in Fig. 3.5 relative

FIG. 3.5.—Wave traveling along a rope.

to the particles of the rope. In other words, if the rope is held stationary and the tube of the form shown in Fig. 3.5 is started to the right with the velocity U as given in (3–19) and then the tube is removed, the wave form shown in Fig. 3.5 will travel to the right with the velocity U and at this velocity the wave form will maintain itself. The velocity of wave propagation along a rope is thus given by (3–19).

3.5. Stationary or Standing Waves.—Let us suppose that the two sets of waves represented by (3–16) and (3–18) are traveling simultaneously along the same rope. The displacement will then be the algebraic sum of the displacement produced by each set of waves, so that

$$u = 2a \sin (2\pi f x/U) \sin 2\pi f t \qquad (3\text{–}20)$$

This is not the equation of progressive waves but of stationary or standing waves. These are shown graphically in Fig. 3.3*c*. Stationary waves are observed when a string which is fixed to a rigid support at each end is vibrating. The string forms itself into nodes and loops, a node being a place of no motion. Stationary waves are produced by the interference of two sets of waves of the same frequency, one set traveling in the opposite direction to the other. Eq. (3–20) represents the stationary waves on a string which is attached to a rigid support at $x = 0$ and which everywhere had a zero displacement at $t = 0$. Comparing (3–20) with (3–1) we see that each part of the string vibrates up and down with simple harmonic motion but that the amplitude depends upon x according to $2a \sin (2\pi f x/U)$. This is zero, or the nodes occur, at those values of x which make $2\pi f x/U$ equal to

$$\cdots, \quad -2\pi, \quad -\pi, \quad 0, \quad \pi, \quad 2\pi, \quad \cdots$$

or x equal to

$$\cdots, \quad -2U/2f, \quad -U/2f, \quad 0, \quad U/2f, \quad 2U/2f, \quad \cdots$$

But in wave motion, the relation connecting the frequency f, the wave-length λ, and the velocity of propagation U is

$$f\lambda = U \tag{3-21}$$

so that the nodes occur at values of x equal to

$$\cdots, \quad -2\lambda/2, \quad -\lambda/2, \quad 0, \quad \lambda/2, \quad 2\lambda/2, \quad \cdots$$

and the length of each loop is $\lambda/2$. If each end of the string is attached to a rigid support, each support must be at a node and the total length L of the string must contain a whole number of loops and must therefore also contain a whole number of half wave-lengths. Hence

$$L = n\lambda/2 \tag{3-22}$$

It is interesting to note that (3–20), the equation which represents stationary waves, consists of two factors—an amplitude factor

$$b \sin\ (2\pi fx/U) \tag{3-23}$$

where $b = 2a$, and a time factor

$$\sin 2\pi ft \tag{3-24}$$

3.6. Energy of a Vibrating String.—Consider an element dx of the string as shown in Fig. 3.3*c*. We shall suppose that the amplitude at the center of a loop is small compared with the length of a loop. Then, if the length and mass of the string are L and M respectively, the mass of the element dx is Mdx/L. According to (3–14) and (3–23) the energy of this element, which is vibrating with simple harmonic motion, is

$$dW = (2\pi^2 f^2 Mb^2/L) \sin^2\ (2\pi fx/U)\ dx \tag{3-25}$$

To find the energy of a loop we integrate this from $x = 0$ to $x = \lambda/2$. Then, in virtue of (3–21), the energy of a loop is

$$(2\pi^2 f^2 Mb^2/L)\int_0^{\frac{\lambda}{2}} \sin^2(2\pi x/\lambda)dx = (2\pi^2 f^2 Mb^2/L)\left[\frac{4\pi x - \lambda \sin \dfrac{4\pi x}{\lambda}}{8\pi}\right]_0^{\frac{\lambda}{2}}$$

$$= \pi^2 f^2 M\lambda b^2/2L$$

This is because differentiation of the expression within the brackets gives the integrand. But $L = n\lambda/2$, where n is a whole number, so that the energy of the whole string is

$$W = \pi^2 f^2 M b^2 \tag{3-26}$$

It should be remembered that M is the mass of the whole string and that b is the amplitude at the center of a loop, as seen by reference to (3–23). Now the mean square amplitude over a loop equals $b^2/2$ and for each element dx of the string the mean square displacement $\overline{u^2}$ equals one-half the square of the amplitude of that element. Hence for the whole loop the relation between the mean square displacement for all elements in the loop and the square of the amplitude at the center of the loop is

$$\overline{u^2} = b^2/4$$

Hence, (3–26) may be written in the form

$$W = 4\pi^2 f^2 M \overline{u^2} \tag{3-27}$$

3.7. Vibrational Modes.—It is possible for a string of a given length L to vibrate in various ways. It is only necessary for each rigidly fixed end of the string to be at a node. The string may vibrate with one loop, with two loops, three loops, and so on. Each of these ways of vibrating is known as a vibrational mode. Moreover, the string may vibrate in two or more modes simultaneously. Thus a string may vibrate so as to give out its fundamental note (one loop) and its second harmonic or overtone (three loops) at the same time. Still further there is no fixed phase relationship between the vibrations of one mode and those of another. In Section 3.5 stationary waves have been represented by (3–20), which includes a time factor given by (3–25). This time factor may contain a phase angle ϕ as in

$$\sin\,(2\pi f t - \phi) \tag{3-28}$$

If a string is set vibrating in the first and third modes and if ϕ_1 and ϕ_3 are the phase angles for the modes respectively, the phase difference $\phi_3 - \phi_1$ will have a certain value. If, however, the string is set vibrating a second time, the new value of the phase difference $\phi_3 - \phi_1$ may have no relation to its former value of the phase difference.

In a later chapter we shall approach Debye's theory of specific heats through the concept of the vibrational modes of a solid substance. The thermal vibrations of a solid are considered as being due to the simultaneous vibration of the solid in many vibrational modes whose phases are at random with respect to each other. Thermal agitation always contains the idea of randomness. Now the vibrational modes in the theory of specific heats are of high frequency so that f is large. From (3–22)

$$n = 2L/\lambda \tag{3–29}$$

This is the relation between the ordinal number n of the vibrational mode and the wave length λ of the mode. If the n_1th mode has a wave-length λ_1 and the n_2th mode a wave-length λ_2, then from (3–29) the number of modes whose wave-lengths are between λ_1 and λ_2 is

$$n_2 - n_1 = 2L\left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right) = \frac{2L(\lambda_1 - \lambda_2)}{\lambda_1\lambda_2} \tag{3–30}$$

If, now, $(n_2 - n_1)$ is small compared to either n_1 or n_2 and likewise $(\lambda_2 - \lambda_1)$ is small compared to either λ_1 or λ_2, we may in the language of the calculus replace $n_2 - n_1$ by dn, $(\lambda_2 - \lambda_1)$ by $d\lambda$, and $\lambda_1\lambda_2$ by λ^2. Then from (3–30) we may say that the number of vibrational modes in the wave-length range from λ to $\lambda + d\lambda$ is

$$dn = 2Ld\lambda/\lambda^2 \tag{3–31}$$

The vibrations of a string are transverse and so can take place in two directions at right angles to the length of the string. Any vibration in an intermediate direction can be considered as the vector sum of two vibrations at right angles to each other. Since each of these vibrations is independent of the other, both in phase and in amplitude, each is counted as a vibrational mode. Hence for transverse waves the number of vibrational modes in the wave-length range λ to $\lambda + d\lambda$ is

$$dn = 4Ld\lambda/\lambda^2 \tag{3–32}$$

while for longitudinal waves such as those in an organ pipe open at both ends the number of vibrational modes is given by (3–31).

3.8. Vibrating Square Sheet and Cube.—We now consider the vibrational modes in a vibrating square sheet such as in the membrane stretched over a square drum. The membrane is rigidly attached along the sides of the square $OXEY$ in Fig. 3.6. The line AB represents the position at time t of the crest of a wave traveling with velocity U in the direction GH which makes an angle α with OX. We shall suppose that the membrane is under equal tension in all directions so that the velocity of wave propagation is the same in all directions. The displacement of a point P whose coordinates are x and y, OX and OY being taken as axes of coordinates, may be represented by

$$u = a \cos 2\pi f\left(t - \frac{x \cos \alpha + y \sin \alpha}{U}\right) \tag{3-33}$$

The stationary waves on the sheet are the result of the interaction of four sets of progressive waves of the type represented by (3–33). In Fig. 3.6, let AB and CD represent two adjacent wave crests so that the perpendicular distance between them is the wave-length λ. These two wave crests cut OX in B and D and OY in A and C. From Fig. 3.6 it is seen that $BD = \lambda/\cos \alpha$ and $AC = \lambda/\sin \alpha$. The conditions for standing waves are that the quotient of the length of OX by half the length of BD must be a whole number n_1 and the quotient of the length of OY by half the length of AC must be another whole number n_2, so that

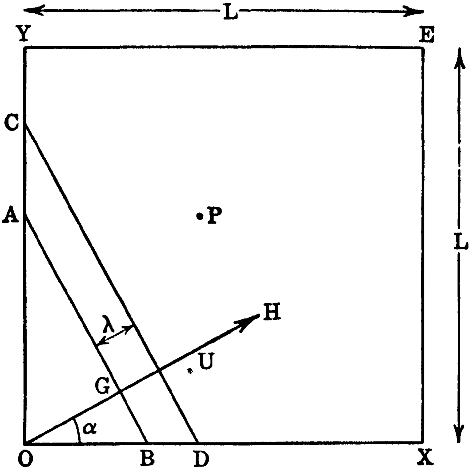


FIG. 3.6.—Waves on a stretched membrane.

$$L = n_1\lambda/(2 \cos \alpha) \tag{3-34}$$

and

$$L = n_2\lambda/(2 \sin \alpha) \tag{3-35}$$

where L is the length of the side of the square in Fig. 3.6. Rearranging (3–34) and (3–35), we have

$$\cos \alpha = n_1\lambda/2L \quad \text{and} \quad \sin \alpha = n_2\lambda/2L$$

Then since

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

we have, after squaring, adding, and rearranging,

$$\sqrt{n_1^2 + n_2^2} = 2L/\lambda \tag{3-36}$$

We can no longer speak of the nth vibrational mode. Instead, we must say that the membrane is vibrating in the n_1th mode along the x axis and in the n_2th mode along the y axis. We shall abbreviate this by saying that the membrane is vibrating in the (n_1, n_2)th mode. In order to determine the number of modes in the wave-length range λ to $\lambda + d\lambda$ we shall take a numerical example. Let the wave-length range be from 2/11 to 2/10 cm and let $L = 1$ cm. Then for the longest wave-length the right side of (3-36) has the value 10 and for the shortest wave-length the value 11. We now find those integral values of n_1 and n_2 for which the value of $(n_1^2 + n_2^2)$ falls between 10^2 and 11^2. To make the problem more definite, we require that $(n_1^2 + n_2^2)$ is one of the numbers 101, 102, $\cdots$ 121. The method is explained in Table 3.1. In this table there are shown 8 ways of choosing n_1 and n_2 so that $(n_1^2 + n_2^2)$ falls in the range from 100 through 121. But for each of these 8 ways n_1 and n_2 may be reversed, so that the total number of ways is $2 \times 8 = 16$. There are thus 16 vibrational modes which in this problem have their wave-lengths between 2/10 and 2/11 cm. By methods which are beyond the scope of this book it can be shown that the number of modes of wave-lengths between λ and $\lambda + d\lambda$ is very nearly given by

TABLE 3.1

n_1	n_2	n_1^2	n_2^2	$n_1^2 + n_2^2$
7	8	49	64	113
6	9	36	81	117
5	9	25	81	106
4	10	16	100	116
3	10	9	100	109
2	10	4	100	104
1	10	1	100	101
0	11	0	121	121

$$dn = 2\pi L^2 d\lambda/\lambda^3 \tag{3-37}$$

In the above example $d\lambda = 2/10 - 2/11 = 2/110$ and we take $\lambda = 21/110$, the average of the two wave-lengths. Substituting in (3-37), we obtain 16.4, which is in reasonably good agreement with the 16 modes obtained from Table 3.1.

Similarly we may consider the vibrational modes in a solid cube of side L. In place of (3–29) and (3–36) we now obtain

$$\sqrt{n_1^2 + n_2^2 + n_3^2} = 2L/\lambda \tag{3–38}$$

where n_1, n_2 and n_3 are whole numbers. As in the case of the membrane, we must say that the solid cube is vibrating in the n_1th mode along the x-axis, in the n_2th mode along the y-axis, and in the n_3th mode along the z-axis; and we shall abbreviate this by saying that the solid cube is vibrating in the (n_1, n_2, n_3)th mode. We could use a numerical example and construct a table similar to Table 3.1, but with three whole numbers instead of two. In this case, we must note the following: if the three numbers n_1, n_2, n_3 are all different (say, 3, 5, 6). the number of ways of arranging these numbers is 6; if two of the three numbers are different (say, 3, 5, 5), the number of ways of arranging the numbers is 3; if all three numbers are the same (say, 5, 5, 5), there is only one way of arranging the numbers. The table would show that the number of modes with wave-lengths between λ and $\lambda + d\lambda$ is very nearly given by

$$dn = 4\pi L^3 d\lambda/\lambda^4 \tag{3–39}$$

As in the case of the vibrating string and organ pipe we must consider whether we are dealing with transverse or longitudinal waves. Eq. (3–39) is for longitudinal waves. For transverse waves, the number of modes with wave-lengths between λ and $\lambda + d\lambda$ is very nearly given by

$$dn = 8\pi L^3 d\lambda/\lambda^4 \tag{3–40}$$

For some purposes it is more convenient to have formulas for the number of modes in a frequency range rather than in a wave-length range. To obtain these formulas we use the relation

$$\lambda = U/f \tag{3–41}$$

where U is the velocity of wave propagation and f is the frequency. From (3–41), if U is a constant, we have

$$d\lambda = -\,U df/f^2 \tag{3–42}$$

Hence for longitudinal waves the number of modes in a solid cube in the frequency range f to $f + df$ is

$$dn = 4\pi L^3 f^2 df/U^3 \tag{3–43}$$

while for transverse waves the number is

$$dn = 8\pi L^3 f^2 df/U^3 \tag{3–44}$$

3.9. Wave and Group Velocities.—If a stone is dropped into the surface of a lake, a group of waves and ripples is excited. These waves and ripples spread out in ever-widening circles. This group of waves appears to have a definite speed and the speed is different from that of a particular wave crest in the group. The group of waves is represented in Fig. 3.7. As the group moves forward, the crest A diminishes and the crest B grows until B becomes the maximum wave-crest at the middle of the group. Still later the crest C becomes the maximum and so on. In the case of water waves whose wave-lengths are not too short the group velocity is less than the wave velocity. The resaon for this phenomenon is that waves of different wave-lengths move with different velocities. The reader is familiar with the fact that when white light is passed through a prism it is broken up into the colors of the rainbow. This phenomenon is called dispersion and is due to the fact that the refractive index depends upon the wave-length. According to the wave theory, this means that the velocity of light in the prism is a function of the wave-length.

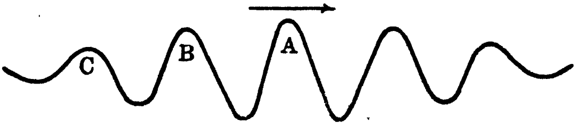


FIG. 3.7.—The center of the group of waves travels with a different velocity from that of a single wave.

Imagine two wave trains of wave-lengths λ and $\lambda + d\lambda$ as shown in Fig. 3.8. Suppose that the train of wave-length λ travels with a velocity U, and that the train of wave-length $\lambda + d\lambda$ travels with velocity $U + dU$. If the two wave trains are superposed, the crest A_2 superposes upon the crest A_1. However, the crests B_2 and C_2 do not exactly superpose upon the crests B_1 and C_1 respectively, so that superposition of the two wave-trains produces a wave group as shown in Fig. 3.7. The maximum crest of Fig. 3.7 results from the superposition of A_1 and A_2 in Fig. 3.8. However, this superposition occurs only at a certain instant of time because the wave train (b) of Fig. 3.8 is traveling faster than train (a) of Fig. 3.8. After a while B_2 will have overtaken B_1, and the middle wave of the group as shown at A in Fig. 3.7 will have moved back a distance λ with respect to the wave train (a) in Fig. 3.8. Let t be the time for B_2 to overtake B_1, then the distance traveled by the middle of the group in this time is $Ut - \lambda$. If we divide this distance by the time t, we obtain the group velocity V, so that

$$V = U - \frac{\lambda}{t} \tag{3-45}$$

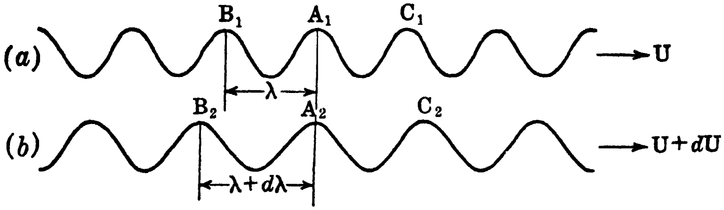


FIG. 3.8.—Waves of different wave-lengths moving with different velocities.

Now, the relative velocity of wave train (*b*) with respect to wave train (*a*) in Fig. 3.8 is dU, so that

$$t = \frac{d\lambda}{dU} \tag{3-46}$$

Hence (3-45) becomes

$$V = U - \lambda\frac{dU}{d\lambda} \tag{3-47}$$

This is the formula for the group velocity.

For water waves which are not too short the wave velocity U is given by

$$U = \sqrt{\frac{g\lambda}{2\pi}} \tag{3-48}$$

whence

$$\frac{dU}{d\lambda} = \sqrt{\frac{g}{8\pi\lambda}} \tag{3-49}$$

Therefore in the case of water waves

$$V = U - \sqrt{\frac{g\lambda}{8\pi}} \tag{3-50}$$

In virtue of (3-48), this may be written

$$V = U(1 - 1/2) = 0.5U \tag{3-51}$$

It is seen that the group velocity in water is one-half the wave velocity when there is a range of wave-lengths from λ to $\lambda + d\lambda$.

In all cases where the velocity of light is measured directly the velocity measured is the group and not the wave velocity. On the other hand, when it is said that the refractive index is equal to the ratio of the velocities in the two media involved, the velocities referred to are the wave velocities. If the medium is free from dispersion so that there is no variation of wave velocity with wave-length and therefore $dU/d\lambda = 0$, the group velocity equals the wave velocity. This is the case for light traveling in a vacuum. When Foucault measured the velocity of light in water he measured the group velocity. Water is not a highly dispersive liquid, so that Foucault's result appeared to agree very well with the requirements of the wave theory. However, had he measured the velocity in carbon bisulphide, as was done later by Michelson, the agreement would not have been so good. In the case of carbon bisulphide, which is a highly dispersive liquid and for which $dU/d\lambda$ is not negligible the difference between V and U should be about 7.5 percent. Michelson found the velocity of light in air to be 1.758 times that in carbon bisulphide. From the refractive indices of air and carbon disulphide the velocity of air should be 1.64 times that in carbon bisulphide. Increasing 1.64 by 7.5 percent we obtain 1.763, which is in reasonably good agreement with 1.758.

3.10. Doppler Effect.—If both a sounding body and the observer are at rest, the number of condensations and rarefactions entering the ear of the observer in one second is the same as the number of vibrations of the sounding body per second, so that the frequency of vibration of the drum of the ear is the same as that of the sounding body. However, if the sounding body is approaching the observer at a velocity v, the waves in the air (or other medium) between the sounding body and the observer are crowded together as shown in Fig. 3.9. In the time t the sound will have traveled a distance Vt where V is the velocity of sound in the medium. If the sounding body is at rest as in (*a*) of Fig. 3.9, the number of waves (a compression plus a rarefaction constitutes a wave) in the distance Vt is Vt/λ, where λ is the wave-length. If, however, the sounding body is moving as in (*b*) the same number of waves is crowded into the length $Vt - vt$ and the new wave-length is λ' where

$$\lambda' = \frac{(V - v)t}{\frac{Vt}{\lambda}} = \frac{\lambda(V - v)}{V} \tag{3-52}$$

Since the relation between the frequency f' of the moving sounding body as heard by the observer who is at rest and the wave-length λ' is $f'\lambda' = V$, we have

$$f' = \frac{fV}{(V - v)} \tag{3-53}$$

for the relation between the apparent frequency f' of the sounding body and the true frequency f.

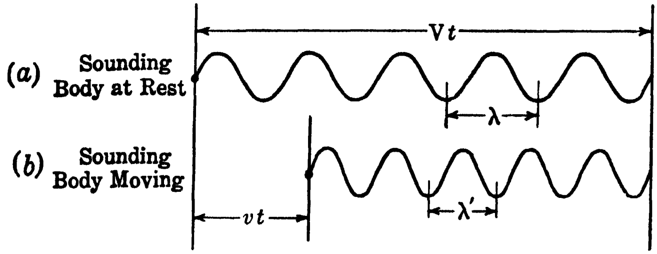


FIG. 3.9.—Representation of the Doppler effect on the classical wave theory.

We next consider the case of an observer traveling with velocity v toward a sounding body at rest with respect to the medium. In a time t, the ear of the observer gathers in vt/λ more wave-lengths than when the observer is at rest. So the total number of wave-lengths entering the ear in the time t is $ft + vt/\lambda$. The apparent frequency is therefore

$$f' = \frac{ft + vt/\lambda}{t} \tag{3-54}$$

But $\lambda = V/f$, so that

$$f' = \frac{f(V + v)}{V} \tag{3-55}$$

CHAPTER III

PROBLEMS

1. A helical spring hangs vertically. When a mass of 100 gm is hung at the lower end of the spring, the spring is extended by 5 cm. (*a*) Find the force constant *k*. Next, the mass is set vibrating up and down. (*b*) Find the frequency of vibration. The acceleration of gravity is 980 cm/sec^2.

2. A disc whose moment of inertia is 2500 gm–cm^2 is suspended by a wire. It is set into rotational vibration and the period of vibration is 5 sec. Calculate the torque constant *K* of the wire.

3. The mass of 100 gm in Problem 1 is set vibrating with an amplitude of 1.5 cm. (*a*) Determine the total energy of the vibrating mass. (*b*) If the total energy is 3000 ergs, determine the root mean square displacement.

4. One end of a wire is rigidly attached at *A*, while to the other end, which hangs vertically from a pulley at *B*, is attached a weight of 500 gm. The wire between *A* and *B* is horizontal. The length and mass of the portion *AB* of the wire are 50 cm and 2 gm, respectively. (*a*) Find the velocity of waves in the wire. (*b*) Determine the energy of the mode whose frequency is 700 vib/sec when the root mean square displacement of the string is 1/100 mm.

5. The velocity of very short waves on the surface of a liquid is given by

$$U = \sqrt{\frac{2\pi T}{\lambda\rho}}$$

where T is the surface tension and ρ the density of the liquid. Derive the formula for the group velocity of these short waves. Determine the group velocity on the surface of water when the wave-length varies within a small range about 0.05 cm. The surface tension of water is 74 dyne/cm.

6. A sounding tuning fork of natural frequency 512 vib/sec is traveling towards an observer at a speed of 30 mi/hr. What is the apparent frequency of the sound as heard by the observer? The velocity of sound is 1100 ft/sec.

7. Consider a cube each of whose edges has a length of 5 cm. Determine the number of longitudinal modes of vibration with wave-lengths between 1 2/3 and 2 cm (*a*) by means of a table similar to Table 3.1, and (*b*) by means of the formula (3–39). The answer of (*b*) depends on the value of λ used. The formula is approximate, not exact.

CHAPTER IV

ALTERNATING CURRENTS

4.1. Systems of Units.—In 1784 Coulomb, a French military engineer, by means of a torsion balance found that the force F between two electric charges, q_1 and q_2, is given by

$$F = \frac{q_1 q_2}{Kd^2} \tag{4–1a}$$

where d is the distance between the two charges and $1/K$ is a constant of proportionality. The quantity K is known as the dielectric constant of the medium between the charges. In the classical electric system of units, K is put equal to unity when the charges are in a vacuum, so that (4–1a) becomes

$$F = \frac{q_1 q_2}{d^2} \tag{4–1b}$$

If the unit of force is taken as the dyne, the definition of unit charge follows from (4–1b). The units of all other quantities in the electric system, such as potential, electric field strength, and electric current, are based on this unit charge.

Coulomb's law for the force F between magnetic poles of strength m_1 and m_2 may be written as

$$F = \frac{m_1 m_2}{\mu d^2} \tag{4–2a}$$

where d is the distance between the poles and $1/\mu$ is a constant of proportionality. The quantity μ is known as the permeability of the medium between the poles. In the classical magnetic system of units, μ is put equal to unity, so that (4–2a) becomes

$$F = \frac{m_1 m_2}{d^2} \tag{4–2b}$$

The definition of unit pole strength is based upon (4–2b). The units of magnetic field strength and magnetic flux are based on this unit pole strength.

Before an electromagnetic system of units could be constructed, it was necessary to find some experimental relation between an electric charge and a magnetic pole. The action of an electric current on a magnet was announced by Oersted in 1820. This discovery led Ampère of France to a mathematical statement of the law of force between a current and a magnetic pole. We shall state Ampère's law in the following simple form: The force F exerted by a long straight wire carrying a current i on a magnetic pole of strength m at a distance r from the wire is

$$F = 2im/cr \tag{4–3a}$$

where $1/c$ is a proportionality constant. However, unlike K and μ in (4–1a) and (4–2a), c does not depend on the medium between the wire and the pole. It should be noted that (4–3a) contains i, not q. Finally, in 1878 Rowland of Johns Hopkins found by rapidly rotating a hard rubber disc carrying charged gilt patches that moving charges have an effect on a magnetic needle similar to that of an electric current flowing in a wire. Hence i in (4–3a) can be replaced by q/t, where q is the charge passing through a cross section of the wire in a small interval t. We can now write (4–3a) in the form

$$F = 2qm/crt \tag{4–3b}$$

We can now construct a system of units for electric and magnetic quantities from (4–2a) and (4–3b) or a second system from (4–1a) and (4–3b). If, in constructing the first system, we put $\mu = 1$ for a vacuum and $c = 1$, we arrive at the absolute electromagnetic system of units. Electric and magnetic quantities measured in this system are given in electromagnetic units (abbreviated to "emu"). If, in constructing the second system, we put $K = 1$ for a vacuum and $c = 1$, we arrive at the electrostatic system. A better name would have been "magnetoelectric" system. Electric and magnetic quantities measured in the electrostatic system are given in electrostatic units (abbreviated to "esu").

Physicists tend to use a system of units which combines certain more convenient features of the electric and magnetic systems. In it, electric quantities, such as current, charge, and potential, are measured in esu, whereas magnetic quantities, such as magnetic field strength and pole strength, are measured in emu. In this mixed system, Rowland found that the numerical value of c in (4–3b) was about 3×10^{10}. Later experiments have shown that $c = 3 \times 10^{10}$ with a high degree of accuracy.

The practical system of units used in electrical engineering is based on the absolute electromagnetic system. The ampere is the unit of current in the practical system. It is the custom among electrical engineers to refer to 1 emu of current as 1 abampere. The prefix "ab" comes from "absolute." Similarly, the electrical engineers speak of abvolts and abohms. Hence we write

$$
\begin{aligned}
1 \text{ ampere} &= 0.1 \text{ emu current} = 0.1 \text{ abampere} \\
1 \text{ volt} &= 10^8 \text{ emu potential difference} = 10^8 \text{ abvolts} \\
1 \text{ ohm} &= 10^9 \text{ emu resistance} = 10^9 \text{ abohms}
\end{aligned}
$$

4.2. Physical Dimensions.—The quantities K and μ which appear in (4–1a) and (4–2a), respectively, measure real physical properties of the media surrounding the electric charges or magnetic poles. From (4–1a) we may infer that

$$[q/\sqrt{K}] = M^{1/2}L^{3/2}T^{-1} \tag{4–4a}$$

where the brackets [ ] mean "the dimensions of" the expression inside the brackets, and M, L, and T represent the dimensions of the fundamental units of mass, length, and time, respectively. Similarly, from (4–2a) we infer that

$$[M/\sqrt{\mu}] = M^{1/2}L^{3/2}T^{-1} \tag{4–4b}$$

Further, from (4–3b) we infer that

$$[qm/c] = ML^2T^{-1} \tag{4–5}$$

Multiplying (4–4a) by (4–4b) and dividing by (4–5), we obtain

$$[c/\sqrt{K\mu}] = LT^{-1} \tag{4–6a}$$

Since c does not depend on the nature of the medium, it seems convenient and plausible to make it a pure number without dimensions. Then (4–6a) becomes

$$[1/\sqrt{K\mu}] = LT^{-1} \tag{4–6b}$$

so that the dimensions of $1/\sqrt{K\mu}$ are those of a velocity.

4.3. Alternating Electromotive Force.—If a circuit represented as *ABCD* in Fig. 4.1 is rotated about a vertical axis in the earth's magnetic field, an electromotive force is produced in the circuit. The number of magnetic lines of force threading their way through a circuit is called the magnetic flux through the circuit. For brevity, we shall write "e.m.f." for "electromotive force." By Faraday's law of electromagnetic induction, the induced e.m.f. E in a circuit is given by

$$E = -\frac{dN}{dt}$$

where N is the magnetic flux and dN/dt is the time rate of change of the flux. The negative sign is a mathematical way of indicating Lenz's law, which states that the induced e.m.f. is in such a direction as to oppose the agency which produces it.

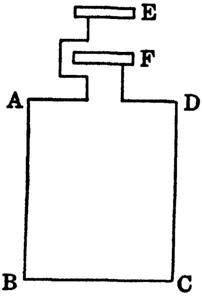


Fig. 4.1.—Simple alternating-current generator.

Let H be the strength of the horizontal component of the earth's magnetic field, A the area of the circuit *ABCD* and f the number of revolutions per second of the circuit about a vertical axis. Let *AD*,

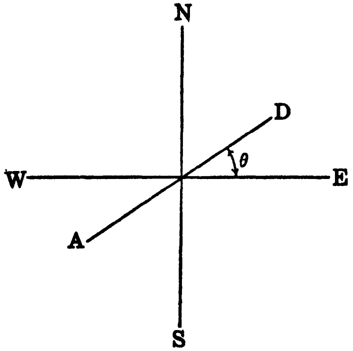


Fig. 4.2.

Fig. 4.2, represent the trace of the circuit *ABCD* and let *NS* represent the magnetic north-south direction, and let the plane of the circuit make an angle θ with the west-east direction. The flux through the circuit is $HA \cos \theta$. If the circuit represented by *AD* is rotating, the angle θ is changing, so that

$$E = -\frac{d(HA\cos\theta)}{dt} = HA\sin\theta\frac{d\theta}{dt} \qquad (4\text{–}7)$$

Now suppose that time is measured from the moment when $\theta = 0$, so that

$$\theta = 2\pi ft \qquad (4\text{–}8)$$

where f is the number of revolutions per second of the circuit. From (4–8)

$$\frac{d\theta}{dt} = 2\pi f \qquad (4\text{–}9)$$

Substituting from (4–8) and (4–9) in (4–7), we obtain

$$E = 2\pi fHA \sin(2\pi ft) \qquad (4\text{–}10)$$

where E is the induced e.m.f. at the time t. It is seen that E varies with the time and is a maximum when $\sin(2\pi ft) = 1$. This maximum value we designate by E_m, so that

$$E_m = 2\pi fHA \qquad (4\text{–}11)$$

We may therefore write (4–10) thus

$$E = E_m \sin(2\pi ft) \qquad (4\text{–}12)$$

If the ends of the circuit $ABCD$ are connected to two collector rings represented as E and F in Fig. 4.1 and these collector rings are connected to an outside circuit by means of brushes, the current in the outside circuit will be alternating. From (4–12) we see that the e.m.f. begins at zero, rises to a maximum, sinks to zero, becomes negative, reaches a negative maximum and then returns to zero. This is known as a cycle. The quantity f in (4–12) is the number of cycles the e.m.f. makes in one second and is known as the frequency. If f is given this meaning, the e.m.f. of most alternating current generators is very nearly given by (4–12). We shall call E in (4–12) the instantaneous e.m.f. so as to distinguish it from the maximum e.m.f. E_m. The frequency of an alternating e.m.f. is given in "cycles per second." However, it has become the practice to omit the phrase "per second" and the frequency is given in "cycles." A frequency of 60 cycles therefore really means a frequency of 60 cycles per second.

If from a fixed point O in a fixed straight line a distance OQ equal (or proportional) to the e.m.f. produced by an a.c. (we shall in future abbreviate "alternating current" to a.c.) generator is marked off, the point Q will vibrate with simple harmonic motion about O, and we see therefore that the root mean square of an alternating e.m.f. is $E_m/\sqrt{2}$. If an a.c. voltmeter is connected across an a.c. generator, the voltmeter reading is the root mean square of the e.m.f. The electric light mains in most houses supply 110-volt a.c. to the lamps. The voltage of 110 is the voltage read on an a.c. voltmeter and is $1/\sqrt{2}$ times the maximum voltage supplied by the mains. The current measured on an a.c. ammeter is also the root mean square of the current and is therefore $1/\sqrt{2}$ times the maximum current.

Comparing (4–12) with the results of Secs. 2.7 and 3.1, we see that the instantaneous value of an a.c. voltage at a time t is given by the real part of a complex quantity v, where

$$v = V\sqrt{2}(\sin 2\pi ft - j \cos 2\pi ft) \tag{4–13}$$

where $V = E_m/\sqrt{2}$ and is the root mean square of the voltage as read on an a.c. voltmeter. It is also found that a current through an a.c. circuit or through a branch of an a.c. circuit is given by the real part of a complex quantity a, where

$$a = A\sqrt{2}\,\{\sin(2\pi ft + \phi) - j\cos(2\pi ft + \phi)\} \tag{4–14}$$

where A is the root mean square of the current flowing in the circuit or branch of the circuit as read by an a.c. ammeter. The quantity ϕ in (4–14) is the angle of phase difference between the rotating vector represented by a and that represented by v when v is the complex voltage across the circuit or branch of the circuit through which the complex current a is flowing. If ϕ is positive, the current is ahead of the voltage in phase and ϕ is called the angle of lead; whereas, if ϕ is negative, the current is behind the voltage in phase and ϕ is then called the angle of lag. Using the symbol p as defined by (2–76), we may replace (4–13) by

$$v = Vp \tag{4–15}$$

and (4–14) by

$$a = A(\cos\phi + j\sin\phi)p \tag{4–16}$$

4.4. Inductance.—If a wire is wound into a coil and a current is sent through the wire, a magnetic field is produced in the coil. For a coil wound upon a certain form or bobbin the field inside of the coil is very nearly proportional to n the number of turns on the bobbin, if there is no magnetic material in the core of the coil. Also for a coil of n turns the flux through each turn is counted when the "total flux" or "flux linkage" through the coil is being determined. Hence the flux linkage is very nearly proportional to n^2 for a given current. If there is no magnetic material such as iron in the core, the flux linkage is exactly proportional to the current through a given coil, so that the flux linkage is Li when a current i is flowing through the coil. L is a proportionality constant and is known as the inductance of the coil. The magnitude of the inductance of a coil is determined by its construction and L is very nearly proportional to n^2 for a given bobbin. If now an alternating current is sent through a coil of inductance L, the flux linkage is continually varying with the time, and this continually varying flux induces an e.m.f. in the coil. This induced e.m.f. by Lenz's law tends to prevent the variation of the current which produces it, and so constitutes what we shall call a "back e.m.f." Obviously, the back e.m.f. $= d(Li)/dt$ or, since L is a constant,

$$\text{Back e.m.f. (inductance)} = L di/dt \tag{4–17}$$

If the current is measured in amperes and the back e.m.f. in volts, the inductance is measured in henrys. A coil is said to have an inductance of one henry when, due to the current varying at the rate of one ampere per second, the back e.m.f. induced in the coil is one volt.

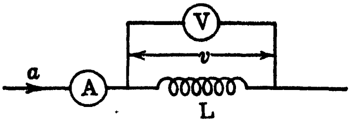


FIG. 4.3.

Since we are dealing with alternating currents, we shall change to complex currents and complex voltages. Consider a part of a circuit containing an inductance L and an a.c. ammeter and voltmeter as shown in Fig. 4.3. Let us assume that a complex current given by

$$a = Ap \tag{4–18}$$

and

$$p = \sqrt{2}\,(\sin 2\pi ft - j \cos 2\pi ft) \tag{4–19}$$

is passing through the inductance. Let us also assume that p is in phase with a, so that A is a real constant. Then a complex back e.m.f. v is developed in the inductance, and the root mean square of the real part of this complex back e.m.f. is measured by the voltmeter V. Hence, applying (4–17) to complex quantities, we have

$$v = L\,da/dt \tag{4–20}$$

which, in virtue of (4–18), gives

$$v = LA\,dp/dt \tag{4–21}$$

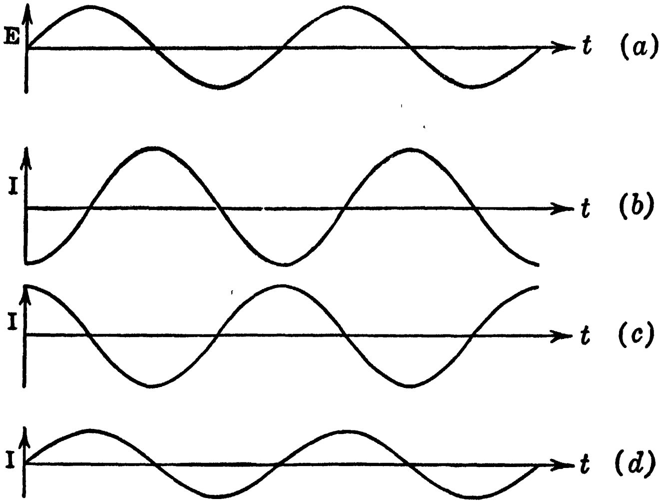


FIG. 4.4.—(a) Graph of the applied alternating voltage. Graphs of the current in (b) an inductance, (c) a capacitance, and (d) a resistance.

But, from (2–81), $dp/dt = j2\pi fp$, so that the complex back e.m.f. developed by an inductance carrying a complex current a is

$$v = j2\pi fLa \tag{4–22}$$

The magnitude of v, or the reading of the voltmeter in Fig. 4.3, is

$$V = 2\pi fLA \tag{4–23}$$

where A is the reading of the ammeter in Fig. 4.3. The presence of j on the right side of (4–22) means that the rotating vector for the voltage is 90° ahead of the rotating vector for the current.

Sometimes it is more convenient to write (4–22) in the form

$$a = -jv/2\pi fL \tag{4–24}$$

which shows that a is 90° behind v in phase. Plotting the instantaneous value of the real part of v as given by (4–13) against the time, we obtain the graph shown in Fig. 4.4*a*. If this represents the voltage across an inductance, then from (4–24) the real part of the current through the inductance is represented by the graph shown in Fig. 4.4*b*. We note that at $t = 0$ the current is a negative maximum while the voltage is zero, that at $t = 1/(4f)$ the current is zero while the voltage is a positive maximum, that at $t = 1/(2f)$ the current is a positive maximum while the voltage is zero, and so on. The phase of the current thus lags behind the phase of the voltage by 90°.

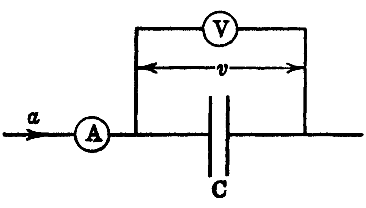


FIG. 4.5.

4.5. Capacitance.—Consider a circuit containing a capacitance C and an ammeter A as shown in Fig. 4.5. In the previous section we have learned that an inductance develops a back e.m.f. given by (4–17). When the plates of a condenser are connected to potential difference V, the plates charge up until finally there are $+Q$ units of electricity on one plate or set of plates and $-Q$ units on the other plate or set of plates. The relation between Q and V is given by

$$V = Q/C \qquad (4\text{–}25)$$

where C is the capacitance of the condenser. If Q is in coulombs and V is in volts, C is in farads. From (4–25) we may say that the back e.m.f. of a charged condenser is Q/C.

Again, as in the previous section, we change to complex quantities and we obtain

$$q/C = v \qquad (4\text{–}26)$$

where q is the complex charge on the condenser and v is the complex voltage across the condenser. However, we measure the current through the ammeter A rather than the charge on the condenser. Since this current equals the rate of change of the charge on the condenser, we may write

$$a = dq/dt \qquad (4\text{–}27)$$

Hence, differentiating both sides of (4–26) with respect to t, we obtain

$$a/C = dv/dt \qquad (4\text{–}28)$$

Now, using the principle that the applied voltage equals the back e.m.f., we have, if $v = Vp$, $a = CV\, dp/dt$. But, from (2–81), $dp/dt = j2\pi fp$, so that

$$a = j2\pi fCVp \tag{4–29}$$

In this case we have assumed that p is in phase with v, so that V is a real constant. Since $v = Vp$, (4–29) may be written in the form

$$a = j2\pi fCv \tag{4–30}$$

This equation shows that a is 90° ahead of v in phase. The graph of the real part of a plotted against the time is shown in Fig. 4.4c.

From (4–30) the complex back e.m.f. produced by a complex current flowing into a condenser is

$$v = -ja/2\pi fC \tag{4–31}$$

The complex back e.m.f. is thus 90° behind a in phase. From (4–29), the reading of the a.c ammeter in Fig. 4.5 is

$$A = 2\pi fCV \tag{4–32}$$

where V is the reading of the a.c. voltmeter.

4.6. Resistance.—According to Ohm's law, when a potential difference V is applied to a resistance R, the current i which is produced is given by

$$Ri = V \tag{4–33}$$

It is convenient to think of a current through a resistance developing a back e.m.f. given by (4–33).

For problems in alternating currents, we change to complex variables and obtain

$$v = Ra \tag{4–34}$$

For the hook-up shown in Fig. 4.6, the complex back e.m.f. produced by a complex current flowing through a resistance is given by (4–34). In this case a and v are in phase. The reading of the a.c. ammeter in Fig. 4.6, is

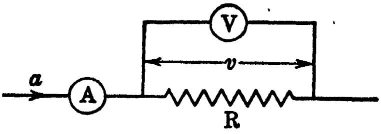


FIG. 4.6.

$$A = V/R \tag{4–35}$$

The following table of complex back e.m.f.'s will be found useful:

Inductance	$j2\pi fLa$
Capacitance	$-ja/2\pi fC$
Resistance	Ra

4.7. Impedance and Reactance.—We now consider the case represented in Fig. 4.7. Before attacking this problem we shall state the following two rules: (1) The sum of the complex back e.m.f.'s in a whole circuit equals the sum of the applied e.m.f.'s around the whole circuit; and (2) the sum of the complex back e.m.f.'s in a part of a circuit equals the complex voltage drop across that part of the circuit.

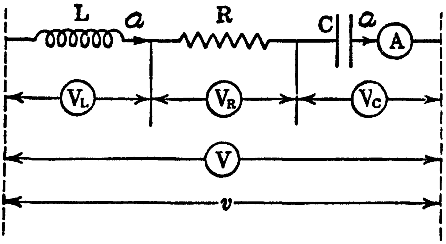


FIG. 4.7.

The sum of the complex back e.m.f's for the resistance, inductance and capacitance of Fig. 4.7 is shown on the left side of (4–36), the complex applied e.m.f. being shown on the right side.

$$aR + ja2\pi fL - ja/2\pi fC = v \tag{4–36}$$

Solving for a, we obtain

$$a = v/z \tag{4–37}$$

where

$$z = R + jX \tag{4–38}$$

and

$$X = (2\pi fL - 1/2\pi fC) \tag{4–39}$$

The complex quantity defined by (4–38) and (4–39) is called the complex impedance while the real quantity X defined by (4–39) is called the reactance. From (2–69), the magnitude of the complex impedance is

$$Z = \sqrt{R^2 + X^2} \tag{4–40}$$

The relation given in (4–37) is often called Ohm's law for alternating currents.

From (4–37) and (4–38), after rationalizing the denominator, we obtain

$$a = \frac{(R - jX)Vp}{R^2 + X^2} \tag{4–41}$$

where v has been replaced by Vp, it being assumed that p is in phase with v. By reference to Sec. 2.7 we see that a represents a rotating vector which lags behind the rotating vector p by a phase angle ϕ given by

$$\tan \phi = X/R \tag{4–42}$$

This is known as the angle of lag. The reading of the ammeter in Fig. 4.7 is the magnitude of the complex coefficient of p in (4–41), or

$$A = V/\sqrt{R^2 + X^2} \tag{4–43}$$

Next we take up the matter of the readings of the voltmeters V_R, V_L and V_C in Fig. 4.7. From the a.c. Ohm's law

$$v_R = az_R \tag{4–44}$$

where v_R is the complex voltage drop across the resistance. In the following, v_L and v_C represent the complex voltage drops across the inductance and capacitance, respectively. But z_R for a resistance R is simply R from (4–38). Hence

$$v_R = aR$$

or, from (4–41),

$$v_R = R\frac{(R - jX)V}{R^2 + X^2}p \tag{4–45}$$

The reading of the voltmeter V_R is the magnitude of the complex coefficient of p in (4–45), or

$$V_R = \frac{RV}{\sqrt{R^2 + X^2}} \tag{4–46}$$

From (4–38) and (4–39) the z_L for an inductance is $j2\pi fL$, so that the a.c. Ohm's law gives

$$v_L = az_L = j2\pi fLa$$

or, from (4–41),

$$v_L = \frac{j2\pi fL(R - jX)V}{R^2 + X^2}p$$

$$= \frac{2\pi fL(X + jR)V}{R^2 + X^2}p \tag{4–47}$$

The reading of the voltmeter V_L is the magnitude of the complex coefficient of p, or

$$V_L = \frac{2\pi fLV}{\sqrt{R^2 + X^2}} \tag{4–48}$$

The z_C for a capacitance is $-j/2\pi fC$, so that we obtain

$$v_C = \frac{(-X - jR)V}{2\pi fC(R^2 + X^2)}p \tag{4–49}$$

and the reading of the voltmeter V_C is

$$V_C = \frac{V}{2\pi fC\sqrt{R^2 + X^2}} \tag{4–50}$$

Adding (4–46), (4–48) and (4–50), we obtain

$$V_R + V_L + V_C = \frac{(R + 2\pi fL + 1/2\pi fC)V}{\sqrt{R^2 + X^2}} \tag{4–51}$$

If numerical values are used, it will be found that the coefficient of V is greater than unity unless two of the quantities R, L and $1/C$ are zero. In other words, usually

$$V_R + V_L + V_C > V$$

The reason for this is that the voltages are not in phase with each other. However, the complex voltages do give

$$v_R + v_L + v_C = v \tag{4–52}$$

From (4–45) v_R lags behind v by a phase angle ϕ_R, where

$$\tan \phi_R = X/R$$

From (4–47), v_L leads v by ϕ_L, where

$$\tan \phi_L = R/X$$

and, from (4–49), v_C lags behind v by $\pi - \tan^{-1}(R/X)$.

In virtue of the a.c. Ohm's law, we may state the following two rules: (1) The complex impedance z of complex impedances $z_1, z_2, z_3 \cdots$ in series is

$$z = z_1 + z_2 + z_3 + \cdots \tag{4-53}$$

(2) The complex impedance z of complex impedances $z_1, z_2, z_3 \cdots$ in parallel is given by

$$\frac{1}{z} = \frac{1}{z_1} + \frac{1}{z_2} + \frac{1}{z_3} + \cdots \tag{4-54}$$

4.8. Illustrative Problems.—The following problems will enable the student to obtain a firmer grasp of the method of complex quantities. Small letters such as a, v, z will represent complex numbers, except that f, the frequency, is a real number. Capital letters such as A, V, R, X represent real numbers. The magnitude of a complex number is represented by the capital of the same letter, which when small is used to represent the complex number. The letter p represents the rotating vector defined by (2–76) and (4–19).

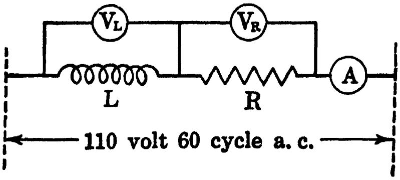


FIG. 4.8.

Problem 1. A resistance R and an inductance L are connected in series to 110-volt 60-cycle mains, as shown in Fig. 4.8. Given that $R =$ 20 ohms and $L = 70$ millihenrys (1 millihenry = 1/1000 henry), it is required to find the reading of the ammeter A and the readings of voltmeters connected across R and L respectively.

We take the complex supply voltage as $v = 110p$. In this problem $X = 2\pi fL = 2\pi \times 60 \times 0.07 = 26.3$ ohms, so that z for the whole circuit is given by

$$z = R + Xj = 20 + 26.3j \tag{4-55}$$

Hence

$$a = v/z = 110p/(20 + 26.3j) \tag{4-56}$$

On rationalizing the denominator by multiplying numerator and denominator by $(20 - 26.3j)$ as explained in Sec. 2.7, we obtain

$$a = (2.02 - 2.63j)p \tag{4-57}$$

The ammeter reading is the magnitude of the complex coefficient of p, or

$$A = \sqrt{2.02^2 + 2.63^2} = 3.33 \text{ amp}$$

Since the coefficient of j in (4–57) is not zero, the current is not in phase with the voltage. It is convenient in most problems of this type to take the position of the rotating vector $110p$ at a time $t = 1/(4f)$ when the vector is along the axis of reals as shown in Fig. 4.9 or 4.10. Then from (4–57) the line a representing the current is in the fourth quadrant of the diagram, and the current lags behind the voltage, the angle of lag being $\tan^{-1}(2.63/2.02) = 52° 40'$.

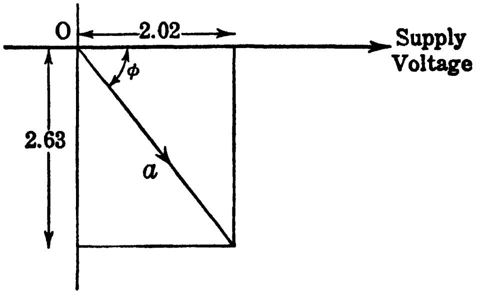


FIG. 4.9.—Geometric representation of the complex current in an a.c. circuit.

Let v_R be the complex number representing the voltage across the resistance. Then

$$v_R = az_R \tag{4-58}$$

where z_R is the complex number representing the impedance of the resistance alone. From (4–38) $z_R = 20 + 0j = 20$, so that in virtue of (4–58)

$$v_R = (2.02 - 2.63j)p \times 20 = (40.4 - 53j)p \tag{4-59}$$

The reading of the voltmeter connected across R is the magnitude of the complex coefficient of p in (4–59), or

$$V_R = \sqrt{40.4^2 + 53^2} = 66.6 \text{ volts}$$

Let v_L be the complex number representing the voltage across the inductance. Then

$$v_L = az_L$$

where z_L is the complex impedance of the inductance alone. From (4–38), $z_L = 0 + 26.3j = 26.3j$, so that

$$v_L = (2.02 - 2.63j)p \times 26.3j = (69.6 + 53j)p$$

The reading of the voltmeter connected across L is then

$$V_L = \sqrt{69.6^2 + 53^2} = 87.5 \text{ volts}$$

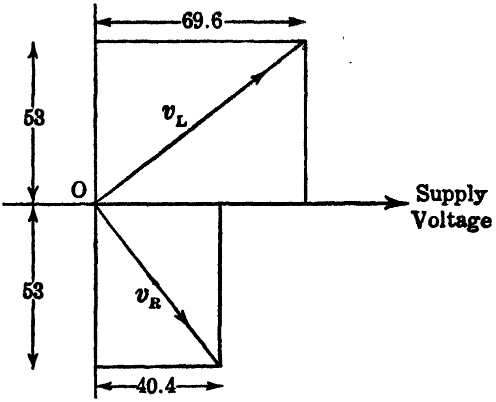


FIG. 4.10.—Geometric representation of the complex voltages across various parts of an a.c. circuit.

It is interesting to note that $V_R + V_L = 66.6 + 87.5 = 154.1$ volts, a voltage greater than the supply voltage of 110 volts. However, the two voltages V_R and V_L are neither in phase with each other nor with the supply voltage. Consider the complex numbers v_R and v_L. The geometrical representation of these at $t = 1/(4f)$ is shown in Fig. 4.10. Since the line representing v_R is in the fourth quadrant, the voltage across the resistance lags behind the supply voltage, the angle of lag being $\tan^{-1}(53/40.4) = 52° 40'$. Since the line representing v_L is in the first quadrant, the voltage across the inductance leads the supply voltage, the angle of lead being $\tan^{-1}(53/69.6) = 37° 20'$. Obviously the angle of phase difference between v_L and v_R is 90°. This is always the case when an inductance and a resistance are in series.

Although the voltmeter readings V_R and V_L do not add to give the supply voltage, yet if we add the complex numbers v_R and v_L we obtain

$$v_R + v_L = (40.4 - 53j + 69.6 + 53j)p = 110p$$

showing that the complex numbers which have been used to represent the voltages across R and L respectively do add up to give the supply voltage.

It is to be noted that if we had omitted the rotating vector p throughout this problem and had treated the problem as one in stationary vectors, we would have obtained the same answers for A, V_R and V_L. When we thus omit p we treat the supply voltage v as being entirely real. This is the same as taking a snapshot of the various rotating vectors, v, a, v_R, v_L, and v_C, at $t = 1/(4f)$. We shall omit p in the following two problems and the student may omit p in working the problems at the end of this chapter. However, it should always be remembered that the current and voltage vectors are to be understood as being multiplied by the rotating vector p.

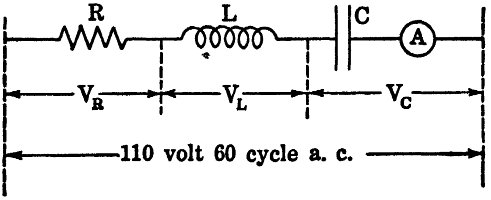


FIG. 4.11.

Problem 2. Consider the circuit of Fig. 4.11 where $R = 10$ ohms, $L = 200$ millihenrys and $C = 33$ microfarads (1 microfarad = 10^{-6} farad). Then

$$X = 2\pi fL - \frac{1}{2\pi fC} = 75.3 - 80.3 = -5 \text{ ohms}$$

so that for the whole circuit

$$z = 10 - 5j$$

Hence

$$a = \frac{110}{(10 - 5j)} = 8.8 + 4.4j$$

and the ammeter reading

$$A = \sqrt{8.8^2 + 4.4^2} = 9.83 \text{ amp}$$

If we construct a diagram for a, we see that the current leads the supply voltage, the angle of lead being $\tan^{-1}(4.4/8.8) = 26° \, 30'$. Now for the resistance, $z_R = 10$, so that

$$v_R = az_R = (8.8 + 4.4j) \times 10 = 88 + 44j$$

whence

$$V_R = \sqrt{88^2 + 44^2} = 98.4 \text{ volts}$$

For the inductance, $z_L = 75.3j$, so that

$$v_L = az_L = (8.8 + 4.4j) \times 75.3j = -331 + 662j$$

whence

$$V_L = \sqrt{331^2 + 662^2} = 740 \text{ volts}$$

For the capacitance, $z_C = -80.3j$, so that

$$v_C = az_C = (8.8 + 4.4j) \times (-80.3j) = 353 - 706j$$

whence

$$V_C = \sqrt{353^2 + 706^2} = 790 \text{ volts}$$

Here is an astounding yet correct result. The voltage across either the inductance or the capacitance is much greater than the supply voltage. However, the three voltages V_R, V_L and V_C are neither in phase with each other nor with the supply voltage. If we construct diagrams for v_R, v_L and v_C, we find that v_R leads the supply voltage by a phase angle of $\tan^{-1}(44/88) = 26° \, 30'$, that v_L leads the supply voltage by a phase angle of $\tan^{-1}(662/-331) = 116° \, 30'$, and that v_C lags behind the supply voltage by a phase angle of $\tan^{-1}(706/353) = 63° \, 30'$. The difference of phase angle between v_C and v_L is 180°. This is always the case when an inductance and a capacitance are in series. If we add the complex numbers v_R, v_L and v_C, we obtain exactly 110, the value of the supply voltage.

Problem 3. Consider the circuit of Fig. 4.12, where $R = 10$ ohms, $L = 50$ millihenrys and $C = 135$ microfarads. Then $2\pi fL = 18.8$ ohms and $-\dfrac{1}{2\pi fC} = -19.6$ ohms. The complex number representing the impedance of the capacitance alone is $-19.6j$ and that for the inductance alone is $18.8j$. The z for the impedance of these impedances in parallel is

$$\frac{1}{z} = \frac{1}{18.8j} - \frac{1}{19.6j} = \frac{1}{460j}$$

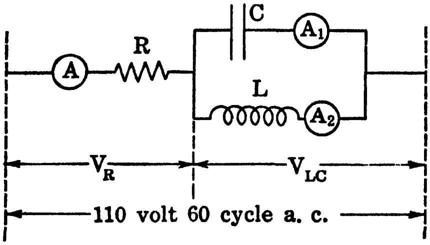


FIG. 4.12.

so that $z = 460j$. This impedance is in series with the impedance R and the complex impedance for the whole circuit is therefore $10 + 460j$. Hence the total current

$$a = \frac{110}{(10+460j)} = .005 - .24j$$

and the ammeter reading

$$A = \sqrt{.005^2 + .24^2} = 0.24 \text{ amp}$$

The total current in the circuit lags behind the supply voltage, the angle of lag being $\tan^{-1}$ (.24/.005) = 88° 48′. For the voltage across the inductance and capacitance in parallel

$$v_{LC} = (.005 - .24j) \times 460j = 110 + 2.3j$$

The complex current a_1 through the capacitance is

$$a_1 = \frac{v_{LC}}{z_C} = \frac{(110 + 2.3j)}{-19.6j} = -0.117 + 5.6j$$

and the ammeter reading

$$A_1 = \sqrt{0.117^2 + 5.6^2} = 5.6 \text{ amp}$$

The complex current a_2 through the inductance is

$$a_2 = \frac{v_{LC}}{z_L} = \frac{(110 + 2.3j)}{18.8j} = 0.122 - 5.84j$$

and the ammeter reading

$$A_2 = \sqrt{0.122^2 + 5.84^2} = 5.84 \text{ amp}$$

Again we have an astounding yet correct result. The current in each branch of the circuit is much greater than the total current through the circuit. However, the two currents A_1 and A_2 are neither in phase with each other nor with the total current A. If we construct diagrams for a_1, a_2 and a, we find that a_1 leads the supply voltage by a phase angle of $\tan^{-1} \{5.6/(-.117)\} = 91° \ 12'$ and that a_2 lags behind the supply voltage by a phase angle of $\tan^{-1} (5.85/.122) = 88° \ 48'$. The difference of phase angle between a_1 and a_2 is 180°. As the current flows to the left through A_1 in Fig. 4.12, it flows to the right through A_2. This is the meaning of the two currents differing in phase by 180°.

If we add a_1 and a_2, we obtain

$$a_1 + a_2 = -0.117 + 5.6j + 0.122 - 5.84j = .005 - .24j$$

which is the value for a.

4.9. Resonance.—If R and V in (4–43) are kept constant and the reactance is varied, the current is a maximum when the reactance is zero. The reactance in (4–38) and (4–39) is zero when

$$2\pi fL = \frac{1}{2\pi fC} \tag{4–60}$$

When f, L and C are related by (4–60), the angle of lag from (4–42) is zero, and the current is in phase with the voltage. In this case the circuit is said to be "in resonance" with the applied a.c. voltage. From (4–60)

$$f = \frac{1}{2\pi\sqrt{LC}} \tag{4–61}$$

and this is known as the natural frequency of a circuit containing an inductance and a capacitance. If R is small, the current becomes very large as L and C are varied so as to cause the circuit to approach resonance.

In the process of tuning a radio set to a certain broadcasting station, either or both L and C are varied until the natural frequency of the set equals the frequency of the broadcasting station. If, in the circuit of Fig. 4.11, the resistance is small, the capacitance is kept constant, the frequency of the applied voltage is constant, L is varied and the ammeter readings are plotted against L, the graph obtained is as in Fig. 4.13 and the reading becomes very large for the critical value of L given by (4–60). For values of L above this critical value, the current lags behind the voltage, and for values of L below the critical value, the current leads the voltage.

The peculiar result in Problem 2 of Sec. 4.8 is due to the fact that the natural frequency of the circuit in Fig. 4.11 is $1/(2\pi\sqrt{0.2 \times 0.000033}) = 61.8$ cycles, which is close to the 60-cycle frequency of the supply voltage.

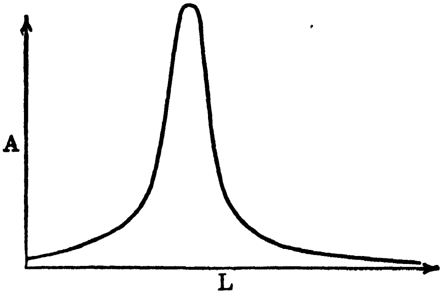


FIG. 4.13.—Variation of the current with the inductance, the capacitance being kept constant. Resonance is indicated by the sharp maximum of the curve.

The circuit is thus near to resonance with the supply voltage. Similarly, in Problem 3 the natural frequency of the closed circuit consisting of L and C in Fig. 4.12 is $1/(2\pi\sqrt{0.05 \times 0.000135}) = 61.3$, which is close to the frequency of the supply voltage. In the language of the radio engineer, the circuit consisting of L and C in Fig. 4.12 has been set oscillating with the supply of very little current (or energy) from the outside source. If the circuit consisting of L and C is tuned to exact resonance with the supply voltage, the impedance of L and C in parallel becomes infinite and no current flows through the ammeter A of Fig. 4.12. On the other hand, the currents through the ammeters A_1 and A_2 are $V/2\pi fL$ and $2\pi fCV$ respectively, where V is the supply voltage. These two currents are equal in magnitude but opposite in phase, since we have supposed exact resonance. Of course in actual practice there is some resistance in the closed circuit consisting of L and C and we then have a small current flowing through the ammeter A.

4.10 Power Factor.—The power dissipated in an a.c. circuit is RA^2 where R is the resistance of the circuit and A is the current through the circuit as measured by an a.c. ammeter. But

$$A = \frac{V}{\sqrt{R^2 + X^2}}$$

where V is the voltage across the circuit as read by an a.c. voltmeter and X is the reactance in the circuit. Hence the power

$$P = AV \times \left(\frac{R}{\sqrt{R^2 + X^2}}\right) \qquad (4\text{–}62)$$

In d.c. circuits, the power in watts is the product of the ammeter and voltmeter readings, but we see from (4–62) that this is not in general the case for a.c. circuits. The factor $R/\sqrt{R^2 + X^2}$ in (4–62) is known as the power factor. The power factor may therefore be defined as the fraction by which the product of the ammeter and voltmeter readings must be multiplied in order to give the power absorbed in an a.c. circuit. Since the angle of lag or lead is given by $\phi = \tan^{-1}(X/R)$, we see that the power factor is equal to $\cos\phi$, so that (4–62) may be written

$$P = AV\cos\phi \qquad (4\text{–}63)$$

If A and V are in amperes and volts respectively, P is in watts.

The power factor is less than unity when the current is not in phase with the voltage. If there is a difference of phase the current is flowing in the opposite direction to the voltage during part of the time. During this part of the time, energy is not being supplied by the outside source to the circuit, but is being supplied by the circuit to the outside source. A wattmeter is an instrument for measuring the net rate at which energy is being supplied by the outside source to the circuit. It should be noted that, when the power absorbed in a portion of a circuit is to be calculated, the product of the current through the portion and the voltage across the portion must be multiplied by the cosine of the angle of phase difference between the current through the portion and the voltage across the same portion. The angle of phase difference must not be taken as the angle of phase difference between the current through a portion of the circuit and the supply voltage for the whole circuit.

When an a.c. voltage is connected to a large reactance with small resistance, the phase angle is nearly 90° and $\cos\phi$ approaches zero. Thus, although there is a current in the circuit, the power absorbed is nearly zero. A current in such a circuit is called a "wattless" current.

4.11. Voltmeters, Ammeters, Wattmeters.—If an a.c. voltmeter or ammeter is examined, it will be noticed that the scales of these instruments have a peculiar appearance. The divisions on the scale become longer at the higher readings. In d.c. instruments, on the other hand, the divisions are uniformly spaced. In fact, the angular deflection of the pointer on an a.c. voltmeter or ammeter is approximately proportional to the square of the respective voltage or current; while on d.c. instruments the deflection is proportional to the first power of the respective voltage or current. Further, if a d.c. voltmeter or ammeter is connected to an a.c. circuit, the pointer shows no deflection but only a slight vibration. There is thus an essential difference between a.c. and d.c. instruments.

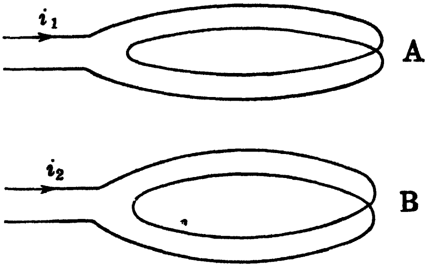


FIG. 4.14.

In one kind of a.c. instruments, the essential part consists of two coils represented as A and B, Fig. 4.14. If a steady (d.c.) current i_1 flows in A and a steady current i_2 flows in B in the same direction as in A, the two coils attract one another with a force proportional to i_1i_2. If the two currents traverse the respective coils in opposite directions, the force is a repulsion instead of an attraction. If, now, one of the coils is fixed and the other attached to a spring, the movement of the second coil will be proportional to i_1i_2 approximately. This movement is made visible by means of a pointer.

In a.c. voltmeters and ammeters the current i_1 is proportional or equal to the current i_2, so that the pointer deflection becomes approximately proportional to i_1^2, or i^2 if we drop the subscript. It will be noticed that, since the deflection is proportional to i^2, the deflection does not change with the sign of i. Hence an alternating current will cause a deflection. The deflection in the case of a slowly alternating current would vary from zero to a maximum proportional to i_m^2. However, with ordinary 60-cycle a.c. the alternations are too rapid for the pointer to follow, and we obtain a deflection which is proportional to the average square of the current.

In a wattmeter the coil *A* is connected as a voltmeter across the circuit and coil *B* is connected as an ammeter in the circuit. In this case we have a current i_1 through *A* which is proportional to the voltage and a current i_2 through *B* which is proportional to (or is) the current in the circuit. In this case, the two currents may not always be flowing in the same direction in the two coils, and in fact do not always flow in the same direction when an a.c. current is out of phase with the voltage. In the case where the angle of lag or lead ϕ is 90°, the current i_2 is flowing in the same direction as i_1 during one half of a cycle and in the opposite direction during the other half of the cycle. In the first half cycle, there is attraction between the coils, and, in the second, repulsion, so that on the average the force is zero and the pointer reads zero. This is the case when we have the so-called "wattless" current. When ϕ is between 0° and 90°, the average attraction is greater than the average repulsion and so there is a net attraction and the pointer gives a reading.

4.12. Mutual Inductance.—Consider the two circuits shown in Fig. 4.15. When the two coils L_1 and L_2 are close enough for some or all of the magnetic flux from one coil to pass through the other coil, the two coils are said to constitute a mutual inductance *M*. When the current in circuit 1 varies at the rate of di_1/dt there is induced in L_2 an e.m.f. Mdi_1/dt. Likewise when the current in circuit 2 varies at the rate of di_2/dt there is induced in L_1 an e.m.f. Mdi_2/dt. In terms of complex quantities the complex back e.m.f. in circuit 1 due to a complex current a_2 in circuit 2 is $j2\pi fMa_2$, and also the complex back e.m.f. in circuit 2 due to a complex current a_1 in circuit 1 is $j2\pi fMa_1$.

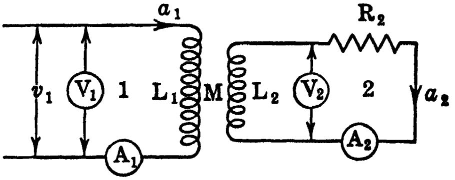


FIG. 4.15.

Now consider the case where $v_1 = V_1p$. For circuit 1 the sum of the back e.m.f.'s is given on the left side of (4–64).

$$ja_1 2\pi fL_1 + ja_2 2\pi fM = V_1 p \tag{4–64}$$

and for circuit 2 the sum of the back e.m.f.'s is given on the left side of (4–65)

$$a_2R_2 + ja_2 2\pi f L_2 + ja_1 2\pi f M = 0 \qquad (4\text{–}65)$$

The right side of (4–65) is zero because no external e.m.f. is applied to the circuit. The above equations are simultaneous linear equations in a_1 and a_2 and may be solved in the ordinary way even though a_1 and a_2 are complex quantities. Eliminating a_1, we obtain

$$a_2\{-L_1R_2 + j2\pi f(M^2 - L_1L_2)\} = MV_1p \qquad (4\text{–}66)$$

Now, the flux produced by a current through coil L_1 does not all pass through coil L_2. Hence we write

$$M^2 = k^2L_1L_2 \qquad (4\text{–}67)$$

where k is the "coupling coefficient." For "close coupling" k approaches unity. We shall examine the case when $k = 1$. Then (4–66) reduces to

$$a_2 = -(\sqrt{L_2/L_1})V_1p/R_2 \qquad (4\text{–}68)$$

Hence a current a_2 flows in circuit 2 as if a voltage

$$v_2 = -(\sqrt{L_2/L_1})\ V_1p$$

were applied to the resistance R_2 and ammeter A_2 in series. The inductances L_1 and L_2 and the mutual inductance M constitute a transformer, and we have

$$v_2/v_1 = -\sqrt{L_2/L_1} \qquad (4\text{–}69)$$

The reading of the voltmeter in Fig. 4.15 is given by

$$V_2/V_1 = \sqrt{L_2/L_1} \qquad (4\text{–}70)$$

The voltages V_1 and V_2 are known as the primary and secondary voltages, respectively. In virtue of (4–68), the secondary current is 180° out of phase with the primary voltage.

We have seen in Sec. 4.4 that the inductance L of a coil is very nearly proportional to n^2. Hence for coils wound upon the same bobbin, one coil over the other, L_2/L_1 very neary equals n_2^2/n_1^2, so that

$$V_2/V_1 = n_2/n_1 \qquad (4\text{–}71)$$

where n_1 and n_2 are the number of turns in the primary and secondary coils respectively. This is the well-known equation for the transformer.

From (4–68) and (4–70) the reading of the ammeter A_2 is

$$A_2 = V_2/R_2 \tag{4–72}$$

The current through the resistance R_2 is in phase with the voltage across R_2 and the power dissipated in the resistance is A_2V_2, since the power factor is unity. Eliminating a_2 from (4–65) by means of (4–68), we obtain

$$a_1 = \frac{(2\pi fL_2 - jR_2)V_1}{2\pi fL_1R_2}p \tag{4–73}$$

in virtue of (4–67). Hence the ammeter A_1 reads

$$A_1 = \frac{\sqrt{(2\pi fL_2)^2 + R_2^2}}{2\pi fL_1R_2}V_1 \tag{4–74}$$

and the power factor in circuit 1 is

$$\cos\phi_1 = 2\pi fL_2/\sqrt{(2\pi fL_2)^2 + R_2^2} \tag{4–75}$$

The power in circuit 1 is

$$A_1V_1\cos\phi_1 = V_1^2L_2/L_1R_2 \tag{4–76}$$

From (4–70), the right side of (4–76) is V_2^2/R_2 or A_2V_2, the power dissipated in circuit 2. We thus see that the power dissipated in circuit 2 equals the power supplied to the transformer in circuit 1.* If the resistance R_2 is made infinite, or in other words if a switch in circuit 2 is opened, A_2 becomes zero from (4–72) but from (4–74)

$$A_1 = V_1/2\pi fL_1 \tag{4–77}$$

so that, although no current is flowing in the secondary, there is a current flowing in the primary of the transformer. However, the power factor in circuit 1 is $\cos\phi = 0$ from (4–75) and the power supplied to the primary is zero. The current given by (4–77) is an example of a wattless current. This current is sometimes known as the "magnetizing current" of the transformer. This term is particularly apt when the transformer contains iron. A transformer containing iron behaves approximately but not exactly according to the theory given in this section.

If we assume that R_2 is negligible with respect to $2\pi fL_2$, as it is in most transformers, we obtain

$$a_2/a_1 = -\sqrt{L_1/L_2} \tag{4–78}$$

after dividing (4–68) by (4–73). Now, if we put $v_1/a_1 = z_1$ and $v_2/a_2 = z_2$, we find that

$$z_1/z_2 = L_1/L_2 = n_1^2/n_2^2 \tag{4–79}$$

From (4–79) we see that, if an impedance z_2 is connected in series to the secondary circuit of a transformer, the effect on the current in the primary circuit is the same as if an impedance z_1, as given by (4–79), had been connected in series to the primary circuit. The impedance z_2 is said to be "reflected" into the primary circuit as an impedance z_1. The impedance z_1 is called the "reflected" impedance. Impedances may be reflected from the secondary into the primary circuit or vice versa. The impedances z_1 and z_2 are external to the transformer inductances L_1 and L_2. The rule described by (4–79) is useful in problems involving transformers.

CHAPTER IV

PROBLEMS

1. A resistance of 20 ohms, a capacitance of 150 microfarads, and an a.c ammeter are connected in series to a.c. mains supplying 110 volts at 60 cycles Find the current registered on the ammeter and determine the readings of a.c. voltmeters connected across the resistance and the capacitance respectively. Also find the current when 110 volts d.c. is applied to the resistance and the capacitance in series.

2. A resistance of 30 ohms, an inductance of 5 millihenrys, and an a.c. ammeter are connected in series to an a.c. generator giving 220 volts at 500 cycles. Find the reading of the ammeter and determine the readings of a.c. voltmeters connected across the resistance and the inductance respectively. Also find the current when 220 volts d.c. is applied to the resistance and inductance in series.

3. A resistance of 10 ohms, an inductance of 35 millihenrys, a capacitance of 200 microfarads, and an a.c. ammeter are connected in series to a.c. mains supplying 110 volts at 60 cycles. Find the reading of the ammeter and determine the readings of a.c. voltmeters connected across the resistance, the inductance, and the capacitance respectively. Also find the current when 110 volts d.c. is applied to the resistance, inductance, and capacitance in series.

4. A resistance of 20 ohms and an inductance of 70 millihenrys are connected in parallel to a.c. mains supplying 110 volts at 50 cycles. Ammeters are connected so as to register the currents through the resistance, through the inductance, and through the resistance and inductance in parallel. Determine the readings of the respective ammeters.

5. A resistance of 5 ohms and an inductance of 50 millihenrys are connected in series. These are then connected in parallel with a capacitance of 150 microfarads. The whole is now connected to a.c. mains supplying 220 volts at 60 cycles. Ammeters are connected so as to register the currents through the whole circuit and through the parts of the circuit. Determine the readings of the respective ammeters. Also find the readings of the respective ammeters when 220 volts d.c. is applied to the circuit.

6. A resistance of 0.5 ohm, a capacitance of 300 microfarads, an a.c. ammeter, and a variable inductance are connected in series to a.c. mains supplying 110 volts at 500 cycles. Plot a graph of the current against the inductance when the inductance varies from 200 to 500 microhenrys. Also plot graphs of the angle of lag against the inductance and of the power absorbed in the circuit against the inductance.

7. A capacitance of 100 microfarads and an a.c. ammeter are connected in series to a.c. mains supplying 110 volts at 60 cycles. Find the current registered by the ammeter and determine the power consumed. If a resistance of 5 ohms is added in series with the capacitance, find the new values of the current and of the power consumed.

8. Consider the points, *A*, *B*, *C*, *D* arranged in the form of a four-sided figure *ABCD*. There is a resistance of 20 ohms between *A* and *B*, an inductance of 40 millihenrys between *B* and *C*, an inductance of 50 millihenrys between *A* and *D*, and a resistance of 30 ohms between *D* and *C*. The points *A* and *C* are connected to a.c. mains supplying 110 volts at 60 cycles and an a.c. voltmeter is connected between *B* and *D*. Determine the reading of the voltmeter. Hint: find v_{AB} and v_{AD}, then $v_{BD} = v_{AB} - v_{AD}$.

9. In the hook-up of Problem 8, *AB* contains a resistance of 20 ohms, *BC* a capacitance of 75 microfarads, *AD* an inductance of 95 millihenrys, and *DC* a resistance of 20 ohms. Determine the reading of the voltmeter between *B* and *D*.

10. In the hook-up of Fig. 4.15, Sec. 4.12, the a.c. supply is 1200 volts 60 cycles, and the inductances L_1 and L_2 are 7200 and 50 millihenrys, respectively. Determine the reading of the ammeter A_1, the power factor and the power in circuit 1 when R_2 is (*a*) infinity, and (*b*) 20 ohms. The coupling between L_1 and L_2 is close. Also determine the reading of the ammeter A_2 in case (*b*).

11. A resistance of 50,000 ohms is connected in series to the secondary circuit of a transformer whose turn ratio (n_2/n_1) is 100. Determine the reflected resistance in the primary.

CHAPTER V

ELECTROMAGNETIC THEORY OF RADIATION

5.1. Electrostatic and Magnetic Fields.—The concept of the "field" surrounding an electrically charged body or a magnetized body is due to Faraday. According to Faraday the space surrounding a charged or magnetized body is different in some way from ordinary space. This particular space Faraday called the "field." Faraday went further and ascribed mechanical properties to the field. The experiment in which iron filings are sprinkled on paper above a magnet suggested to Faraday the concept of lines of force, and further suggested that these lines of force represented lines of strain in the field. Because his mind required some mechanism to produce the observed action of one charge upon another or of one magnet upon another, Faraday supposed a medium to exist between the charged bodies and between the magnets. This hypothetical medium is called the "ether." Faraday postulated that the ether in the vicinity of charged or magnetized bodies is in a state of strain and that the observed effects on the charged bodies and magnets are results of the strain in the ether. To Faraday's mechanical type of mind, the idea of action at a distance was abhorrent.

Maxwell adopted Faraday's ideas and put them into mathematical form. The theory which we shall describe in the present chapter is essentially the Maxwell, or perhaps more correctly the Faraday-Maxwell, electromagnetic theory of radiation. Nowadays, this theory is frequently referred to as the classical theory as distinct from the more modern quantum theory.

5.2. Lines of Force.—The reader is already familiar with lines of force. A line of force in an electric or magnetic field is a line so drawn that the direction of the tangent at every point on the line gives the direction of the electric or magnetic force respectively at that point. The electric field strength R at a point P is defined as the ratio of the small force F acting on a small positive charge q placed at the point P to the charge q. Thus,

$$R = \frac{F}{q}, \tag{5-1}$$

The direction of the electric field at P is the direction of the force F acting on the charge q. Likewise, the magnetic field strength H at a point P is defined as the ratio of the force F acting on a north magnetic pole of strength m placed at the point P to the pole strength m. Thus,

$$H = \frac{F}{m} \tag{5-2}$$

The direction of the magnetic field at P is the direction of the force F acting on the pole m. It is essential that both the charge q and the pole strength m shall be so small that the field which is being measured is not altered to any noticeable extent by the presence of q or m as the case may be.

The reader has already learned that the lines of force of an electric field leave a conductor in a direction at right angles to the surface of the conductor. These lines of force are less crowded together in parts of the field where the field strength is small and are more crowded together in parts of the field where the field strength is large. There is thus a correlation between the crowding together or the density of the lines of force and the field strength. In a uniform electric field of strength R the number of lines of force threading their way through an area A taken perpendicular to the direction of the field is proportional to A. The number of lines of force threading their way through an area is known as the flux through that area. The convention is adopted that, if the strength of the uniform electric field is R, the flux N through the area A is given by

$$N = RA \tag{5-3}$$

This relation is true in the case of a non-uniform field if the area A is taken so small that the field strength does not change by a sensible amount from one point to another in the area. If in this case we represent the small area by dA and the small number of lines of force by dN, we have the more general relation

$$R = \frac{dN}{dA}$$

But dN/dA is the number of lines of force per unit area, or, as it is called, the flux density, so that the electric field strength at a point is equal to the flux density at that point.

Let us consider a charge Q at the point O. Draw a sphere of radius

a about the point O. The field strength at any point on the sphere is then

$$R = \frac{Q}{a^2} \tag{5–4}$$

Since the total area of the sphere is $4\pi a^2$, the flux across the sphere is by (5–3)

$$N = \frac{4\pi a^2 \times Q}{a^2} = 4\pi Q \tag{5–5}$$

Although the relation (5–5) between N and Q has been proved for a special case, the relation is generally true, so that the total flux from a charge Q is $4\pi Q$ lines.

Similarly in magnetism the number of magnetic lines of force or magnetic flux across an area A taken at right angles to a uniform magnetic field of strength H is given by

$$N = HA \tag{5–6}$$

where N is the magnetic flux. If the magnetic field is non-uniform we have, as in the corresponding electrical case,

$$H = \frac{dN}{dA}$$

so that the magnetic field strength is equal to the magnetic flux density. Also, corresponding to the electrical case, the total magnetic flux from a magnetic pole of strength m is $4\pi m$.

The difference of electrical potential between the points A and B is defined as the ratio of the work done by the electric force of the field, or, briefly, by the field, when a small positive charge q is moved from A to B, to the charge q, so that

$$V_A - V_B = \frac{W}{q} \tag{5–7}$$

where V_A and V_B are the electric potentials at A and B respectively. It is again essential that the charge q shall be so small that any change in the electric field due to the presence of q is negligible. It is now necessary to take account of the algebraic signs of the quantities in (5–7). If the work done by the field is positive when the charge moves from A to B, then $V_A - V_B$ is positive. In particular, if A and B are on the same line of force and are close together, and if the direction of the electric field is from A to B, the work done by the field is positive. Since, in this case, $V_A - V_B$ is positive, V_B is less than V_A so that the potential at a point P decreases as the point P moves along a line of force in the direction of the field. The positive direction of a line of force is the direction in which a small positive charge is impelled to move along the line of force. Now, it is a convention of the calculus that a small increase in a quantity V is represented by $+dV$, while a small decrease is represented by $-dV$. When A and B are on the same line of force and are close together, the work done by the field is given by $Rqds$, where ds is the length of AB, so that (5–7) becomes

$$-dV = \frac{Rqds}{q} = Rds \tag{5–8}$$

where $-dV$ is the decrease of the potential in going from A to B. Hence, we may write

$$R = -\frac{dV}{ds} \tag{5–9}$$

We therefore say that the electric field strength is equal to minus the rate of increase of potential with respect to the distance along a line of force. Because difference of potential in the practical system of units is measured in volts, electric field strength is frequently measured in volts per centimeter, and the direction of the field is in the direction of the greatest rate of fall of potential.

If the point B is at infinity, it is usually assumed that $V_B = 0$, and we then speak of the potential at the point A. In words, the potential at the point A is defined as the ratio of the work done by the field when a small positive charge q is moved from A to infinity to the charge q.

The electric field strength at a point P is sometimes defined as the force acting on a unit positive charge placed at P, and, likewise, the difference of potential between points A and B as the work done by the field when a unit positive charge is moved from A to B, and the potential at a point A as the work done by the electric force of the field when unit positive charge is moved from A to infinity. These definitions are less exact than the definitions given in the previous paragraphs.

5.3. Capacitance of a Parallel Plate Condenser.—Fig 5.1 represents a condenser each of whose plates has an area A, one plate having a charge $+Q$ and the other a charge $-Q$. Excepting near the edges of the plates, the lines of force proceed straight across from the positive to the negative plate. If the linear dimensions of each plate are large compared with the distance s between the plates, the effect of the edges may be neglected. Since the charge on the positive plate is Q and the area of the plate is A, the number of lines of force leaving the positive plate is

$$N = 4\pi Q \tag{5-10}$$

Since the lines of force between the plates are straight, the number of lines crossing unit area taken parallel to either plate at a point P between the plates is $4\pi Q/A$, so that the field strength between the plates is

$$R = \frac{4\pi Q}{A} \tag{5-11}$$

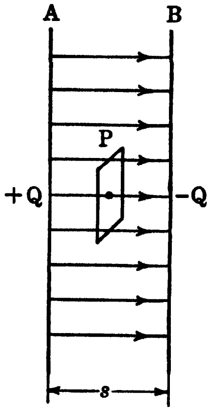


FIG 5.1.—Field between the plates of a parallel plate condenser.

The work done by the electric force of the field when a small charge $+q$ is moved from the positive plate to the negative plate is Rqs, where s is the distance from plate A to plate B. From (5-7) the potential between A and B is therefore

$$V = \frac{Rqs}{q} = Rs \tag{5-12}$$

From (5-11) and (5-12)

$$V = \frac{4\pi Qs}{A} \tag{5-13}$$

But capacitance is defined by

$$C = \frac{Q}{V} \tag{5-14}$$

so that

$$C = \frac{A}{4\pi s} \tag{5-15}$$

This is the formula for the capacitance of a parallel plate condenser.

5.4. Energy of a Charged Condenser.—The energy of a charged condenser is the total amount of work done in developing a charge $-Q$

on the negative plate and a charge $+Q$ on the positive plate. If the condenser is discharged by means of discharging tongs, the energy appears as heat in the spark. Let us imagine that we begin with the condenser of Fig. 5.1 totally discharged. Now suppose that a very small charge $+dq$ is removed from plate B and transferred to plate A. After the transfer A will have a charge $+dq$ and B a charge $-dq$. Next, suppose that a second charge dq is transferred. This charge will be in a field during the transfer, and work must be done against the electric force of the field. A third charge $+dq$ is then transferred and more work is done against the field. This process is repeated many times until there is a charge $+q$ on the positive plate and a charge $-q$ on the negative plate. The potential between the two plates is now q/C. When the next charge dq is transferred from B to A, Fig. 5.1, the work done against the field will be

$$dW = \left(\frac{q}{C}\right) dq \tag{5–16}$$

Hence the total work done by the transfer of successive charges $+dq$ from B to A is the sum of the quantities dW given by (5–16) where q increases with each transfer. Therefore the total work W done in charging the condenser from zero to $+Q$ on A and $-Q$ on B is

$$W = \int_0^Q \left(\frac{q}{C}\right) dq = (\tfrac{1}{2}) \left[\frac{q^2}{C}\right]_0^Q = (\tfrac{1}{2}) \frac{Q^2}{C} \tag{5–17}$$

In virtue of (5–14) the energy may also be written

$$W = (\tfrac{1}{2})QV \quad \text{and} \quad W = (\tfrac{1}{2})CV^2 \tag{5–18}$$

The formulas (5–17) and (5–18) for the energy of a charged condenser are valid for a condenser of any shape.

5.5. Elastic Properties of the Medium.—According to Faraday and Maxwell, the energy of charged bodies resides in the medium about them. The energy of a parallel plate condenser is, from (5–15) and (5–17),

$$W = (\tfrac{1}{2}) 4\pi s \frac{Q^2}{A} \tag{5–19}$$

Eliminating Q by means of (5–11), we have

$$W = \left(\frac{R^2}{8\pi}\right) sA \tag{5–20}$$

But sA is the volume of the medium between the plates of the condenser, so that the energy per unit volume of the medium is $R^2/8\pi$. This expression for the energy per unit volume at a point is also true for the medium surrounding any system of charged bodies.

Imagine the two plates of a parallel plate condenser to be initially very close together. The capacitance is then very large and the energy is so small that we may take it to be zero. Let the charge on the condenser be Q. The plates are then pulled apart to a distance s. If the attractive force of one plate on another is F, the work done will be Fs. But the energy of a charged condenser is $(\frac{1}{2})Q^2/C$, so that

$$Fs = (\tfrac{1}{2})\frac{Q^2}{C} \tag{5-21}$$

Substituting from (5-15), we obtain

$$F = 2\pi\frac{Q^2}{A} = 2\pi\left(\frac{Q^2}{A^2}\right)A \tag{5-22}$$

The pull per unit area on each plate is F/A, so that the pull per unit area on a plate is

$$T = 2\pi\frac{Q^2}{A^2} \tag{5-23}$$

According to Faraday and Maxwell, the pull of one plate on the other is transmitted through the medium. In the present case the attractive force between the two condenser plates is the result of a tension along the lines of force. On this theory the tension across unit area taken perpendicular to the lines of force is given by (5-23). Eliminating Q/A from (5-23) by means of (5-11), we obtain

$$T = \frac{R^2}{8\pi} \tag{5-24}$$

Although this formula has been developed for the particular case of the medium between the plates of a parallel plate condenser, it is generally true. The tension per unit area along the lines of force at a point P in the medium is given by (5-24) when R is the electric field strength at P.

Faraday and Maxwell also supposed that the lines of force exert a sideways pressure against each other. It can be shown that the magnitude of this sideways pressure across unit area taken parallel to the field is $R^2/8\pi$.

In the magnetic field, we have similar formulas for the energy per unit volume of the medium, the tension along the lines of force, and the sideways pressure of the lines of force. In the magnetic formulas R is replaced by H. It should be remembered that the above formulas are for the case where the medium is a vacuum.

5.6. Electric Field Between Coaxial Cylinders.—Consider two coaxial conducting cylinders with their common axis perpendicular to the plane of the paper as shown in Fig. 5.2. Let a and b be the radii of the inner and outer cylinders, respectively. Let the cylinders extend to infinity above and below the plane of the paper and let the charges per unit length be $-q$ on the outer surface of the inner cylinder and $+q$ on the inner surface of the outer cylinder, with a difference of potential V between the cylinders. From symmetry, the lines of force extending from the inner to the outer sphere are in or parallel to the plane of Fig. 5.2 and they diverge radially as if coming from O, which is the point where the axis of the cylinders cuts the plane of the figure. Consider the space enclosed between the two conducting cylinders and two planes perpendicular to the axis and at unit distance apart. Since the charge on the inner cylinder between the planes is q, the electric flux from this charge is $4\pi q$. This flux passes through an imaginary cylinder of radius r, where $a < r < b$. The part of this cylinder between the planes has an area $2\pi r$, so that the flux density at a point P on the imaginary cylinder is $4\pi q/2\pi r = 2q/r$. Hence the electric field strength R is given by

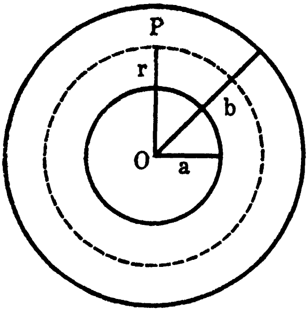


FIG. 5.2.

$$R = 2q/r \tag{5-25}$$

This is the force per unit charge on a small charge placed at P. It the small charge is moved outward through a distance dr the work done by the field per unit charge is

$$dV = (2q/r)dr \tag{5-26}$$

Integrating this from $r = a$ to $r = b$, we have

$$\begin{aligned} V &= 2q\int_a^b dr/r = 2q\Big[\log r\Big]_a^b \\ &= 2q\log(b/a) \end{aligned} \tag{5-27}$$

The capacitance per unit length of the coaxial condenser is

$$C = q/V = 1/[2 \log (b/a)] \tag{5-28}$$

If V_r is the difference of potential between a point at $r = r$ and a point at $r = a$, integration of (5–26) gives

$$V_r = 2q \log (r/a) \tag{5-29}$$

Elimination of q from (5–25) and (5–27) gives

$$R = V/[r \log (b/a)] \tag{5-30}$$

while elimination of q from (5–27) and (5–29) gives

$$V_r = \frac{V}{\log (b/a)} \cdot \log (r/a) \tag{5-31}$$

5.7. Electromagnetic Induction.—In Chapter IV, we have discussed Faraday's law of electromagnetic induction. Suppose that a wire is bent in the form *ABCD*, Fig. 5.3, and that a wire *FG* makes sliding contact with *ABCD* as shown.

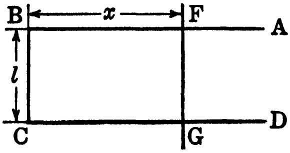


FIG. 5.3.—Electromotive force induced in a wire moving in a magnetic field.

Let the plane of *FBCG* be perpendicular to a uniform magnetic field of strength H. Then, the flux through *FBCG* is Hxl, where x is the length of *BF* and of *CG* and l is the length of *BC*. If *FG* slides with a velocity v along *BA* and *CD* in a way such that it is always parallel to *BC*, the flux through *FBCG* varies and we have the induced e.m.f.

$$E = -\frac{d}{dt}(Hlx) = -Hl\frac{dx}{dt} = -Hlv \tag{5-32}$$

since H and l are constants and $dx/dt = v$. If the direction of the magnetic field is from the reader into the paper, the induced e.m.f. is counterclockwise around *FBCG*. We may say that an e.m.f. is induced in *FG* because it is cutting the lines of force of the magnetic field and that this induced e.m.f. is given by (5–32), where v is the velocity of the wire sideways across the field and l the length of the wire. This e.m.f. may be considered as being between the two ends of the wire F and G. This means that the work done in taking unit charge from F to G is E. If the electric field strength in the wire is R, then $E = Rl$

and $Rl = Hlv$, whence, omitting the minus sign,

$$R = Hv \tag{5-33}$$

In the present theory it is supposed that, even if an imaginary line is thought to cut across the magnetic lines of force, there will be an e.m.f. of magnitude Hlv between the two ends of the line and an electric field produced along the line of amount given by (5–33). As v is the relative velocity of the line across the field, it matters not whether the line or the field is moving. Hence we may generalize and say that an electric field is produced at each point of a moving magnetic field. It should be remarked that the quantities R and H in (5–33) are measured in electromagnetic units.

The question now arises as to whether a moving electric field will produce a magnetic field. We have noted in Sec. 4.1 that a current of i emu flowing along a long straight wire produces a magnetic field of strength $2i/r$ at a point P which is at a distance r from the wire. But a current is simply a series of charges, q, traveling along the wire. Each of these charges, q, sends out $4\pi q$ lines of force which travel with it. These electric lines of force move past the point P and produce a magnetic field. Hence we may again generalize and say that a magnetic field is produced at each point of a moving electric field. It can be shown that an electric field moving with velocity v produces a magnetic field of strength H at a stationary point P such that

$$H = Rv \tag{5-34}$$

where R is the strength of the electric field. The quantities H and R are here measured in electrostatic units.

It should be noted that (5–33) and (5–34) are reciprocal relations. However, the quantities in (5–33) are in emu, while those in (5–34) are in esu. Now, it is usual to measure magnetic field strength in emu and to measure electric field strength in esu. It has been found that 1 esu of charge equals $1/c$ emu of charge, and that 1 esu of potential equals c emu of potential, where c is a certain number. As a consequence of these relations between the two systems of units, it can be shown that

$$cR \text{ (in esu)} = R \text{ (in emu)} \tag{5-35}$$

and

$$cH \text{ (in emu)} = H \text{ (in esu)} \tag{5-36}$$

Let us distinguish quantities measured in esu from those measured in emu by adding a prime (′) to the former. Thus, R' signifies that R is measured in esu. Then (5–33) and (5–34) become

$$cR' = Hv \text{ (magnetic field moving)} \tag{5–37}$$

$$cH = R'v \text{ (electric field moving)} \tag{5–38}$$

In these, c, the ratio of the electromagnetic unit of charge to the electrostatic unit of charge, has been found experimentally to have a numerical value of 3×10^{10}.

The electrical engineers call the esu of current the statampere, the "stat" being taken from "electrostatic." Similarly, they use the terms statvolt and statohm. We have the following relations:

$$1 \text{ statampere} = (1/c) \text{ abampere} = 3 \times 10^{-9} \text{ ampere} \tag{5–39}$$

$$1 \text{ statvolt} = c \text{ abvolts} = 300 \text{ volts} \tag{5–40}$$

The relation (5–40) is useful and should be remembered.

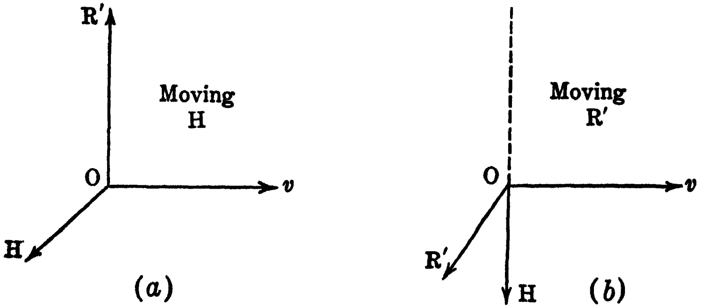


FIG. 5.4.—(a) Direction of the electric field produced by a moving magnetic field. (b) Direction of the magnetic field produced by a moving electric field.

The relations between the directions of R' and H are shown in Figs. 5.4*a* and 5.4*b*. The first figure represents the case where an electric field R' is produced by the motion of a magnetic field H which is moving with a velocity v in a direction at right angles to itself. The second figure represents the case where a magnetic field H is produced by the motion of an electric field R' which is moving with velocity v in a direction at right angles to itself.

Referring to Sec. 2.7, it is seen that the multiplication of a quantity a by j, the symbol for $\sqrt{-1}$, has the effect of rotating the direction of the line representing a through a right angle. The field R' which is produced by the moving field H is at right angles to H as shown in Fig. 5.4*a*. Suppose we look along the axis of v in this figure; then a rotation from OH to OR' in the plane perpendicular to Ov requires a turn through a right angle in the counterclockwise direction. A rotation in the counterclockwise direction is represented by multiplication by j, while a rotation in the clockwise direction is represented by multiplication by $-j$. Hence (5–37) and (5–38) are respectively replaced by

$$R' = \left(\frac{Hv}{c}\right) j \text{ (magnetic field moving)} \qquad (5\text{–}41)$$

and

$$H = -\left(\frac{R'v}{c}\right) j \text{ (electric field moving)} \qquad (5\text{–}42)$$

5.8. Dielectric Constant and Permeability.—So far we have considered a vacuum as the medium surrounding the charged bodies or magnets. Now, however, we shall consider media other than a vacuum. If we have two similarly shaped condensers of equal size, the first condenser having a vacuum as the dielectric and the second having some other substance as the dielectric, the ratio of the capacitance of the second condenser to that of the first is called the dielectric constant of the substance and is represented by K. In a medium of dielectric constant K the relation (5–3) becomes

$$D' = KR' \qquad (5\text{–}43)$$

where D', the electric displacement, is written for N'/A and is therefore the electric flux density (number of electric lines of force per unit area). In a medium of permeability μ the relation (5–6) becomes

$$B = \mu H \qquad (5\text{–}44)$$

where B, the magnetic induction, is written for N/A and is therefore the magnetic flux density (number of magnetic lines of force per unit area).

In media where K and μ are not unity, the formulas for the energy, tension and pressure in the medium become as shown in Table 5.1.

Formulas (5–41) and (5–42) now become

$$cR' = \mu Hvj \text{ (magnetic field moving)} \qquad (5\text{–}45)$$

$$cH = -KR'vj \text{ (electric field moving)} \qquad (5\text{–}46)$$

TABLE 5.1

	Electric	Magnetic
Energy per unit volume....................	$KR'^2/8\pi$	$\mu H^2/8\pi$
Tension across unit area	$KR'^2/8\pi$	$\mu H^2/8\pi$
Sideways pressure	$KR'^2/8\pi$	$\mu H^2/8\pi$

Now let us suppose that the electric field is produced by a moving magnetic field. In the present theory it is supposed that this electric field moves with the original magnetic field, and therefore produces a magnetic field. If the magnetic field so produced equals the original magnetic field, this field is moving with such a speed as to maintain itself. In this case the R' in (5–46) is given by (5–45), so that

$$H = -K\left(\frac{\mu Hv}{c}\right)\cdot\left(\frac{v}{c}\right)j^2 = \frac{\mu KHv^2}{c^2}$$

Cancelling out H and taking the square root, we have

$$v = \pm\frac{c}{\sqrt{\mu K}} \tag{5–47}$$

Hence, if the magnetic field moves with a velocity given by (5–47), it produces an electric field which moves with it, and this produces the magnetic field, and so the field keeps moving automatically. The field maintains itself. The meaning of the ± sign in (5–47) is that the field may move either forward or backward with the velocity v. We have noted in Sec. 4.2 that the physical dimensions of $1/\sqrt{\mu K}$ are those of a velocity. The number c is the numerical value of the velocity of electromagnetic waves in a vacuum.

5.9. Velocity of Electromagnetic Waves.—It is a characteristic of wave motion that the form of the wave is maintained when the wave travels with a particular velocity, which is known as the velocity of propagation. For example, consider a long string of mass m per unit length and under a tension T. The string will set itself along a straight line between its two ends A and B, as in Fig. 5.5. Now suppose that the end A is moved quickly to C (Fig. 5.5*a*) and back. A disturbance or wave form ADE will begin to move along the string as in Fig. 5.5*b*. Some time later the disturbance will have traveled to FGH as in Fig. 5.5*c*. The form of the disturbance is similar to the form ADE and the velocity of propagation is such that the wave form remains unchanged.

In the case of sound waves in air, a vibrating tuning fork produces a series of compressions and rarefactions, which are propagated through the air with a velocity such that the wave form remains constant. In this case the size of the variations of the pressure of the air from the normal pressure decreases as the distance of the wave from the tuning fork increases; but, if a given condensation and the succeeding rarefaction are represented by a curve, the general shape of the curve and therefore the wave form will remain constant as the condensation and rarefaction travel out from the tuning fork. In Sec. 3.4, we have seen that, if a rope is stretched with tension T and it has a mass m per unit length, the velocity which is necessary in order for a given wave form to maintain itself is

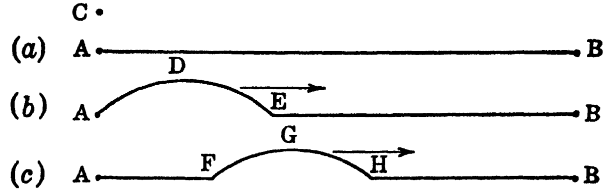


FIG. 5.5.—Propagation of a wave along a stretched string.

$$v = \sqrt{\frac{T}{m}}$$

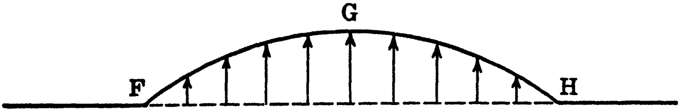


FIG. 5.6.—Propagation of displacement along a stretched string.

Again, considering the stretched string in Fig. 5.5, the wave *FGH* may be considered as a series of displacements as represented by the arrows in Fig. 5.6, the length of each arrow denoting the displacement. We may think of these arrows as traveling sideways with a velocity $v = \sqrt{T/m}$.

Now, let us consider a bundle of electric lines of force which leave a charged body, such as an electron. Let this bundle of lines of force be contained within the cone shown in Fig. 5.7, the axis of the cone containing the bundle being represented by *OB*. The cross section of the cone varies as the square of the distance from *O*. If the electron is quickly moved in a direction at right angles to *OB*, so that *O* moves to *C* and back to *O*, a kink will be sent out along the line of force *OB* and a situation similar to Fig. 5.5*c* will develop. But what is meant by a kink in a line of force? Such a kink is represented in Fig. 5.8. Consider the situation at *P*, Fig. 5.8. The direction of the tangent *PQ* gives the direction of the electric field at *P*. But the direction of the undisturbed electric field is in the direction *AB* or *PR*, which is parallel to *AB*. To obtain the vector *PQ* from the vector *PR* we must add the vector *RQ* which represents the strength of an electric field in a direction at right angles to *AB*. Hence each point of the kink has associated with it a transverse electric field at right angles to *AB*. This is represented by Fig. 5.8*b*, the lengths of the arrows representing the strength of the transverse electric field at various points at a given instant of time. If the heads of the arrows are joined by a curve, a wave form is obtained. If each arrow of Fig. 5.8*b* moves sideways with a velocity given by (5–47), the transverse electric field will maintain itself and the wave form will remain constant. Hence, by analogy with the cases of a stretched string and of sound waves in air, we may say that the velocity of propagation of electromagnetic waves is $c/\sqrt{\mu K}$.

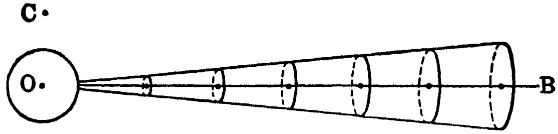


FIG. 5.7.—Bundle of lines of electric force leaving a portion of a charged body.

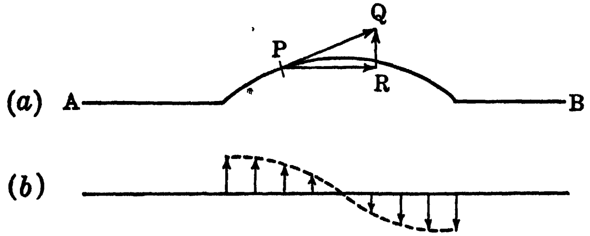


FIG. 5.8.—Propagation of a transverse electric field along a thin bundle of lines of electric force.

In the case of a vacuum, $K = \mu = 1$, so that the velocity of electromagnetic waves in a vacuum is c. We have previously remarked that the experimental value of c, the ratio of the emu of charge to the esu of charge, is 3×10^{10}, making $v = 3 \times 10^{10}$ cm/sec. Since the measured velocity of light is 3×10^{10} cm/sec, we may conclude that light is of electromagnetic origin. From (5–45) and (5–46), it is seen that for a light wave traveling in air (or a vacuum) R' is numerically equal to H. The velocity of light in a vacuum is usually represented by c, so that the symbol c is used to represent both the ratio of the units and the velocity of light.

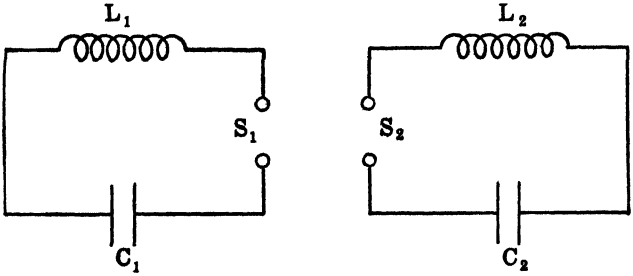


FIG. 5.9.—Experimental arrangement used by Hertz.

5.10. Experimental Verification.—In 1887 Hertz began experiments which were essentially as follows. A capacitance C_1 was charged to a voltage V_1. In series with the capacitance was an inductance L_1, and a spark gap S_1, as shown in Fig. 5.9. The voltage V_1 was sufficiently great to cause a spark to jump across S_1. As soon as a spark jumps, the air at S_1 becomes a fairly good conductor and the capacitance C_1 discharges through the inductance L_1. Since the resistance is small, the discharge is oscillatory and the circuit has a natural frequency

$$f = \frac{1}{2\pi\sqrt{L_1C_1}}$$

At some distance away, a second circuit of capacitance C_2 and inductance L_2 was set up. The length of the spark gap S_2 could be adjusted by means of a micrometer screw. It was found by Hertz that, with an adjustment of L_2 and C_2 such that L_2C_2 was nearly equal to L_1C_1, small sparks would jump across the spark gap S_2, and that the length of the sparks as measured by the micrometer was a maximum when L_2C_2 was exactly equal to L_1C_1. The frequency of the waves developed by Hertz was of the order of 100,000 kilocycles, which corresponds to a wave-length of 3 meters. Hertz showed that these electromagnetic waves could be reflected and refracted, that they showed interference phenomena and that they were polarized. In other words, he showed that they had many of the properties of light. This confirms the conclusion in Sec. 5.9 that light is of electromagnetic origin.

5.11. Refraction of Light.—Consider two media, (1) and (2). The wave or phase velocity of light in (1), according to (5–47), is

$$v_1 = \frac{c}{\sqrt{\mu_1 K_1}} \tag{5–48}$$

the subscripts (1) referring to medium (1). The wave or phase velocity of light in medium (2) is

$$v_2 = \frac{c}{\sqrt{\mu_2 K_2}} \tag{5–49}$$

On the wave theory of light, the index of refraction n from medium (1) to medium (2) is v_1/v_2. Hence from (5–48) and (5–49)

$$n = \sqrt{\frac{\mu_2 K_2}{\mu_1 K_1}} \tag{5–50}$$

However, in most transparent media, $\mu_2 = \mu_1 = 1$. If medium (1) is air, so that $K_1 = 1$, we have

$$n = \sqrt{K_2} \tag{5–51}$$

The relation (5–51) is approximately true in a medium showing little dispersion. If the dispersion is not negligible, account must be taken of the difference between wave and group velocities. This has been discussed in Sec. 3.9 for the case of carbon bisulphide (CS_2). However, another factor enters in the case of water, where $n = 1.33$ and $K_2 = 76$, so that $\sqrt{K_2} = 8.7$. The value 76 for K_2 is for steady electric fields. If, however, the dielectric constant for water is measured for rapidly alternating fields, it is found that its magnitude diminishes with increasing frequency. Light consists of very rapidly alternating electric and magnetic fields and thus the disagreement is qualitatively accounted for.

5.12. Pressure of Light.—If electromagnetic waves are traversing space, there is an electrostatic field R' and a magnetic field H at any point P in the space. From (5–45) and (5–46), it is seen that, if $v = c$

and $\mu = K = 1$ as is the case in a vacuum, then $R' = H$, or, in words, the numerical value of the electrostatic field in esu equals the numerical value of the magnetic field in emu. But the electrostatic and magnetic fields have energy densities of $R'^2/8\pi$ and $H^2/8\pi$ respectively, so that the total energy density is $R'^2/8\pi + H^2/8\pi = R'^2/4\pi = H^2/4\pi$. This energy moves forward with a velocity c, the velocity of light. Hence, if a unit area is taken perpendicular to the direction of propagation of the electromagnetic radiation, the rate of energy flow across this unit area will be $cR'^2/4\pi = cH^2/4\pi$. If the radiation is allowed to fall on an absorbing surface, energy flows into the surface at the rate of $cR'^2/4\pi$ erg/cm^2 sec. This energy is transformed into heat and the temperature of the absorbing material rises. However, it must be remembered that R' and H vary with the time according to $\sin 2\pi ft$, where f is the frequency of the light, and that therefore R'^2 and H^2 must be understood as representing the average values of R'^2 and H^2 respectively in these formulas. Let m be the mass of the absorbing material, T the rise in temperature, s the specific heat of the absorber, A the area presented perpendicularly to the radiation and t the time during which the radiation is being absorbed, the the rate at which heat is absorbed is msT/At cal/cm^2 sec. Multiplying by J, the mechanical equivalent of heat, the energy absorbed is $JmsT/At$ erg/cm^2 sec. Since this equals $cR'^2/4\pi$, we can obtain R'. For sunlight, the heat received on 1 cm^2 in 1 sec is about 0.03 cal/cm^2 sec. Since $J = 4.187 \times 10^7$ erg/cal, the energy received is 1.3×10^6 erg/cm^2 sec, so that

$$\frac{cR'^2}{4\pi} = 1.3 \times 10^6 \text{ erg/cm}^2 \text{ sec}$$

Remembering that $c = 3 \times 10^{10}$ cm/sec, we obtain the energy density of the radiation.

$$\frac{R'^2}{4\pi} = 4.3 \times 10^{-5} \text{ erg/cm}^3$$

Hence

$$R' = 0.023 \text{ esu}$$

and, since $R' = H$,

$$H = 0.023 \text{ emu}$$

In Sec. 5.5 we have mentioned that according to the Faraday-Maxwell theory an electrostatic field R' exerts a sideways pressure of

$R'^2/8\pi$ per unit area. Likewise a sideways pressure of $H^2/8\pi$ per unit area is produced by a magnetic field H. In electromagnetic radiation there is therefore a total sideways pressure of $R'^2/8\pi + H^2/8\pi = R'^2/4\pi = H^2/4\pi$ per unit area. Since the lines of force travel sideways in electromagnetic radiation, the pressure produced by such radiation is in the direction of propagation of the radiation. We have seen that $R'^2/4\pi = H^2/4\pi$ is the energy density in a beam of sunlight, so that the pressure produced by a beam of sunlight should equal the energy density in the beam. Since the energy density in sunlight is 4.3×10^{-5} erg/cm^3, the pressure exerted by a beam of sunlight on an absorbing surface placed perpendicular to the beam should be 4.3×10^{-5} dyne/cm^2. If, instead of allowing the light to fall upon an absorbing surface, the light falls upon a mirror, the reflected beam has as much energy density as the incident beam, so that the total energy density in front of the surface, and consequently the pressure on it, are doubled.

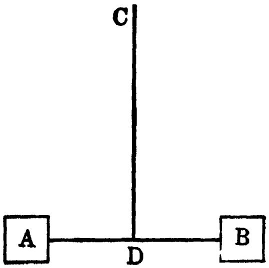


FIG. 5.10.—Apparatus for measuring the pressure of light.

In 1900 and 1901, Lebedew in Europe and Nichols and Hull in America experimentally proved the existence of the pressure of light. In the experiment of Nichols and Hull, two vertical mirrors, A and B, Fig. 5.10, are joined by a horizontal rod and the whole suspended by a quartz fibre CD. The whole system is in a vacuum. If a beam of light falls on A and is reflected, the mirror A is pushed back in the direction of propagation of the incident light. This causes a twist in the fibre. The angle of twist can be measured, and, if the torsional strength of the fibre is known, the pressure exerted by the light can be found. In one case, the observed energy density was 6.86×10^{-5} ergs/cm^3 and the observed pressure was 6.94×10^{-5} dyne/cm^2, which is an excellent agreement.

The pressure of sunlight at the earth is 4.3×10^{-5} dynes/cm^2 if the light is absorbed. This is very small. Atmospheric air pressure is 1,013,000 dynes/cm^2 and the pressure of sunlight is therefore only 4.3×10^{-11} atmosphere. The fact that the tails of comets always curve away from the sun has been explained by the pressure of the sunlight on the very attenuated matter of which comets' tails are supposed to consist.

CHAPTER V

PROBLEMS

1. A sphere of radius 5 cm carries a charge of 75 esu. How many lines of force leave the sphere? What is the flux density at a point distant 50 cm from the center of the sphere and what is the field strength at this point?

2. A condenser consists of two parallel plates, each of area 200 cm^2, separated by a dielectric plate of 1 mm thickness. Given that the dielectric constant of the plate is 2.5, find the capacitance of the condenser.

3. The plates of the condenser in Problem 2 are connected to a difference of potential of 10 esu. Determine the charge on each plate, the energy of the charged condenser, and the energy in a cubic millimeter of the dielectric.

4. Find the tension along the lines of force in the dielectric between the plates of the condenser of Problem 2 when the difference of potential between the plates is 20 esu.

5. A uniform magnetic field of 1000 emu is traveling at a speed of 10^8 cm/sec in a direction at right angles to itself. Find the strength in esu of the electric field produced.

6. What is the wave or phase velocity of electromagnetic radiation in a medium whose dielectric constant is 3.5?

CHAPTER VI

PROPERTIES OF MOVING CHARGED BODIES

6.1. Moving Charged Sphere.—In Secs. 5.6 and 5.7 it has been assumed that the electric or magnetic field was moving in a direction at right angles to the direction of the field. If, however, the field is moving in a direction making an angle θ with the direction of the field, formulas (5–38) and (5–39) become

$$cR' = j\mu Hv \sin\theta \text{ (moving magnetic field)} \tag{6–1}$$

$$cH = -jKR'v \sin\theta \text{ (moving electric field)} \tag{6–2}$$

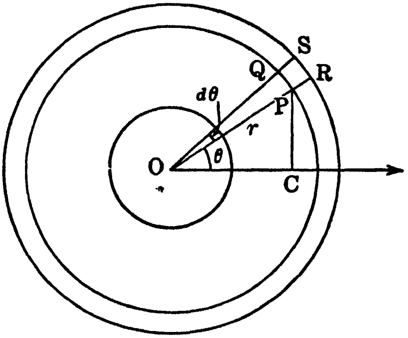


Fig. 6.1.—Field produced by moving charged sphere.

Let us now consider a sphere of radius a, carrying a charge Q' and moving with velocity v as shown in Fig. 6.1. It is supposed that v is small compared with the velocity of light c.

The center of the sphere is at O. Next, let us consider the electric field strength at a point P, such that $OP = r$ and makes an angle θ with OC, the direction of the velocity of the sphere. The electric field strength at P is Q'/r^2 and is along OP. This field travels with the sphere and thus moves with a velocity v in a direction θ with itself. Hence, a magnetic field of strength H is produced at P, such that

$$cH = K\left(\frac{Q'}{r^2}\right) v \sin\theta \tag{6–3}$$

The symbol $-j$ is omitted in (6–3) because at present we are interested only in the magnitude and not in the direction of H. If the medium at P is a vacuum, $K = 1$, so that

$$H = \frac{Q'v \sin \theta}{cr^2} \tag{6–4}$$

The magnetic energy per unit volume at P is therefore

$$\frac{H^2}{8\pi} = \frac{Q'^2v^2 \sin^2 \theta}{8\pi c^2r^4} \tag{6–5}$$

Let us rotate the radius vector OP of Fig. 6.1 to the position OQ. Let the angle $QOP = d\theta$. Also let OP be extended to R and OQ to S and let $PR = QS = dr$. If $d\theta$ and dr are both small, the value of $H^2/8\pi$ is sensibly the same over the area $PQSR$. Now rotate the lines OP and OQ and the area $PQSR$ about the axis Ov. The area $PQSR$ then carves out a ring-shaped volume about Ov as axis. The area of $PQSR$ which is the area of cross section of the ring is $rd\theta dr$. The length of the ring is the length of the circumference of the circle whose radius is PC and is therefore $2\pi r \sin \theta$. Hence the volume of the ring is $2\pi r^2 \sin \theta d\theta dr$, and the magnetic energy in the ring is

$$\frac{Q'^2v^2 \sin^2 \theta \cdot 2\pi r^2 \sin \theta d\theta dr}{8\pi c^2r^4} = \frac{Q'^2v^2 \sin^3 \theta d\theta dr}{4c^2r^2} \tag{6–6}$$

If we integrate this from $\theta = 0$ to $\theta = \pi$, we obtain the magnetic energy in a spherical shell of thickness dr and radius r. While we are integrating with respect to θ, r is to be treated as a constant. Now, omitting limits,

$$\int \sin^3 \theta d\theta = - \cos \theta + \frac{(\cos^3 \theta)}{3} \tag{6–7}$$

since, on differentiating the right side, we obtain the integrand $\sin^3 \theta$. Inserting the limits, we obtain

$$\left[- \cos \theta + \frac{(\cos^3 \theta)}{3}\right]_0^\pi = \frac{4}{3} \tag{6–8}$$

From this result, the energy in the spherical shell becomes

$$\frac{Q'^2v^2dr}{3r^2c^2} \tag{6–9}$$

To obtain the total magnetic energy in the space outside the charged sphere of radius a, we add the energy in all the shells of thickness dr between $r = a$ and $r = \infty$, that is, we integrate (6–9) with respect to r between the limits $r = a$ and $r = \infty$. Hence, if W represents the total magnetic energy,

$$W = \int_{r=a}^{r=\infty} \left(\frac{Q'^2v^2}{3r^2c^2}\right) dr \tag{6–10}$$

Now, omitting limits, we have

$$\int \frac{dr}{r^2} = -\frac{1}{r} \tag{6–11}$$

since, on differentiating the right side, we obtain the integrand $1/r^2$. Inserting the limits, we obtain

$$\left[-\frac{1}{r}\right]_a^\infty = \frac{1}{a} \tag{6–12}$$

since $1/\infty = 0$. The energy is then

$$W = \frac{Q'^2v^2}{3ac^2} \tag{6–13}$$

This may be written in the form

$$W = (\tfrac{1}{2}) \cdot \left(\frac{2Q'^2}{3ac^2}\right) v^2 \tag{6–14}$$

The formula for the kinetic energy of a body of mass m is

$$W = (\tfrac{1}{2})mv^2 \tag{6–15}$$

Comparing (6–14) with (6–15), we see that the quantity $2Q'^2/3ac^2$ behaves as a mass. The mass of a sphere when charged is therefore greater than that of the same sphere when uncharged by an amount

$$m = \frac{2Q'^2}{3ac^2} \tag{6–16}$$

The question then arises as to whether all mass may not be electrical in nature. It is the modern belief that the entire mass of the electron is electrical. The magnetic energy of a moving charge is of the nature of kinetic energy.

6.2. Mass of Energy.—Consider a charged sphere at rest. The electric field strength at a point P as in Fig. 6.1 is Q'/r^2 where Q' is the charge on the sphere and r is the distance of P from the center of the sphere. The work done by the field when unit charge is moved from the surface of the sphere to infinity is

$$V' = \int_a^\infty \frac{Q'}{r^2}\,dr = -\left[\frac{Q'}{r}\right]_a^\infty = \frac{Q'}{a} \qquad (6\text{–}17)$$

where a is the radius of the sphere. The quantity V' is the potential of the sphere so that the capacitance of the sphere is

$$C' = Q'/V' = a \qquad (6\text{–}18)$$

The capacitance of a sphere is thus equal to its radius in centimeters. For this reason the unit of capacitance in esu is the centimeter. The relation of this unit to the micro-microfarad (10^{-12} farad) is

$$1 \text{ cm capacitance} = 1.11 \text{ micro-microfarads} \qquad (6\text{–}19)$$

The electrostatic energy of the charged sphere is

$$W = \frac{Q'^2}{2C'} = \frac{Q'^2}{2a} \qquad (6\text{–}20)$$

Eliminating Q'^2/a from (6–16) by means of (6–20),

$$m = \left(\frac{4}{3}\right)\frac{W}{c^2} \qquad (6\text{–}21)$$

Now when a soap bubble is charged the bubble expands slightly. This expansion against the surface tension of the bubble and the difference of the air pressures inside and outside the bubble means that in charging the bubble not only has work to be done to assemble the charge on the bubble but also to expand the bubble. The total work done in charging the bubble is somewhat greater than the quantity on the right side of (6–20) and the factor $4/3$ in (6–21) is replaced by a factor γ, the value of γ being less than $4/3$. Earnshaw's theorem in electrostatics states that electrical charges can never be in stable equilibrium under the action of electrical forces alone. For example, the electron cannot hold together because each part of the electron repels each other part unless it is held together by non-electrical forces. Since the electron must be held together in somewhat the same way as a charged soap bubble, the relation between the mass and the charge on the electron must be

$$m = \frac{\gamma W}{c^2} \qquad (6\text{–}22)$$

where γ depends on the elastic and non-electrical forces holding the electron together and also on the way the charge is distributed throughout the electron. For instance, it makes a difference whether the charge of an electron exists on the surface of a sphere or is uniformly distributed throughout a sphere. The stability of the electron is one of the mysteries of modern physics.

According to Einstein's special theory of relativity, which was published in 1905, the value of γ in (6–22) is rigorously unity and the relation holds not only for the electron but also for any energy. The relativity theory states that all energy has mass and that the relation between mass and energy is

$$\text{mass} = \text{energy}/c^2 \tag{6–23}$$

where c is the velocity of light in cm/sec, the energy is in ergs and the mass in grams. This formula is valid for both potential and kinetic energy. A wound watch spring has a greater mass than the same spring unwound, while a rapidly rotating fly-wheel has a greater mass than the same fly-wheel at rest. The increase of mass in these two cases is very small owing to the large value of c in (6–23).

6.3. Effect of External Magnetic Field.—We have seen in Sec. 6.1 that a moving charged body produces a magnetic field. If, however, an external magnetic field due to permanent magnets or to other causes is present, what is the effect of this field upon the motion of the charged body? In Fig. 6.1, the magnetic field produced at P by a moving charge is, according to (6–4),

$$H = \frac{Q'v \sin \theta}{cr^2} \tag{6–24}$$

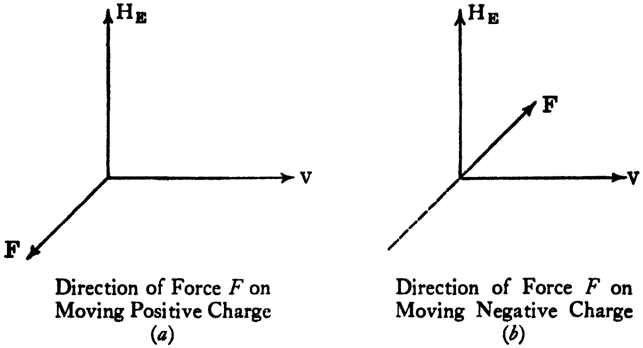
If Q' is a positive charge, the direction of this field is towards the reader from the plane of the figure. Now, if a magnetic north pole of strength m is placed at P, the force acting on this pole is

$$Hm = \frac{Q'mv \sin \theta}{cr^2} \tag{6–25}$$

But m/r^2 is the strength of the magnetic field due to the pole at the center of the moving charged body. This field may be considered as an external magnetic field through which the charged body is moving. Let us therefore represent m/r^2 by H_E. The force acting on the magnetic pole at P is towards the reader in Fig. 6.1. By Newton's third law, which states that action and reaction must be equal and opposite, there must be a reaction on the moving charge in the direction away from the reader. The force acting on the charge must be of the same magnitude as the force acting on the pole and so is

$$\frac{H_E Q' v \sin\theta}{c} \tag{6-26}$$

If the external magnetic field H_E is at right angles to the direction of motion of the charged body, the magnitude of the force acting on the charged body is $H_E Q' v/c$ when Q' is in esu and is $H_E Q v$ when Q is in emu. The direction of the force acting on a positively charged body is shown in Fig. 6.2*a*.



Direction of Force *F* on Moving Positive Charge (*a*)

Direction of Force *F* on Moving Negative Charge (*b*)

Fig. 6.2.

The direction of the force acting on a moving negatively charged body is in the opposite direction to that acting on a moving positively charged body. This direction is shown in Fig. 6.2*b*

6.4. Rate of Radiation of Energy.—In Sec. 5.9, it is shown that an electrostatic field of strength R' which moves at right angles (or transverse) to itself with a velocity c maintains itself. The electric energy per unit volume in this field is $R'^2/8\pi$ and the magnetic energy is $H^2/8\pi$. But in this case $H = R'$, so that the total energy is $R'^2/4\pi$ or $H^2/4\pi$. This energy travels with a speed c. Hence the rate at which energy crosses a unit area which is held stationary and perpendicular to the direction of the velocity is c times the energy per unit volume, so that

$$\text{Rate of radiation across unit area} = \frac{R'^2 c}{4\pi} = \frac{H^2 c}{4\pi} \tag{6-27}$$

By methods beyond the scope of this book it can be shown that, when a body carrying a charge Q' is accelerated in the direction OY, Fig. 6.3, the rate of radiation across unit area at P taken perpendicular to OP is

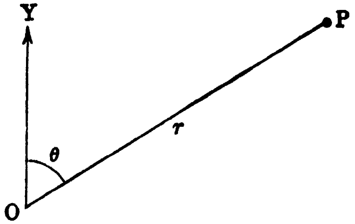


Fig. 6.3.

$$\frac{Q'^2a^2 \sin^2 \theta}{4\pi r^2c^3} \tag{6-28}$$

where a is the acceleration of the charged body. It is to be noted that the rate of radiation varies inversely as the square of the distance, which agrees with experience in the case of light, and also that there is no radiation in the direction of the acceleration, that is, when $\theta = 0$. If we integrate the radiation over the whole surface of a sphere surrounding the accelerated charged body at O, we obtain

$$\text{Total rate of radiation} = \frac{2Q'^2a^2}{3c^3} \tag{6-29}$$

Formulas (6–28) and (6–29) apply to an accelerated electron. An electron which is vibrating about a center with simple harmonic motion is moving with acceleration and so sends out radiation. In this case the radiation has a frequency which is the same as that of the vibrating electron. Since the vibrating electron is sending out energy in the form of radiation the electron itself is losing energy, so that the vibrations die down unless energy is supplied to the electron from some other source. In mechanics, it is proved that a body moving with a uniform speed v in a circle of radius r is moving with an acceleration v^2/r toward the center of the circle. Hence an electron which is moving with uniform speed in a circle is moving with acceleration and so must radiate. An electron which is stopped during a very small interval of time is subject to a retarding acceleration and should therefore radiate. This is one theory of the production of x-rays when fast moving electrons hit the target in an x-ray tube. We shall return to this matter later.

We thus see that the classical electromagnetic theory of radiation requires that an electron moving with acceleration must radiate and in particular that an electron moving with uniform speed in a circle must radiate. We shall see later when we come to study the Bohr theory of the hydrogen atom that Bohr makes the assumption that an electron moving in a circular orbit does not radiate. The Bohr theory thus starts out with a denial of one of the fundamental requirements of the classical theory. We shall return to this later.

6.5. Radius of the Electron.—If the electron is considered as a charged sphere of radius a, then from (6–16) its mass is $2e'^2/3ac^2$. This may be written

$$a = \frac{2e'^2}{3mc^2} \tag{6–30}$$

where e' is the charge in esu and m is the mass of the electron. We shall learn in Chapter VII that both e'/m and e' can be measured and that their values are 5.28×10^{17} esu/gm and 4.8×10^{-10} esu respectively. Substituting in (6–30), we find that $a = 1.88 \times 10^{-13}$ cm, which is the radius of a spherical electron. Of course, we have no reason for supposing the electron to be actually spherical, but, at least, we feel that its linear dimensions must be of the order of 10^{-13} cm. In 1946, it became the practice to replace the factor ⅔ in (6–30) by unity and so obtain

$$a = 2.82 \times 10^{-13} \text{ cm}$$

for the radius of the electron.

CHAPTER VI

PROBLEMS

1. A sphere of radius 10 cm carries a charge of 150 esu. What is the capacitance of the sphere? Determine the potential of the sphere and the energy of the charge on the sphere.

2. A sphere of radius 0.5 cm carries a charge of 20 esu. Find the increase in its mass due to its charge.

3. An electron is a small body whose radius is of the order of 10^{-13} cm and which carries a charge of -4.8×10^{-10} esu. A certain electron is traveling with a speed of 10^8 cm/sec. Determine the magnetic field produced by this electron at a point distant 10^{-7} cm from the center of the electron and in a direction from the center of the electron at right angles with the motion of the electron.

4. An electron moves with a speed of 5×10^8 cm/sec through a magnetic field of strength 10 emu. The motion of the electron is at right angles to the field. Determine the force acting on the electron.

5. An electron traveling with a speed of 2×10^9 cm/sec is brought to rest in a distance of 10^{-8} cm. Assuming that it is brought to rest with uniform acceleration, find the acceleration. Next, find the rate at which radiation is given out by the electron as it is being brought to rest. Finally, find the total radiation given out.

CHAPTER VII

THE ELECTRON

7.1. Electric Discharge in Gases.—A glass tube *A* is constructed with two metal electrodes *B* and *C* as shown in Fig. 7.1. *D* is a side tube which is connected to an air pump so that the tube *A* may be exhausted. A voltage sufficient to cause a discharge to pass between the electrodes is applied. When the tube is filled with gas at atmospheric pressure, the discharge takes the form of a spark which jumps between the electrodes provided that the potential between the electrodes is of the order of 50,000 to 100,000 volts. The spark is accompanied by a sharp crackling noise, appears to be intermittent, and follows a continually varying zig-zag course. On gradually exhausting the gas from the tube, the character of the discharge changes. As the pressure decreases, the voltage necessary to force a discharge decreases, the spark becomes fatter, and the sharp crackling sound is replaced by a continuous buzzing sound. As the exhaustion proceeds further, still less voltage is necessary to force a current through the tube, and a glow fills the tube. The color of the glow depends on the gas originally in the tube—reddish in the case of air, bright red in the case of neon, and bluish in the case of carbon dioxide. At further exhaustion the voltage necessary to produce a discharge increases. This decrease and later increase of the voltage as the pressure decreases is shown by Fig. 7.2. The voltage is plotted against the pressure and the curve is for the case of a current of 10 milliamperes forced through a tube in which the electrodes are

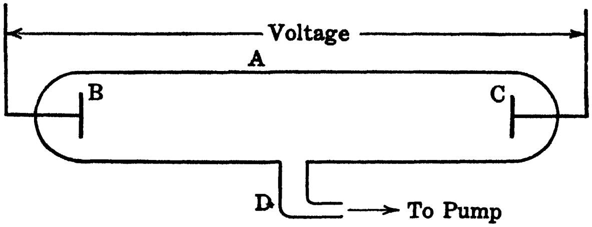


Fig. 7.1.—Diagram of the electric discharge tube.

11.5 cm apart. It is seen that the potential across the tube drops to a minimum of 470 volts when the pressure is about 1 mm of mercury. As the pressure decreases below 1 mm, a greater and greater voltage is necessary to force the current through the tube. As it becomes harder to force a current through the tube as the pressure decreased after the minimum is passed, the tube is said to become harder. A tube in which the pressure is low and which requires a high voltage to force a discharge through it is therefore said to be hard. In distinction, a tube in which the pressure is relatively high and which consequently requires a low voltage is said to be soft.

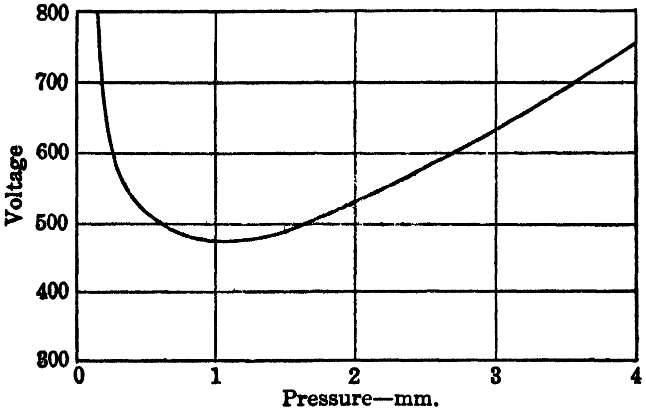


FIG. 7.2.—Variation of the voltage across a certain discharge tube with the pressure when a current of 10 ma is maintained through the tube.

7.2. Cathode Rays.—At atmospheric pressure and down to a pressure of about 10 mm of mercury, there is no particular difference between the discharge near one electrode and that near the other. However, at a pressure of about 3 or 4 mm a difference between the two ends becomes evident. At a pressure of about 1 mm, the tube has a striking and characteristic appearance as shown in Fig. 7.3. The luminous discharge between the electrodes is seen to break up into five distinct parts: (1) around the negative electrode or cathode at *A* there is a thin layer of luminosity which is sometimes described as a velvety glow, then (2) a sharply defined dark space *B*, followed by (3) another luminous region *C*, then (4) an ill-defined dark region *D*, and finally (5) a column of luminosity *E* extending to the positive electrode or anode. The column *E* is called the positive column and is sometimes striated. The first dark space *B* is called Crookes' dark space, while the second dark space *D* is called Faraday's dark space. The most important of these regions in our present consideration is Crookes' space. As the pressure decreases below 1 mm and approaches 1/100 mm, the Crookes' space continually increases in width while the luminous regions decrease in extent and finally vanish. When the pressure is such that Crookes' space partially or wholly fills the tube, it is noticed that the glass walls of the tube fluoresce with a green light. The pressure is now about 1/100 mm. Certain salts when placed in the Crookes' space are found to fluoresce. If a screen covered with a layer of one of these fluorescent salts is placed in the tube, it is found that, if an object is placed between the cathode and the fluorescent screen, a shadow is thrown on the screen. It appears as though something is streaming from the cathode, and hence "cathode rays" are said to be shot off by the cathode. These rays appear to leave the cathode almost at right angles to its surface, the direction being independent of the position of the anode.

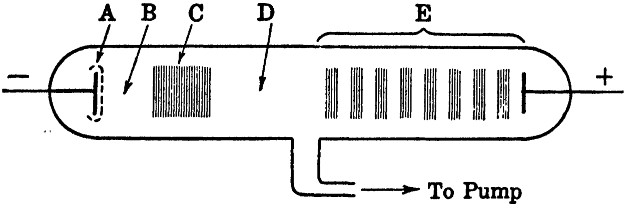


Fig. 7.3.—Appearance of a discharge tube at a pressure of about 1 mm of mercury.

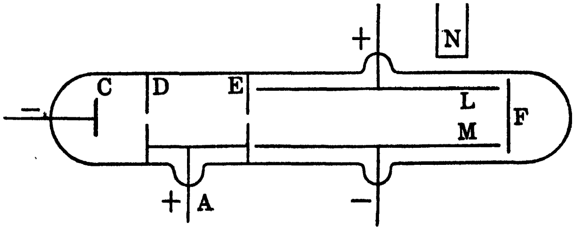


Fig. 7.4.—Discharge tube for showing the electrostatic or magnetic deflection of a beam of cathode rays.

7.3. Electrostatic Deflection.—We have just noticed that the cathode rays come off at right angles to the surface of the cathode. Advantage is taken of this fact in the construction of the tube shown in Fig. 7.4. The tube is constructed with a cathode *C*, two metal diaphragms *D* and *E*, and fluorescent screen *F*. The diaphragms *D* and *E* are electrically connected to the anode *A*. The pressure is about 1/100 mm. Some of the cathode rays pass through the aperture in the first diaphragm *D* and some of these pass through the aperture in the second diaphragm *E*. The rays which pass through both apertures form a well-defined beam which produces a definite spot of fluorescence on the screen *F*. If a positively charged rod is brought near the space *EF*, the fluorescent spot moves towards the rod, indicating that the particles coming from the cathode are charged negatively. If a negatively charged rod is presented to the tube, the spot moves away from the rod, verifying our former deduction.

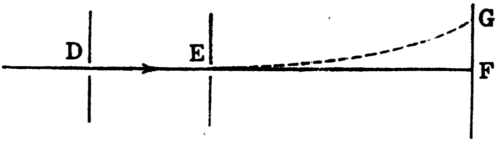


Fig. 7.5.—Electrostatic deflection of a beam of cathode rays.

If two parallel plates *L* and *M* are placed in the space between *E* and *F*, Fig. 7.4, and if a difference of potential is applied to these plates, an electrostatic field is obtained which is uniform except at the edges. When a difference of potential V' esu is applied to the plates *L* and *M* the fluorescent spot moves from *F* to *G* as in Fig. 7.5. If s is the distance between the plates *L* and *M*, the electrostatic field between the plates is V'/s. Further, if we suppose that each cathode particle carries a charge e', the upward force acting on each particle due to the field is

$$\frac{V'e'}{s} \tag{7-1}$$

This force remains upward even though the path of the particles is bent as shown by the broken line *EG*, Fig. 7.5. If the mass of each particle is m, the upward acceleration produced by the force will be

$$\frac{V'e'}{sm} \tag{7-2}$$

Since the original velocity of the cathode particles entering the space *EF* is horizontal in Fig. 7.4 and the acceleration is upward, the path of the cathode particles is similar to that of a projectile fired horizontally from a gun on a hill—in other words, the path is parabolic.

If, for simplicity, we suppose that the plates L and M extend very nearly all the way from E to F, Fig. 7.4, and that the distance from E to F is l, the time for the cathode particles to traverse the space EF is l/v, where v is the velocity of the particles entering the space. During this time the particles are moving upwards with an acceleration $V'e'/sm$. The distance moved upwards is the electrostatic deflection FG, Fig. 7.5, which we shall represent by d_E, so that

$$d_E = \frac{1}{2}\cdot\frac{V'e'}{sm}\cdot\frac{l^2}{v^2} = \frac{V'e'l^2}{2smv^2} \tag{7-3}$$

The quantities, d_E, V', s can be measured, so that we have an equation with three unknowns, e', m and v. One equation does not suffice for the finding of three unknown quantities, so other equations containing these unknowns must be developed.

7.4. Magnetic Deflection.—If, instead of applying an electric field to the space between E and F, Fig. 7.4, a north magnetic pole is presented to the tube as at N, it is found that the spot on the fluorescent screen is deflected. The deflection, however, is neither away from nor toward the magnet. If the north pole is presented as shown in Fig. 7.4, the deflection is perpendicular to the plane of the paper and outwards towards the reader. By referring to Fig. 6.2*b*, it is seen that this is the direction in which negatively charged bodies traveling from C to F are deflected, verifying the deduction we drew from the direction of the electrostatic deflection. Further confirmation of the sign of the charge on the cathode particles was found by Perrin who allowed the particles to enter a metal chamber connected to an electroscope. The electroscope became charged negatively.

Instead of presenting the north pole of a bar magnet to the tube as in Fig. 7.4, let us arrange to have a uniform magnetic field of strength H applied to the space EF and let the direction of H be perpendicular to the plane of the diagram in Fig. 7.4 and outwards towards the reader. Negatively charged particles will then according to Sec. 6.3, be acted on by a force toward the top of the page in Fig. 7.4. If each particle carries a charge e' esu and moves with velocity v, the force acting upon each particle, according to (6–26), is

$$\frac{He'v}{c} \tag{7-4}$$

This causes the path to turn upwards as shown by the broken line EG

in Fig. 7.5. The force is always perpendicular to the direction of motion of the charged particle, so that the broken line EG, Fig. 7.5, is an arc of a circle. The force acting on the particle is always toward the center of the circle and this force must equal the centrifugal force due to the circular motion of the particle, so that

$$\frac{He'v}{c} = \frac{mv^2}{r} \tag{7-5}$$

where m is the mass of the particle and r is the radius of the circular path EG. Consider the circle whose center is at C, Fig. 7.6. It is shown in geometry that $BA \times AE = AG^2$. However, if AE is small compared to the diameter, then, since $AE = GF$ and $AG = EF$,

$$GF = \frac{EF^2}{2r}$$

where r is the radius of the circle. Applying this result to Fig. 7.5, we have

$$d_M = \frac{l^2}{2r} \tag{7-6}$$

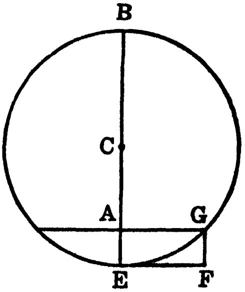


FIG. 7.6.

where $d_M = GF$, the deflection of the fluorescent spot in Fig. 7.5 when the uniform magnetic field is applied, and $l = EF$. Eliminating r from (7-5) and (7-6), we obtain

$$d_M = \frac{He'l^2}{2mvc} \tag{7-7}$$

Comparing (7-3) and (7-7) we note that these equations can be solved simultaneously for e'/m and v. The solution is

$$v = \frac{V'cd_M}{Hd_{ES}} \tag{7-8}$$

and

$$\frac{e'}{m} = \frac{2V'c^2d_M{}^2}{H^2l^2d_{ES}} \tag{7-9}$$

These quantities on the right sides of (7-8) and (7-9) can be measured experimentally, and, when the experimental values are substituted, it is found that, while the velocity v depends on the conditions of pressure and voltage under which the discharge tube is operated, the value of e'/m always comes out the same, this value being 5.28×10^{17} esu/gm. The value of v is of the order of 10^9 cm/sec. We see therefore that these negatively charged cathode particles have a charge and mass which are such that the ratio of the charge to the mass is always the same.

7.5. Digression on Electrolysis.—The reader is already familiar with Faraday's laws of electrolysis. When an electric current is passed through water which is slightly acidified, oxygen gas is given off at the anode and hydrogen gas is given off at the cathode. The electric current is supposed to be carried through the water by ions, in the present case by hydrogen ions which are charged positively and by oxygen ions which are charged negatively. Faraday's first law states that the mass of any material deposited electrolytically on an electrode is proportional to the quantity of electricity which has passed through the electrolyte, while his second law states that if the same quantity of electricity is passed through different electrolytes the mass of each material deposited is proportional to the chemical equivalent of the material. The electrochemical equivalent of an ionic substance is defined as the mass of the substance deposited when unit quantity of electricity is sent through the electrolyte. The electrochemical equivalent for the hydrogen ion is .00001046 gram per coulomb. If the unit of quantity (or charge) is taken as the electromagnetic unit instead of the practical unit or coulomb which is 1/10 emu, the electrochemical equivalent of hydrogen is .0001046 gm/emu. All that we know from electrolysis in the above case is that hydrogen is deposited on the cathode of the electrolytic cell and that the ratio of the mass deposited to the charge sent through the electrolyte is .0001046 gm/emu. We do not know that each of the hydrogen ions in the electrolyte has the same mass and carries the same charge and that the mass and charge are such that the ratio of the mass to the charge is the figure given above. However, if we make the assumption that each hydrogen ion in electrolysis has a mass M and carries a charge E emu, the result is that for each hydrogen ion $M/E =$ .0001046 gm/emu. If we invert this, we obtain for the hydrogen ion in electrolysis

$$\frac{E}{M} = 9565 \text{ emu/gm} \tag{7-10}$$

If we express the ratio of the charge to the mass of the cathode particles in electromagnetic units, then, since $e' = ce$, the ratio of the charge to the mass for the cathode particle is

$$\frac{e}{m} = 1.76 \times 10^7 \text{ emu/gm} \tag{7-11}$$

Comparing (7–10) and (7–11), it is seen that e/m for the cathode particle is 1840 times E/M for the hydrogen ion in electrolysis. It should always be remembered that the charge on the cathode particle is negative while that on the hydrogen ion is positive. It is, however, the practice to omit the negative sign in the value given for e/m in (7–11).

7.6. Early History of Cathode Rays.—Crookes' dark space is only relatively dark as compared with the bright luminosity in other parts of the tube. At a pressure somewhat below 1/100 mm the space is filled with a faint bluish light. In 1859 the German physicist Plücker found that this stream of bluish light which starts from the cathode is deflected by a magnet outside the tube. Further, he found that the glass walls of the tube fluoresce with a greenish-yellow color at a spot to which the faint bluish light extends and that the position of this spot was affected by a magnet outside the tube. In 1879 the English physicist Crookes began examining the phenomenon found by Plücker. Two schools of thought grew up: the German physicists fixed their attention on the faint bluish light streaming from the cathode whereas the English physicists fixed their attention on the fluorescent glass or the fluorescent screen which was placed in the tube. Plücker believed that he had discovered "magnetic light." Crookes on the other hand believed that the effect was caused by negatively charged molecules proceeding from the cathode.

In 1874 Johnstone Stoney in addressing the British Association pointed out that it is almost a necessary consequence of Faraday's laws of electrolysis and of the atomic theory that electricity should be atomic in character, that is, that any charge of electricity must always contain an integral number of elementary charges which cannot be further subdivided. Stoney suggested the name "electron" for one of these elementary charges of either sign.

The magnetic deflection of cathode rays was observed by Plücker in 1859. No certain electrostatic deflection of cathode rays was observed before 1897 when J. J. Thomson, at the Cavendish Laboratory, Cambridge, England, performed the experiment described in Sec. 7.3. Placing a charged hard-rubber rod outside a discharge tube did not deflect the cathode rays. We know now that this lack of deflection is caused by charges which collect on the surface of the glass. As early as 1890 there were indications that e/m for the cathode rays was at least several hundred times larger than E/M for the hydrogen ion in electrolysis. Such indications implied the then appalling idea that the mass of a cathode ray particle was several hundred times less than that of the hydrogen atom or ion. At that time any particle smaller than the hydrogen atom was inconceivable. But the brilliant experiments of J. J. Thomson as described in Secs. 7.3 and 7.4 forced the physicists of 1897 to revise their ideas.

In 1897 the question arose as to the interpretation of e/m for the cathode particle being 1840 times E/M for the hydrogen ion. This could only be answered by devising an experiment to measure e, the charge on the cathode particle. This was done by Thomson and his collaborators who found that e for the cathode particle has the same value as E for the hydrogen ion. Stoney's term "electron" began to be applied to the cathode particles and from now on we shall refer to these particles as electrons. At the end of 1897 it was conceded that the electron had a mass 1/1840 times that of the hydrogen atom or ion. The hydrogen atom or ion was no longer the smallest particle. Thus the invasion of the subatomic world began.

7.7. The Charge on the Electron.—Thomson's method of measuring the charge on the electron has been improved by Millikan and we shall describe Millikan's method of determining the electronic charge. Both Thomson's and Millikan's methods depend upon the velocity with which small liquid drops fall in air. Raindrops falling from a cloud a mile high do not hit the surface of the earth with the velocity of free fall given by $v^2 = 2gh$. If they did hit the surface of the earth with this velocity it would be dangerous to be out in a rain storm. The raindrops begin to fall with the acceleration g but as they speed up the friction, or viscosity, of the air comes into play and the drop finally attains a constant speed of a few yards per second. Stokes investigated this problem for a falling sphere of radius a and found that the final constant velocity is given by

$$v = \frac{2ga^2\rho}{9\eta} \tag{7-12}$$

where ρ is the density of the sphere, η is the coefficient of viscosity of the gas and g is the acceleration of gravity.

If the drops are small so that v is only a fraction of a millimeter per second, Stokes' law can be used to determine a, the radius of the drop. Knowing the radius, the mass and the weight of the spherical drop can be determined, the mass being $(4/3)\pi a^3\rho$. Thomson used water drops, which, however, evaporate quickly and are unsuitable. Millikan used drops of oil and mercury which do not evaporate quickly. If an electron can be attached or detached from one of Millikan's oil drops, the drop will be charged negatively or positively respectively. An electron can be detached from a drop by means of illuminating the drop with ultraviolet light or x-rays as we shall see in the chapters on The Photo-electric Effect and on X-Rays. Also, when the drops are produced by an atomizer, they are usually charged by friction. A charge is given in one of these ways to a drop C placed between two plates A and B as shown in Fig. 7.7.

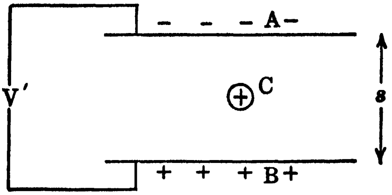


FIG. 7.7.

The drop is observed by means of a microscope which contains a scale. The potential difference between the plates A and B is at first made zero and the velocity of the falling drop determined by means of the scale in the microscope and a stop-watch. Using (7–12) the radius of the drop can be found from this velocity. Next, a potential difference V' esu is applied to A and B, and V' is adjusted until the image of the drop in the field of the microscope remains stationary. The electric field strength acting on the drop is V'/s and the force acting on the drop is $V'e'/s$ if the charge on the drop is e' esu. This force must just balance the weight of the drop so that

$$\frac{V'e'}{s} = (4/3)\pi a^3\rho g \qquad (7\text{–}13)$$

Eliminating a from (7–12) and (7–13), we obtain

$$e' = \frac{18\pi s}{V'}\sqrt{\frac{v^3\eta^3}{2g\rho}} \qquad (7\text{–}14)$$

All the quantities on the right of (7–14) can be measured and so e' can be determined. The value of e' as found by Millikan's method is

$$e' = 4.8 \times 10^{-10} \text{ esu} \tag{7–15}$$

or

$$e = 1.6 \times 10^{-20} \text{ emu} \tag{7–16}$$

In Millikan's experiments it was not always possible to charge an oil drop with one elementary charge. Sometimes two, three or perhaps four elementary charges are given to the drop. In the beginning the observer does not know how many elementary charges are on the drop. However, Millikan always found that if the charge on the drop was not 4.8×10^{-10} esu it was some integral multiple of this amount; the charge was never one-half or three-halves this amount. It therefore seems that electricity is atomic and that the smallest possible charge of electricity is given by (7–15) or (7–16).

7.8. Avogadro's Number.—How does the elementary charge given by (7–15) compare with the charge on the hydrogen ion in electrolysis? We know that it takes 95,650 coulombs or 9565 emu of electricity to liberate 1 gm of hydrogen in electrolysis. If this charge of 9565 emu is carried by hydrogen ions and there are n ions in one gram of hydrogen, the charge on each ion is $9565/n$ emu. This brings us to the consideration of Avogadro's number, which is one of the most important constants of physics and chemistry. Before defining Avogadro's number, however, we shall describe what is meant by the terms "gram-atom" and "gram-molecule."

In chemistry it is known that all substances are either compounds or elements and that the compounds are made up of elements. According to the atomic theory each element consists of atoms all of the same kind and weight, and the weight of each atom when expressed in terms of the weight of the hydrogen atom was called the atomic weight. In 1815 Prout of England suggested the hypothesis that, if the atomic weight of hydrogen be taken as unity, the atomic weights of other elements would be whole numbers. However, accurate investigations showed that, if the atomic weight of hydrogen is taken as unity, the atomic weight of oxygen is not the whole number 16 but the number 15.87. Also the atomic weight of carbon was found to be 11.91 instead of 12. It is desirable if possible to express the atomic weights of the elements as whole numbers and it has been found that, if the atomic weight of oxygen is taken as 16, the atomic weights of many other elements are whole numbers or very nearly whole numbers. With O = 16, the atomic weight of hydrogen is 1.0080, that of carbon is 12, nitrogen 14, sodium 23. However, even this scheme is not perfect, for the atomic weight of chlorine is 35.46. The atomic weights based upon O = 16 are known as the international atomic weights.

A molecule is defined as the smallest portion of a compound which can exist without the compound breaking up into its elements. The chemical formula for water is H_2O, which means that the molecule of water consists of two atoms of hydrogen and one atom of oxygen. The molecular weight of a compound is the sum of the atomic weights of the atoms of which it is composed. The molecular weight of water is therefore $2 \times 1.0080 + 1 \times 16 = 18.0160$. For the purposes of chemistry it is usually not necessary to know the weights of the atoms in grams. The weight of the oxygen atom is taken as 16 and the weights of all other atoms are expressed in terms of oxygen without knowing the weight of the oxygen atom in grams. However, substances are weighed out in grams in the laboratory and the terms gram-atom and gram-molecule have come into use. A gram-atom is defined as that mass of an element measured in grams which is numerically equal to the atomic weight. A gram-atom of hydrogen is thus 1.0080 gm of hydrogen, a gram-atom of oxygen is 16 gm of oxygen, and a gram-atom of chlorine is 35.46 gm of chlorine. Similarly a gram-molecule is defined as that mass of a compound measured in grams which is numerically equal to the molecular weight. A gram-molecule of water (H_2O) is thus 18.0160 gm of water.

Avogadro's number is defined as the number of atoms in a gram-atom of an element or as the number of molecules in a gram-molecule of a compound. Since the atomic weight of hydrogen is 1.0080, the charge necessary to liberate one gram-atom of hydrogen in electrolysis is $9565 \times 1.0080 = 9642$ emu. If we assume that the charge on the electron numerically equals the charge on the hydrogen ion in electrolysis, Avogadro's number is

$$N = \frac{9642}{1.60 \times 10^{-20}} = 6.03 \times 10^{23} \qquad (7\text{–}17)$$

Knowing Avogadro's number we can determine the mass of the hydrogen atom. Since there are 6.03×10^{23} atoms in 1.0080 gm of hydrogen, the mass of each atom is 1.67×10^{-24} gm. Moreover, since the charge on the electron is the same as that on the hydrogen ion and since

e/n for the electron is 1840 times E/M for the hydrogen ion, the mass of the electron must be 1/1840 times the mass of the hydrogen ion or atom, so that

$$m = 9.1 \times 10^{-28} \text{ gm} \tag{7-18}$$

CHAPTER VII

PROBLEMS

1. A beam of electrons is deflected by means of a magnetic field whose magnitude is 10 emu and whose direction is perpendicular to that of the beam. If each electron in the beam is traveling with a speed of 4×10^9 cm/sec, determine the radius of the circle into which the beam is bent by the magnetic field.

2. Find the maximum velocity of fall of a raindrop whose diameter is 1 mm through the air of the atmosphere at 23° C. Look up the necessary constants in a book of tables.

3. Find the maximum rate of fall of an oil drop whose diameter is 0.001 mm and whose density is 1.8. If the drop carries a charge equal to that of three electrons, what potential must be applied between the plates A and B in Millikan's experiment in order to support the drop against gravity when the distance between plates is 1 mm? Look up the necessary constants in a book of tables. The temperature of the air between the plates is 0° C.

CHAPTER VIII

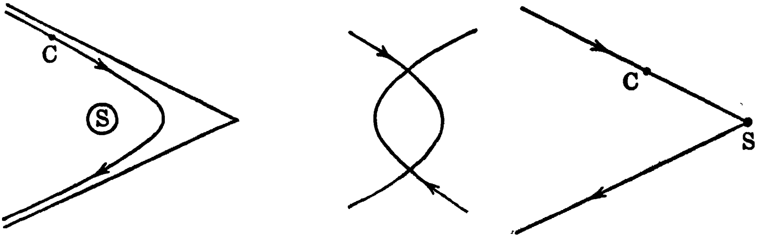
KINETIC THEORY OF GASES

8.1. The Three States of Matter.—All substances may exist in any of the three states—solid, liquid, or gas. We have already mentioned in Sec. 7.8, that compound substances consist of molecules and that molecules consist of atoms. According to the kinetic theory of matter, the molecules of a substance, such as water, are in motion and they attract each other. In the case of a solid, such as ice, each molecule is so tightly bound by the attractive forces to the surrounding molecules that it remains very nearly in constant position relative to the surrounding molecules, although it is allowed to vibrate within narrow limits about its average position. However, the forces between the molecules are not entirely attractive, because, on bringing the molecules too close together, we find that repulsive forces come into play, and that these repulsive forces are very great. Usually we describe the effect of these repulsive forces as being due to the "size" of the molecules. When one molecule is in "contact" with another, the two molecules cannot be made to approach more closely except by very large forces, and then the distance of approach is very small compared with the distance between centers of the two molecules. This is proved by the fact that solids and liquids are well-nigh incompressible—the molecules can only with great difficulty be made to penetrate one another. For many purposes we may think of the molecules as being hard elastic spheres. In a solid these spheres are held together by attractive forces and the spheres are in contact or almost in contact with each other.

It will be noted that in the previous paragraph the words "size" and "contact" were put in quotation marks. We shall see later that, according to modern views, an atom consists of a central positively charged nucleus around which a few electrons revolve, the structure of an atom resembling somewhat that of the solar system. There are vast stretches of empty space between the planets and the sun and between one planet and another, so that the solar system consists principally of empty space. Likewise, compared with the linear dimensions of the nucleus and of the electrons in an atom, the distances between the electrons and the nucleus and between one electron and another are large, so that an atom consists principally of empty space. Since molecules are made up of atoms, they too consist principally of empty space. However, when the centers of two molecules approach within a certain distance s of each other, the attractive forces suddenly change to repulsive forces, and s may be called the radius of the sphere of exclusion of a molecule. When the center of one molecule enters the sphere of exclusion of another molecule, large repulsive forces come into play, and the effect is the same as if two hard spheres were in contact.

Now consider what happens when a solid is heated. According to the kinetic theory the average amplitude of the vibrations of each molecule increases with rise of temperature. This increase of the amplitude means that on the average two adjacent molecules are further apart and so the substance expands. The expansion takes place against the attractive forces between the molecules. Finally, when a certain temperature is reached, the vibrations become sufficiently energetic for a particular molecule to break free from its surrounding molecules. However, after breaking free from one group of molecules, the molecule is almost immediately caught by the attractive forces of a second group, and so on. When this state of affairs occurs, the substance is in the liquid state. In this state the molecules are in contact or almost in contact with each other, but any two particular molecules do not stay bound together for any great length of time. On further heating, the liquid expands until the vibrations of the molecules become sufficiently energetic for each molecule to break completely free from all the other molecules. Each molecule flies hither and thither independently or almost independently of each other molecule. The substance is now in the gaseous or vapor state.

In the gaseous state, the molecules are flying about in random directions with all sorts of speeds. When liquid water changes to vapor at 100° C the volume of the vapor at a pressure of one atmosphere is 1600 times the volume of the water. The molecules are evidently much further apart than when the water was in the liquid state. In fact, the average distance between the center of one molecule and that of a molecule nearest to it is $\sqrt[3]{1600}$ or about 12 times the average distance apart of two such molecules in the liquid state. Because of the very much greater separation of the molecules in the gaseous state, the molecules are rarely close enough for the attractive forces to come into play, and they are still more rarely close enough for the repulsive forces to come into play. A collision occurs when one molecule of a gas hits another and rebounds. However, it is possible for an apparent collision to take place under the attractive forces of two molecules. A comet sometimes describes a path about the sun as in Fig. 8.1*a*. In this case the path is a hyperbola, the asymptotes of which are shown. If this phenomenon is observed from a great distance, it appears as in Fig. 8.1*b*. In this case the appearance is that of a collision. However, we shall not distinguish between real collisions where the molecules actually hit one another and those which are really due to the two molecules swinging around each other as in Fig. 8.1*c*.



(*a*) (*c*) (*b*)

FIG. 8.1.—Different types of collisions.

8.2. Simple Theory of the Pressure.—We have seen in Sec. 8.1 that, according to the kinetic theory, gaseous molecules are rushing about with random speeds in random directions, that each molecule has a size, and that there are attractive forces between the molecules. However, the effect of the size and of the attractive forces is very small at ordinary pressures. In this section we shall therefore assume that the molecules are points having no size and that there are no attractive forces. Let us assume that the gas is contained in a cubic box whose edges are of unit length. Such a box is shown in Fig. 8.2. Further, let us assume that each molecule of the gas has a mass m and that there are n molecules in the box. Since the molecules have all kinds of velocities, we shall assume that at a given instant of time molecule No. 1 has a velocity V_1, molecule No. 2 a velocity V_2, molecule No. 3 a velocity V_3, and so on for all the molecules. However, the directions of these velocities are at random. The best way of taking account of these

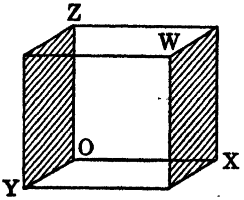


FIG. 8.2.

directions is to consider the components of the velocity of each molecule.

Let u_1, v_1, w_1 be the components of V_1 in the directions OX, OY, and OZ respectively and so on for the components of V_2, V_3, $\cdots$ V_n. We then have

$$V_1^2 = u_1^2 + v_1^2 + w_1^2$$

$$V_2^2 = u_2^2 + v_2^2 + w_2^2 \quad \text{and so on.} \tag{8-1}$$

Let us consider the motions of the molecules between the shaded faces YZ and WX in Fig. 8.2. Only those molecules which have their u's$\neq 0$ will hit the shaded faces. We shall further assume that when a molecule hits a shaded face of the cube the impact is elastic. The x-component of the velocity of the molecule then remains unchanged in magnitude but is reversed in sign, the y- and z-components remaining unchanged both in magnitude and sign. The time for molecule No. 1 to move from the face YZ to the face WX and back again is $2OX/u_1$. But $OX = 1$, so that the time is $2/u_1$. Hence the number of impacts of molecule No. 1 upon the face YZ per unit time is $u_1/2$. At each impact of molecule No. 1 on the face YZ the x-component of the velocity changes from $-u_1$ to $+u_1$, so that the change is $u_1 - (-u_1)$ or $2u_1$. Therefore the change in the x-component of the momentum is $2mu_1$. This change of momentum occurs each time molecule No. 1 hits face YZ, so that the total change of momentum experienced by molecule No. 1 at face YZ in one second is

$$2mu_1 \times \frac{u_1}{2} = mu_1^2 \tag{8-2}$$

But there are n molecules and each molecule in unit time experiences a change of momentum given by an equation similar to (8-2), the subscript of course being changed. Hence, taking all the n molecules into account, the total change of momentum occurring in unit time at face YZ is

$$m(u_1^2 + u_2^2 + \cdots + u_n^2) \tag{8-3}$$

Now, since (8-3) is the rate of change of momentum occurring at face YZ, a force towards the right must act on the face YZ, in order to keep the face from being pushed back to the left by the gas. Since the area of the face YZ is unity, this force is numerically equal to the pressure p. Hence

$$p = m(u_1^2 + u_2^2 + \cdots + u_n^2) \tag{8-4}$$

In the above we have considered the rate of change of momentum occurring at the face *YZ*. If also we consider the rate of change of momentum occurring at the faces *ZX* and *XY*, we shall obtain

$$p = m(v_1^2 + v_2^2 + \cdots + v_n^2) \tag{8-5}$$

and

$$p = m(w_1^2 + w_2^2 + \cdots + w_n^2) \tag{8-6}$$

respectively. Adding (8–4), (8–5) and (8–6) and making use of (8–1), we obtain

$$3p = m(V_1^2 + V_2^2 + \cdots + V_n^2) \tag{8-7}$$

Let us now introduce the mean square velocity C^2, such that

$$C^2 = \frac{(V_1^2 + V_2^2 + \cdots + V_n^2)}{n} \tag{8-8}$$

According to Sec. 2.6, C is the root mean square of the velocities V_1, V_2, etc. We now write (8–7) in the form

$$3p = nmC^2 \tag{8-9}$$

But, if m is the mass of each molecule and if there are n molecules in 1 cm³ of the gas, the quantity nm is the mass of unit volume of the gas. The quantity nm is thus ρ the density of the gas. Hence the pressure is given by

$$p = (1/3)\rho C^2 \tag{8-10}$$

If we take a mass M of the gas occupying a volume v, then, since $\rho = M/v$, (8–10) may be written

$$pv = (1/3)MC^2 \tag{8-11}$$

If the root mean square velocity of the molecules of the gas depends upon the temperature alone, the right side of (8–11) becomes a constant for a given mass M of a gas at constant temperature. We thus arrive at Boyle's law, which states that the product of the pressure and volume of a given mass of gas at a given temperature is a constant.

From (8–10) it is seen that, if we know or can measure p and ρ for a gas at a given temperature, we can calculate the root mean square velocity of the molecules of the gas. At 0° C and a pressure of 76 cm of mercury the density of air is 0.0013 gm/cm³. Now, it must be remembered that the pressure p in (8–10) is in absolute units, that is, in dynes/cm^2. One atmosphere in these units is

$$76 \times 13.6 \times 980 = 1{,}013{,}000 \text{ dynes/cm}^2$$

since the density of mercury is 13.6 gm/cm^3 and the acceleration of gravity is 980 cm/sec^2. This value of the pressure of one atmosphere in absolute units is useful to remember. We therefore obtain

$$C = \sqrt{\frac{3p}{\rho}} = \sqrt{\frac{3 \times 1{,}013{,}000}{0.0013}} = 48{,}400 \text{ cm/sec}$$

which is a speed of about one-third of a mile per second. The molecules are thus rushing about with very high speeds. It must be remembered that this is the root mean square speed and that some molecules are traveling with greater and others with smaller speeds than this.

8.3. Perfect Gas Equation.—If the pressure, volume, and temperature of any real gas are varied, these quantities very nearly obey the relation

$$pv = RT \tag{8-12}$$

where T is the absolute temperature and R is known as the gas constant. A perfect gas is defined as one which exactly obeys the relation (8-12), and so this relation is called the Perfect Gas Equation. In the kinetic theory a perfect gas is one whose molecules have no size and whose molecules exert no attractive forces on each other. The behavior of real gases departs from (8-12), but the departure in the case of gases far removed in temperature from their liquefaction point is small.

The equations derived in Sec. 8.2 apply to the case of a perfect gas. Equating the right sides of (8-11) and (8-12), we obtain

$$MC^2 = 3RT \tag{8-13}$$

The value of R depends upon the mass and nature of the gas being considered. If, however, a gram-molecule of a gas is taken, it is found that the value of R for all gases is very nearly the same. It has been found that the volume occupied by a gram-molecule of any gas at a temperature of 0° C and a pressure of 76 cm of mercury is very nearly 22,400 cm^3. Hence, in virtue of (8-12), the gas constant for a gram-molecule of gas is

$$R = \frac{1{,}013{,}000 \times 22{,}400}{273} = 8.31 \times 10^7 \text{ erg/deg} \tag{8-14}$$

If M in (8–13) is the mass of a gram-molecule of gas, we see from (8–13) that, since R has the same value for a gram-molecule of any gas, MC^2 is a constant for all gases at the same temperature. But, according to Sec. 7.8, a gram-molecule contains N molecules, where N is Avogadro's number, so that $M = Nm$, where m is the mass of each molecule. Hence (8–13) may be written

$$(\tfrac{1}{2})mC^2 = \frac{3RT}{2N} = \text{const} \times T \qquad (8\text{–}15)$$

Hence the kinetic theory in conjunction with the gas equation requires that the average kinetic energy of the molecules of a gas is the same as that of a second gas provided that the two gases are at the same temperature. If the gases are at different temperatures, the ratio of these temperatures equals the ratio of the average kinetic energies of the molecules of the respective gases.

From (8–13), the value of C for any gas is given by

$$C = \sqrt{3RT/M} \qquad (8\text{–}16)$$

where R is the gas constant for a gram molecule of the gas and M is the molecular weight of the gas.

8.4. Boltzmann's Constant.—If for R/N in (8–15) we write k we obtain a constant since both R and N are constant. We may look upon k as the gas constant for a single molecule. The importance of k is such that it is given a name, being called Boltzmann's constant. In terms of this constant the average kinetic energy of a molecule of any gas at temperature T is

$$(\tfrac{1}{2})mC^2 = (\tfrac{3}{2})kT \qquad (8\text{–}17)$$

We have already mentioned the value of N in Sec. 7.8. Using this value of N, Boltzmann's constant has a value

$$k = \frac{R}{N} = \frac{8.31 \times 10^7}{(6.03 \times 10^{23})} = 1.38 \times 10^{-16}\ \text{erg/deg} \qquad (8\text{–}18)$$

8.5. The Electron-Volt.—A new unit of energy has recently come into common use. This unit of energy is the electron-volt. It is the amount of energy equal to the kinetic energy acquired by an electron when it falls through a potential of one volt. When a particle or a vibration has an energy equal to the kinetic energy of an electron which has fallen through V volts, the particle or vibration is said to have an energy of V electron-volts. The kinetic energy acquired by an electron with a charge e' esu when it falls through V' esu of potential is $e'V'$ erg. From (5–33), 1 volt = 1/300 esu of potential, so that

$$\begin{aligned} 1 \text{ electron-volt} &= (1/300) \times 4.8 \times 10^{-10} \\ &= 1.6 \times 10^{-12} \text{ erg} \end{aligned} \tag{8–19}$$

It is sometimes useful to express Boltzmann's constant k in electron-volts. From (8–18) and (8–19), we have

$$\begin{aligned} k &= (1.38 \times 10^{-16})/(1.6 \times 10^{-12}) \\ &= 8.63 \times 10^{-5} \text{ ev/deg} \end{aligned} \tag{8–20}$$

where ev is the symbol for electron-volt. A gas molecule at a room temperature of 22° C has an average energy of $(3/2) \times 8.63 \times 10^{-5} \times 295 = 0.037$ ev.

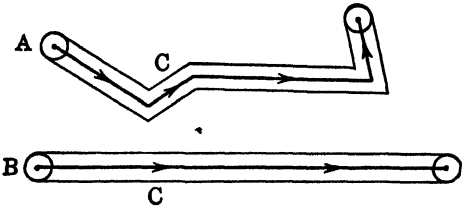


FIG. 8.3.—Successive collisions of a given molecule.

8.6. Mean Free Path.—In real gases the molecules have size and exert attractive forces on each other in contradiction to the assumptions made in Sec. 8.2 and used in developing formula (8–9) for the pressure. In Sec. 8.1 we have already spoken of the radius s of the sphere of exclusion of a molecule. Whenever the center of one molecule approaches within a distance s of another molecule, there is a collision. As a given molecule moves about, it collides with other molecules. The average distance which a molecule travels between collisions is called the mean free path, and the average number of collisions which it makes per second with other molecules is called the collision frequency.

Consider the given molecule as moving with the root mean square velocity C. The sphere of exclusion associated with the given molecule moves forward with velocity C and in one second carves out a space as shown at A in Fig. 8.3. The kinks occur where a collision takes place. When this space is straightened out, a space as shown at B is obtained. If the center of any other molecule is in this space, a collision occurs. The number of collisions which occur in one second is equal to the number of other molecules whose centers are in the volume $\pi s^2 C$. This number is $\pi n s^2 C$, since the number of molecules per unit volume is n. The collision frequency is thus $\pi n s^2 C$, and the mean distance between collisions is L, where

$$L = \frac{C}{\pi n s^2 C} = \frac{1}{\pi n s^2} \tag{8-21}$$

L is the value of the mean free path.

In this derivation we have treated the given molecule as moving with the velocity C and all the other molecules in the volume $\pi s^2 C$ as being at rest. However, the other molecules are neither at rest, nor, being in motion, are they all moving with the same velocity C. Taking these facts into account, the mean free path is given by

$$L = \frac{1}{\sqrt{2}\pi n s^2} \tag{8-22}$$

The mean free path which we have been discussing is that for a molecule moving among molecules of the same kind. In certain modern experiments, however, a stream of electrons is shot into a gas and it is desired to know the mean free path of an electron in the gas. This is given by

$$L_E = 4\sqrt{2} L_G \tag{8-23}$$

where L_G is the mean free path of a molecule of the gas as given by (8–22), and L_E is the mean free path of an electron in the gas.

We have seen in Sec. 8.2 that the root mean square speed of the molecules of a gas is large, so that, if, for instance, a quantity of ammonia gas (NH_3) is released at a point in a room, the ammonia should almost immediately diffuse throughout the room. Actually, if there are no draughts in the room, it is found that considerable time elapses before the ammonia becomes evident in the remote parts of the room. The slow rate of diffusion is due to the fact that the mean free path of the ammonia molecules among the air molecules is very small.

When different parts of a liquid or gas are moving relatively to each other, frictional forces come into play, and these forces tend to destroy the relative motion. This phenomenon is called viscosity. In the case of a wind blowing over the ground, the layer of gas in contact with the ground is at rest, and other layers slide over it and over each other.

Molecules of the gas diffuse from one layer to another. The gas molecules are moving in random directions with high velocities but they do not proceed very far before a collision occurs. In the case of a wind toward the right, the velocity of each molecule has vectorially added to it a velocity to the right equal to the velocity of the wind. We shall distinguish these two velocities by the terms gas-kinetic velocity and wind velocity. The gas-kinetic velocity of an air molecule is of the order of ⅓ mi/sec, while a wind velocity of more than 60 mi/hr or 88 ft/sec is unusual. At the layer *C* in Fig. 8.4 the wind velocity is v_1 and at the layer *D* the wind velocity is v_2, where $v_2 > v_1$. As a molecule moves up through the plane *E* it carries with it a wind velocity which is appropriate to the level from which it came. The molecule arrives among other molecules whose wind velocities are greater than its own wind velocity and so exerts a dragging effect upon the other molecules. On the other hand molecules are also diffusing from the space above *E* down through *E* into the space below. One of these molecules, upon arriving among other molecules whose wind velocities are less than its own wind velocity, exerts an accelerating effect upon the other molecules. On the average a single molecule has the wind velocity appropriate to the place of its last collision. It is evident, therefore, that the viscosity will increase with the mean free path. Applying these general principles, Maxwell, in 1860, obtained the relation

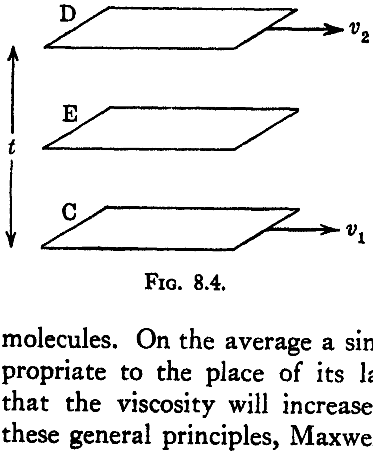


FIG. 8.4.

$$\eta = (1/3)\rho CL \tag{8-24}$$

where η is the coefficient of viscosity. However, more recent analysis by Chapman of London over the period 1912 through 1918 has shown that the factor ⅓ should be replaced by 0.499.

The coefficient of viscosity can be determined experimentally by suspending a horizontal disc by means of a wire. The disc is set into rotational oscillations about the vertical wire as axis. A second horizontal disc is placed just below the oscillating disc. The second disc is fixed so that the frictional effect of the air between the discs comes into play and the oscillations of the upper disc are damped. From the damping, η can be calculated. From (8-24) it is seen that η for a given temperature varies as ρL. But, according to (8–22), L varies inversely as ρ, so that ρL is constant when the pressure is varied. Hence η is independent of the pressure. Experiments show this to be true over a wide range of pressures. The value of η for air at 0° C is .00017. Substituting in (8–24) and using the known values of ρ and C for air, we find that the mean free path in air at 0° C and 76 cm is 0.8×10^{-5} cm. Owing to the uncertainty of the factor $\frac{1}{3}$, this result must be taken as giving the order of magnitude.

According to (8–22), L varies inversely as n, the number of molecules in a cubic centimeter. But, according to (8–9), n varies as the pressure p, so that the mean free path varies inversely as the pressure. Hence the mean free path at a pressure of 1/100 mm of mercury is $76{,}000 \times 0.8 \times 10^{-5}$ or 0.61 cm, and the mean free path of an electron in air at this pressure, according to (8–23), is of the order $4\sqrt{2} \times 0.61$ or 3.4 cm. In Chapter VII the appearance of the electrical discharge in a tube at low pressure was described. We saw that at a pressure of 1/100 mm the cathode rays, which are made up of swiftly moving electrons, extend throughout the length of the tube. This is explained by the fact that the mean free path of the electrons is comparable with the length of the tube. In modern vacuum technique, pressures as low as 1/10,000 and 1/100,000 mm are common. At these pressures the mean free paths of air molecules are 76 and 760 cm respectively. However, even at these low pressures the number of molecules is still very large. At 1/100,000 mm pressure, the number of molecules at 0° C is 3.6×10^{11} per cubic centimeter.

8.7. Van der Waals' Equation.—It is found that although all gases approximately obey the relation

$$pv = RT \tag{8–25}$$

yet no gas exactly follows this law. Van der Waals found that real gases more nearly obey the relation

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT \tag{8–26}$$

where a and b are constants characteristic of each gas. The perfect gas equation as we have seen in Secs. 8.2 and 8.3 can be derived on the assumptions that the molecules have no size and that they exert no attractive forces on each other, together with the assumption that the average kinetic energy of a molecule of a gas is proportional to the absolute temperature.

However, molecules do have size. The available volume in the box of Fig. 8.2 in which a molecule can move about is not the geometrical volume of the box but the geometrical volume minus the sum of the volumes of the other molecules. We may call this sum the total volume of exclusion. Since the center of a given molecule cannot approach closer than s to the center of another molecule, the total volume of exclusion is $(n - 1)4\pi s^3/3$. But n is very large, so that we can replace $n - 1$ by n. Thus the volume from which a particular molecule is excluded may be written $4\pi n s^3/3$. However, when molecule 1 is excluded from the sphere of exclusion surrounding molecule 2, molecule 2 is also excluded from the sphere of exclusion surrounding molecule 1. Hence in calculating the total volume of exclusion as $4\pi n s^3/3$, we have counted each molecule twice and the correct total volume of exclusion is one-half of $4\pi n s^3/3$. We shall represent the total volume of exclusion by b, so that

$$b = \frac{2\pi n s^3}{3} \tag{8-27}$$

Hence the available volume for each molecule is not v, the volume of containing vessel, but $(v - b)$. Hence (8-25) becomes

$$p(v - b) = RT \tag{8-28}$$

Let us suppose that attractive forces come into play when the centers of two molecules approach each other within a distance D and that repulsive forces predominate when the centers approach each other within a smaller distance s. There is thus a certain space surrounding each molecule such that, if the center of another molecule enters this space, attractive forces come into play. In the central portions of the box in Fig. 8.2 a molecule is attracted equally in all directions, so that the total attraction is zero. However, as a molecule approaches the wall of the containing vessel, the part of the space of attraction which is between the center of the molecule and the wall contains less molecules than the part behind the molecule, so that the molecule experiences a pull back into the gas as it approaches the wall. Consequently the velocity with which the molecule hits the wall is less than the velocity when there are no attractive forces between the molecules of the gas. The pressure is therefore less than that calculated in Sec. 8.2. Hence the experimental value of the pressure must be increased by a certain amount x in order to conform with the theoretical value of p found in Sec. 8.2. The quantity x may be called the deficiency of pressure. Now, the backward pull on a gas molecule as it approaches the wall of the containing vessel depends upon the number of molecules in the part of the space of attraction which is behind the molecule and this number is proportional to the number of molecules per unit volume of the gas. When the volume of a given mass of gas is halved, the backward pull on each molecule approaching the wall is doubled, and also the number of molecules hitting unit area of the wall in unit time is doubled. Therefore the deficiency of pressure is quadrupled when the volume of the gas is halved. Hence x varies inversely as v^2, so that, putting $x = a/v^2$, where a is a constant of proportionality, and adding x to p in (8–28), we obtain

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT \tag{8–29}$$

which is Van der Waals' equation.

8.8. Digression on Population Statistics.—So far we have made use only of the root mean square velocity of the molecules of a gas. However, it is necessary in many problems that not only the average velocity but also the distribution of the velocities of the molecules be known. In such problems, we need to know the number or rather the fraction of the molecules whose speeds are, say, between 20,000 and 30,000 cm/sec or perhaps we need to know the fraction of the molecules whose speeds are greater than a certain amount.

A good example of a distribution function is that for the age distribution of people in the United States. In this case the function is represented by the curve in Fig. 8.5. From this curve, for instance, the fraction of individuals between the ages of 10 and 11 years or the fraction between the ages of 25 and 35 years can be found. The reader will note that in order to determine the fraction a certain range must be given. In everyday language an age of 25 years usually means an age somewhere between 25 and 26 years, but in exact mathematical language an age of 25 years is taken to be an age of 25 years no days no hours no minutes and no seconds—in fact, exactly 25 years. In this sense the fraction of individuals whose ages are exactly 25 years is zero and is meaningless. On the other hand, the fraction of individuals whose ages are between 24.99 and 25.01 years is quite definite, because here we have given a range. Instead of speaking of the fraction between 24.99 and 25.01 years, we shall for brevity speak of the fraction within a range of 0.02 year at 25 years, or still more briefly, the fraction in 0.02 year at 25 years. In symbols, if y represents years of age, we shall speak of the fraction in dy at y. Now, if the range dy is small, the fraction will also be small. Let us represent this small fraction by dp. However, although both dp and dy are small, the ratio dp/dy is not small. The fraction of the population between 24.99 and 25.01 years is 0.0356 per cent or 0.000356 expressed as a fraction, and in this case

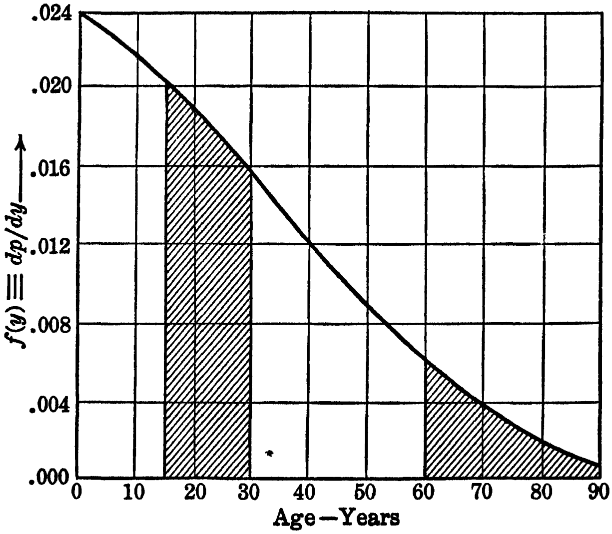


FIG. 8.5.—Graph showing the age distribution of the people in the United States.

$$\frac{dp}{dy} = \frac{0.000356}{0.02} = 0.0178 \tag{8-30}$$

This is the value of dp/dy for an age of 25 years. If values of dp/dy are found for other ages and a curve is plotted between dp/dy and y the age, the curve of Fig. 8.5 is obtained. If the curve can be expressed as a function $f(y)$, this function is called the distribution function for the ages of individuals. It is interesting that the particular individuals whose ages are between 24.99 and 25.01 are constantly changing as time goes on, yet it is a law of statistics that the number within the range remains practically constant.

By definition,

$$f(y) = \frac{dp}{dy} \tag{8-31}$$

so that the fraction of the population in a range dy at y years is

$$dp = f(y)dy \tag{8-32}$$

If, however, the range is large such as from 15 to 30 years, the fraction of individuals whose ages are within this range is p, where

$$p = \int_{y=15}^{y=30} f(y)dy \tag{8-33}$$

If $f(y)$ is expressed as a curve, the value of the integral is represented by the shaded area between the ordinates at $y = 15$ and $y = 30$ in Fig. 8.5. Similarly the proportion of individuals whose ages are above 60 years is

$$p = \int_{y=60}^{y=\infty} f(y)dy \tag{8-34}$$

and is represented by the shaded area to the right of the ordinate at $y = 60$. The total area under the curve in Fig. 8.5 must be unity since it must represent the total population. Likewise the integral of any distribution function over the whole range must be unity.

It should be noted that in all statistical problems of the type just described the range dy must be neither too small nor too large. If the time of birth could be measured accurately to one minute, the number of people of the 130 millions in the United States whose ages are within a range of 10 minutes at 25 years is 43 on the average. However, the number might easily fluctuate from zero to 100 for successive intervals of 10 minutes. In such a case, dy is too small. The criterion for the right size of dy is that the number of people with ages between y and $y + dy$ is very nearly proportional to dy. Large numbers must be used in statistical problems.

8.9. Distribution of Component Velocities.—The molecules of a gas are continually colliding with each other. At each impact the speed and component velocity of a given molecule are changed. Considering only the component velocity in the x-direction, the molecule at one impact receives an impulse or bump to the left, while at another impact it receives an impulse or bump to the right. Hence at any instant of time the x-component velocity is the result of many previous bumps to the left or right. It is reasonable to assume that the chance of an impulse to the right is equal to that of an impulse to the left. The problem is then similar to a problem in the throwing of coins.

Let there be n coins, each of which has one face marked $+g$ and the other face marked $-g$. The coins are thrown and the result of the throw is r coins showing $+g$ and $(n - r)$ coins showing $-g$. The algebraic sum of the numbers showing is $rg - (n - r)g = (2r - n)g$. If n is an even number, the values which can be thrown are 0, $\pm 2g$, $\pm 4g$, $\cdots$, and $\pm ng$. In the theory of probability it is shown that the chance of throwing $(2r - n)g$ is given by the $(r + 1)$th term in the binomial expansion of $(a + b)^n$, where $a = b = \frac{1}{2}$, since the chance of one coin showing $+g$ is the same as that for it showing $-g$. The probabilities for 8 coins are given by the various terms of $(a + b)^8$ as shown in Table 8.1. If we plot the probability against the value thrown we obtain the crosses shown in Fig. 8.6. The square of each value is shown in the third column of Table 8.1. To obtain the mean square the square of each value is multiplied by the probability of throwing the value and the resultant quantities added as in the fourth column of Table 8.1. It is seen that the mean square is $7.984g^2$. When n is made large and g is made small in such a way that the mean square remains the same, it is shown in the theory of probability that the chance of throwing a value between x and $x + dx$ is $f(x)dx$, where

TABLE 8.1

Value Thrown	Probability	Square of Value Thrown	To Obtain Mean Square
$-8g$	$1/2^8 = .004$	$64g^2$	$.256g^2$
$-6g$	$8/2^8 = .031$	$36g^2$	$1.116g^2$
$-4g$	$28/2^8 = .109$	$16g^2$	$1.744g^2$
$-2g$	$56/2^8 = .219$	$4g^2$	$.876g^2$
$0g$	$70/2^8 = .273$	$0g^2$	$0g^2$
$2g$	$56/2^8 = .219$	$4g^2$	$.876g^2$
$4g$	$28/2^8 = .109$	$16g^2$	$1.744g^2$
$6g$	$8/2^8 = .031$	$36g^2$	$1.116g^2$
$8g$	$1/2^8 = .004$	$64g^2$	$.256g^2$
		Mean Square	$7.984g^2$

$$f(x) = 1/\sqrt{2\pi\overline{x^2}}\ \exp\left(-x^2/2\overline{x^2}\right) \qquad (8\text{–}35)$$

and $\overline{x^2}$ is the mean square value. The curve shown in Fig. 8.6 is $f(x) \times$

$2g$ plotted against x for the case where $\overline{x^2} = 7.984g^2$ and the range dx is taken as $2g$ because in the case of coins the values jump by $2g$. It is seen how well the points plotted from Table 8.1 fall upon the curve.

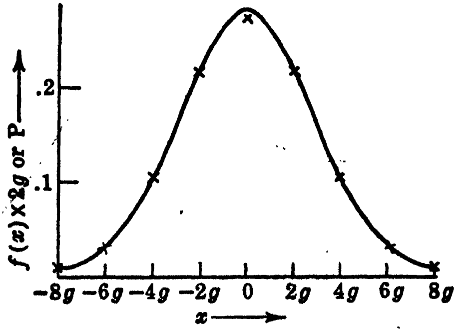


FIG. 8.6.

By reasoning similar to that for the coins it can be shown that the chance of a molecule having an x-component velocity between u and $u + du$ is $f(u)du$, where

$$f(u) = (1/\alpha\sqrt{\pi}) \exp(-u^2/\alpha^2) \tag{8-36}$$

and α^2 is twice the mean square x-component velocity. Not only is $f(u)du$ the chance of a single molecule having an x-component velocity u in the range du but it is also the fraction of all the molecules having their x-component velocities in this same range at any instant of time. Eq. (8-36) is Maxwell's distribution function for component velocities. The shape of the distribution curve is shown in Figs. 8.6 and 8.8. In Fig. 8.8 the abscissa $x = 1$ corresponds to $u = \alpha$.

8.10. Distribution of Speeds.—It is also important to know the distribution of speeds irrespective of direction. Maxwell's distribution function for this case is $F(V)$, where

$$F(V) = (4/\alpha^3\sqrt{\pi})V^2 \exp(-V^2/\alpha^2) \tag{8-37}$$

The fractions of molecules whose speeds are in the range dV at a speed V is $F(V)dV$. The shape of the distribution curve for speeds is shown in Fig. 8.9. It is to be noted that this curve does not extend to negative values of V. This is because we are not interested in the direction of V.

8.11. Most Probable Speed.—In Fig. 8.9, the curve rises to a maximum. The value of V for which $F(V)$ is a maximum is called the most probable speed, for the fraction having their speeds in a given small range dV at this velocity is greater than at any other velocity. We have shown in Sec. 2.5 that the value of x for which a function $f(x)$ is a maximum is found by differentiating $f(x)$ with respect to x, putting the derivative equal to 0, and then solving for x. This rule has already been applied to $F(V)$ in Sec. 2.5, where we have shown in (2–59) et seq. that $F(V)$ is a maximum when

$$V = \alpha \tag{8–38}$$

Hence α in the preceding formulas is the most probable velocity. The relation between α, the most probable velocity, and C, the root mean square velocity, can be shown by means which are beyond the scope of this book to be

$$\alpha = C\sqrt{2/3} = 0.816 \times C \tag{8–39}$$

In Sec. 8.2 we have calculated that C for air at 0° C is 48,400 cm/sec, so that α for air at this temperature is 39,400 cm/sec. In Fig. 8.9, the abscissa $x = 1$ corresponds to $V \mp \alpha$.

From (8–16) and (8–39), the value for α for any gas is given by

$$\alpha = \sqrt{2RT/M} \tag{8–40}$$

where R is the gas constant for a gram-molecule of the gas and M is its molecular weight.

8.12. Velocity of Escape.—A projectile which is fired from the surface of the earth with a velocity having an upward component will leave the earth if its speed is greater than a certain critical velocity. This critical velocity is known as the velocity of escape. We shall consider the simple case of firing the projectile directly upwards. The kinetic energy of the projectile must be greater than the work done in removing the projectile from the earth against the attraction of gravity. The pull of gravity on the projectile varies inversely as the square of the distance from the center of the earth, so that at a distance r shown in Fig. 8.7 the pull is mga^2/r^2, where a is the radius of the earth, m is the mass of the projectile and g is the acceleration of gravity at the

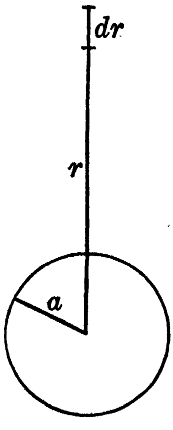


FIG. 8.7.

surface of the earth. As the projectile passes over the small distance dr the work done is $(mga^2/r^2)dr$, so that the total work from $r = a$ to $r = \infty$ is

$$\int_{r=a}^{r=\infty} \left(\frac{mga^2}{r^2}\right) dr = -mga^2\left[\frac{1}{r}\right]_a^\infty = mga \qquad (8\text{–}41)$$

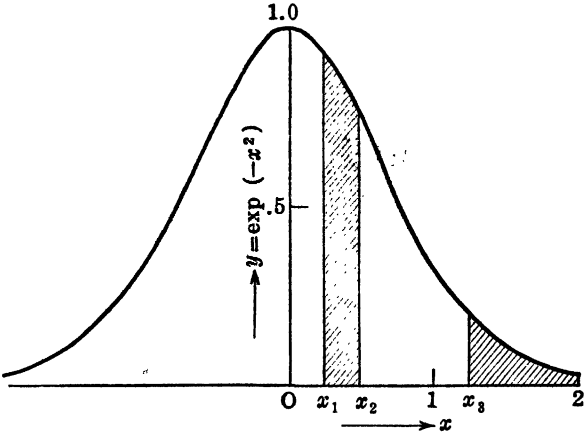


FIG. 8.8.—Graph of $y = \exp(-x^2)$.

Equating this to $(\frac{1}{2})mV_E^2$, where V_E is the velocity of escape, we obtain

$$V_E = \sqrt{2ga} \qquad (8\text{–}42)$$

For the earth a = 4000 miles or 2.1×10^7 ft and g = 32 ft/sec^2, so that V_E = 36,700 ft/sec or approximately 7 mi/sec.

The most probable speed for air molecules at 0° C is 39,400 cm/sec. A velocity of 7 mi/sec = 1.1×10^6 cm/sec. In Fig. 8.9, the abscissa $x = 1$ corresponds to the velocity 39,400 cm/sec, so that the value of x for the velocity of escape is $1.1 \times 10^6/39{,}400 = 28$. From Fig. 8.9 it is therefore seen that the fraction of molecules having sufficient velocity to escape is exceedingly small. It is for this reason that the earth retains its atmosphere. A planet whose mass is too small cannot retain an atmosphere. The velocity of escape from the moon is 2.4×10^5 cm/sec. The fraction of molecules in an atmosphere at 0° C on the moon which have sufficient velocity to escape, although small, is not excessively small, so that the moon has lost its atmosphere long ago.

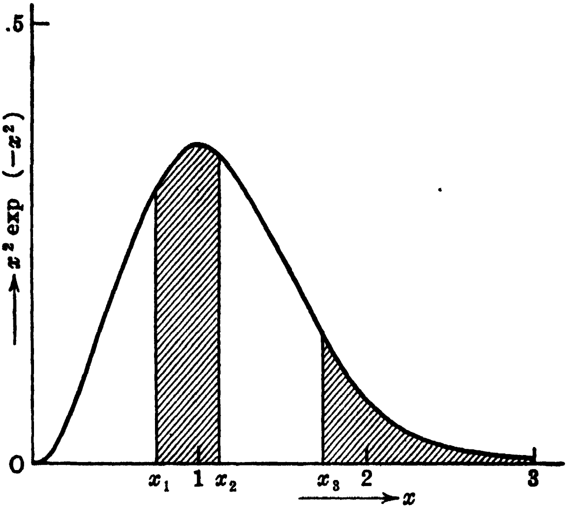


FIG. 8.9.—Graph of $y = x^2 \exp(-x^2)$.

8.13. Illustrative Problems.—Perhaps the best way to teach the real meaning of the distribution functions we have been considering is by means of problems.

Problem 1. Find the fraction of the molecules of ammonia gas (NH_3) at 100° having component velocities in the upward direction (*a*) between 25,000 and 40,000 cm/sec, and (*b*) greater than 70,000 cm/sec.

Let us write x for u/α or for V/α according as we are interested in component velocities or speeds respectively. The fractions $f(u)du$ and $F(V)dV$ now become $f(x)dx$ and $F(x)dx$ respectively, where from (8–36) and (8–37)

$$f(x) = (1/\sqrt{\pi}) \exp(-x^2) \tag{8–43}$$

and

$$F(x) = (4/\sqrt{\pi})x^2 \exp(-x^2) \tag{8–44}$$

In the following we do not need to use the multiplying constants $1/\sqrt{\pi}$ and $4/\sqrt{\pi}$, so we plot the two curves $y = \exp(-x^2)$ and $y = x^2 \exp(-x^2)$ as shown in Figs. 8.8 and 8.9 respectively. The molecular weight of NH_3 is $14 + 3 \times 1 = 17$, so that, from (8–40), $\alpha = 60{,}500$ cm/sec for a temperature of 373° Abs. When $u = 25{,}000$, $u/\alpha = 25{,}000/60{,}500$

= .43, so that x_1 = .43. Similarly x_2 = 40,000/60,500 = 0.66. The shaded area is obtained as shown in Fig. 8.8. The reader should plot the curve in Fig. 8.8 and should count the squares in the shaded portion between x_1 and x_2. Then he should count the squares under the whole curve in Fig. 8.8. Dividing the area of the shaded portion between x_1 and x_2 by the total area under the whole curve, we obtain the fraction of the molecules for which x lies between .43 and .66. Because we obtain the fraction by dividing the area of the shaded portion by the whole area the multiplying constant cancels out, and thus we do not need to know it.

When $u = 70{,}000$, $u/\alpha = 70{,}000/60{,}500 = 1.16$, so that $x_3 = 1.16$. Counting squares in the shaded portion to the right of x_3 in Fig. 8.8, the fraction of the molecules for which is x greater than 1.16 is the ratio of the area to the right of x_3 to the area under the whole curve.

Problem 2. Find the fraction of molecules in hydrogen gas at −200° C having speeds (*a*) between 50,000 and 100,000 cm/sec, and (*b*) greater than 120,000 cm/sec.

A molecule of hydrogen gas consists of two atoms of hydrogen as is represented by the chemical formula H_2, so that the molecular weight of hydrogen is 2. Hence, from (8–40), α = 78,000 cm/sec at 73° Abs. We now obtain x_1 = 50,000/78,000 = 0.64, x_2 = 100,000/78,000 = 1.28, and x_3 = 120,000/78,000 = 1.54. The reader should plot the curve in Fig. 8.9 and should count the squares in the shaded portions and under the whole curve as in Problem 1. From the proper areas the fractions can be calculated. This is left to the reader to work out in detail.

8.14. Concentration Law.—Boltzmann has shown that, if the potential energy of a molecule at a certain place P in a gas is w_P and at another place Q is w_Q, the ratio of the concentrations at P and Q is

$$\left(\frac{n_P}{n_Q}\right) = \exp\left\{-\frac{(w_P - w_Q)}{kT}\right\} = \exp\left(-\frac{w}{kT}\right) \qquad (8\text{–}45)$$

where w is the difference between the potential energy of a molecule at P and that of a molecule at Q, k is Boltzmann's constant, and T is the temperature of the gas. It is assumed that T is constant throughout the gas. As an example, consider two points P and Q at different levels in the atmosphere. If P is at a height h above Q, the difference between the potential energy of a molecule at P and that of a molecule at Q is mgh, where m is the mass of a molecule. The ratio of the concentrations is then

$$\frac{n_P}{n_Q} = \exp\left(-\frac{mgh}{kT}\right) \tag{8-46}$$

The density of the earth's atmosphere decreases with the height according to this law so long as the temperature remains constant with the height. The concentration law is sometimes called the law of atmospheres.

Suppose we ascend in the atmosphere to a height h_1 where $n_P/n_Q = \frac{1}{2}$. Then, since $\exp(-x) = \frac{1}{2}$ when $x = .693$ (see tables of $\exp(-x)$ or e^{-x}),

$$\frac{mgh_1}{kT} = 0.693$$

or

$$k = \frac{mgh_1}{0.693T} \tag{8-47}$$

But, according to Sec. 8.4, $k = R/N$, so that

$$N = \frac{0.693RT}{mgh_1} \tag{8-48}$$

The value of N, Avogadro's number, can therefore be found, provided that m and h_1 can be measured.

8.15. Brownian Movements.—In 1827 Brown, an English botanist, found that fine particles in aqueous suspension appeared to be dancing about with all sorts of random motions when these particles were observed by means of a microscope. It was proposed and is now believed that these dancing particles are merely partaking of the general heat motions of the liquid in which they are suspended. The particles act like huge gas molecules and should obey the laws of gases. In particular, they should obey Boltzmann's concentration law. Since m the mass of a Brownian particle is large compared with the mass of an air molecule, it may be that h_1 in (8–47) and (8–48) can be made very small, say, a millimeter or so. The value of h_1 for air at 0° C is about 3.5 miles.

In 1908 Perrin tested this theory of the Brownian movements. He used colloidal suspensions of gum mastic and of gum gamboge. In order to test whether the particles obey the concentration law, it is necessary to have all the particles of the same size and mass. Perrin accomplished this by means of fractional centrifuging. Having obtained particles all of the same size, it was then necessary to measure the mass of each particle. The density of the particles was found by determining that strength of salt solution for which the density of the particles and that of the solution is the same. If the densities were the same, the particles would not separate out when the solution was violently centrifuged. The volume of one of these small particles cannot be determined by observation with the microscope because the particles are so small that diffraction effects are produced. Perrin therefore determined the volume by evaporating a dilute suspension of the particles on a microscope slide, when the effect of surface tension is such as to cause the particles to run together in rows. The number in a row was counted, and the length of a row was measured, and so the diameter of the spherical particles was found and the volume determined. Knowing the density the mass was immediately found.

Perrin placed the particles in small glass cells kept at constant temperature. A cell was placed on the stage of a microscope. What is known as the depth of focus of a microscope is very sharp. Only particles within this depth can be seen. If the microscope is raised or lowered different numbers of particles come into the field of view. Perrin found that the number of particles observed in the microscope decreased according to Boltzmann's concentration law as the microscope was raised. In one case he found the concentration of an "atmosphere" of particles decreased to one-half in a height of 0.03 mm.

In Perrin's experiments the particles whose density is D are suspended in a liquid whose density is d. The effective weight of each particle is less than mg because of the buoyant force of the liquid. The effective weight is $mg(D - d)/D$, and mg in (8–48) is replaced by this expression, so that

$$N = \frac{0.693RTD}{mgh_1\,(D - d)} \tag{8–49}$$

Perrin counted thousands of particles at different levels and finally obtained a value of N between 6.5×10^{23} and 7.2×10^{23} with an average of 6.85×10^{23}.

We have already mentioned in Sec. 7.8 that Avogadro's number can be obtained from the values of the charge on the electron and of the charge necessary to liberate one gram-atom of hydrogen in electrolysis. The value of Avogadro's number so found is 6.03×10^{23} which is in good agreement with that found by Perrin.

CHAPTER VIII

PROBLEMS

1. The density of oxygen gas at 20° C and a pressure of 70 cm of mercury is .00123 gm/cm^3. Find the root mean square velocity of oxygen molecules at 20° C.

2. The density of hydrogen gas at 100° C and a pressure of 120 cm of mercury is .000104 gm/cm^3. Find the root mean square and the most probable velocity of hydrogen molecules at 100° C.

3. The radii of the earth and Mars are 3960 and 2100 miles respectively, and the mass of Mars is 0.106 that of the earth. Find the velocity of escape from the surface of Mars.

4. Ceres is the largest asteroid (see Chapter XXIII). It has a diameter of 480 miles and a mass of about 1/8000 that of the earth. Given that the radius of the earth is 3960 miles, find the velocity of escape from the surface of Ceres.

5. Imagine the asteroid Ceres (see Problem 6) to be surrounded by an atmosphere of argon at $-50°$ C. Argon is a monatomic gas of atomic weight 40. Determine the fraction of argon molecules which at any instant of time are moving upwards with a velocity more than sufficient to escape from the asteroid.

6. Carbon dioxide gas has the chemical formula CO_2. Find the fraction of molecules of CO_2 which at 70° C have total speeds (not component velocities) between (*a*) 10,000 and 20,000 cm/sec and (*b*) 40,000 and 60,000 cm/sec.

7. Find the volume of 5 grams of methane gas (CH_4) at a temperature of 150° C and a pressure of 40 cm. Determine the number of molecules in this mass of methane and find the number of molecules whose upward component velocities are (*a*) between 40,000 and 55,000 cm/sec, and (*b*) greater than 60,000 cm/sec.

8. Imagine the earth to be surrounded with an atmosphere of carbon dioxide (CO_2) at a temperature of 20° C. Determine the height at which the density would be one-third that at the surface.

9. In an experiment with Brownian particles it is found that the concentration of the particles which are suspended in a solution of density 1.10 gm/cm^3 falls to one-half in a height of 0.05 mm when the temperature is 20° C. The density of the particles is 1.15 gm/cm^3. The size of the particles is determined by evaporating the suspension of the particles as in the method of Perrin and 34 particles are counted in a row of length 0.02 mm. Assuming the particles to be spheres in contact when they are counted in the row, determine Avogadro's number.

10. Determine the respective heights at which the concentrations of the suspended Brownian particles in Problem 9 are reduced to 0.8, 0.6, 0.4 and 0.2 of the concentration at the lowest level.

CHAPTER IX

SPECIFIC HEATS AND HEAT RADIATION

9.1. Specific, Molecular and Atomic Heats.—The specific heat of a substance is the amount of heat measured in calories which is required to raise one gram of the substance through one degree centigrade; the molecular heat is the amount of heat required to raise one gram-molecule of the substance through one degree; and the atomic heat of an elementary substance is the amount of heat required to raise one gram-atom of the substance through one degree. Hence, the molecular heat of a substance is the product of the molecular weight and the specific heat of the substance, and the atomic heat is the product of the atomic weight and the specific heat of an elementary substance.

In the above paragraph the specific, molecular, and atomic heats are defined in terms of the calorie. However, since the calorie is equivalent to an energy of 4.187×10^7 ergs, the specific, molecular, and atomic heats may be defined in terms of the heat measured in ergs required to raise the respective masses through one degree centigrade. The relation between any one of these quantities c measured in calories and the same quantity C measured in ergs is

$$C = Jc \tag{9–1}$$

where J is the mechanical equivalent of heat having the value 4.187×10^7 erg/cal. In this chapter we shall assume that the various heats are measured in ergs per degree and not in calories per degree.

9.2. Specific Heats at Constant Volume and Constant Pressure.—The specific heat of a gas is an ambiguous term because the specific heat depends upon the conditions of volume and pressure under which it is measured. In order to be definite we must say under what conditions of volume and pressure the specific heat is measured. Usually we confine ourselves to two sets of conditions, (1) specific heat at constant pressure C_p, and (2) specific heat at constant volume C_v. When a gas is heated at constant pressure it expands and does work. This extra work is done at the expense of the heat energy supplied and so more heat must be supplied to the gas to raise its temperature through one degree. Hence the specific heat at constant pressure is greater than that at constant volume. We shall now show that the difference between the two specific heats is equal to the gas constant per gram R.

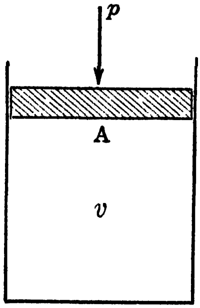


FIG. 9.1.

Consider a gas of volume v enclosed in a cylinder of area of cross section A. Suppose the piston to exert a pressure p as in Fig. 9.1. If the pressure is kept constant, then on heating the gas from a temperature of $T°$ Abs to $(T + dT)°$ Abs, the volume increases from v to $v + dv$. The increase of volume is obtained by differentiating

$$pv = RT \tag{9-2}$$

and holding p constant so that

$$pdv = RdT \tag{9-3}$$

The piston in Fig. 9.1 moves through a distance dv/A. But the total force on the piston is pA, so that the work done against this force is

$$pA \times \frac{dv}{A} = pdv \tag{9-4}$$

From (9–3) this work equals RdT. Now, it requires $C_v dT$ ergs to raise the temperature of the gas at constant volume through $dT°$, so that the work required to raise the gas through $dT°$ at constant pressure is $C_v dT + RdT$. But the work required to raise the gas at constant pressure through $dT°$ is $C_p dT$, so that

$$C_p dT = C_v dT + RdT$$

Dividing throughout by dT and transposing C_v, we have

$$C_p - C_v = R \tag{9-5}$$

This may be considered either as a relation between the specific or the molecular heats of a gas. If the C's are specific heats, R is the gas constant per gram; while, if the C's are molecular heats, R is the gas constant per gram-molecule. The value of R in this latter case is the same for any gas, namely, 8.315×10^7 erg/deg. Since 4.187×10^7 ergs = 1 calorie, the value of R for a gram-molecule may be expressed as $8.315 \times 10^7/4.187 \times 10^7 = 1.986$ cal/deg. Since 1.986 is nearly equal to 2, the value of R for a gram-molecule sometimes is remembered as 2 calories per degree.

9.3. Degrees of Freedom.—The number of degrees of freedom of a body is equal to the number of coordinates which must be given to describe its position completely. A point needs three coordinates to describe its position completely and so has three degrees of freedom. A rod needs the three coordinates of its center of gravity and two angles to describe its position completely and so has five degrees of freedom. A spheroid has five degrees of freedom. An ellipsoid which has all three of its axes unequal needs three coordinates for its center of gravity, two angles for its major axis and one angle for one of its smaller axes to describe its position completely, so that an ellipsoid has six degrees of freedom.

In Chapter VIII on the Kinetic Theory of Gases we have seen that the average translational kinetic energy per molecule of a gas at temperature T is $(3/2)kT$, where k is the gas constant for a single molecule and is known as Boltzmann's constant. The relation between k and R, the gas constant per gram-molecule is

$$k = \frac{R}{N} = 1.38 \times 10^{-16} \text{ erg/deg} \tag{9–6}$$

where $N = 6.03 \times 10^{23}$ is Avogadro's number. Since the velocity of a molecule has components in the directions of three mutually perpendicular axes of coordinates, the average kinetic energy of a molecule associated with the x-direction is one-third of $(3/2)kT$ or $(1/2)kT$. By methods which are beyond the scope of this book, it can be shown that the average energy associated with a degree of freeedom for a molecule of gas at temperature T is $(1/2)kT$. This is known as the Maxwell-Boltzmann principle of the equipartition of energy.

9.4. Monatomic Gases.—In the case of monatomic gases, the atoms are considered either as points or spheres, so that they have only three degrees of freedom. Consequently, the average kinetic energy of each atom (or molecule) is $3 \times (1/2)kT = (3/2)kT$. The total energy for all the atoms (or molecules) in a gram-atom, which is also a gram-molecule in this case, is therefore

$$W = N \times (3/2)kT = (3/2)RT \tag{9–7}$$

In order to raise the temperature of the gas from T to $T + dT$ without the gas doing external work, the energy given to the gas must be

$$dW = (3/2)RdT \tag{9–8}$$

Under the condition of no external work, the volume of the gas must be kept constant, so that $C_v dT = dW$. Hence from (9–8)

$$C_v = \frac{dW}{dT} = (3/2)R \tag{9–9}$$

The molecular (or atomic) heat at constant volume for a monatomic gas is therefore $(3/2)R$. From (9–5) the molecular heat at constant pressure is $(5/2)R$. Since R for a gram-molecule is approximately 2 calories, the molecular heats of a gas at constant volume and constant pressure are respectively 3 and 5 cal/gm-mol. The gases helium, neon, and argon are monatomic gases.

FIG. 9.2.—Representation of a diatomic molecule.

9.5. Diatomic and Polyatomic Gases.—The atoms in the molecules of a diatomic gas may be considered as being joined together to form a body as in Fig. 9.2. To specify completely the position of such a molecule, three coordinates are needed for the center of gravity and two angles for the direction of the axis, so that the molecule has five degrees of freedom. According to Boltzmann's rule, the average energy of such a molecule must be $5 \times (1/2)kT = (5/2)kT$. The energy of a gram-molecule of a diatomic gas at a temperature T is therefore

$$W = N \times (5/2)kT = (5/2)RT \tag{9–10}$$

which on differentiation gives

$$C_v = \frac{dW}{dT} = (5/2)R \tag{9–11}$$

The molecular heats of a gas at constant volume and constant pressure are therefore respectively $(5/2)R$ and $(7/2)R$, or approximately 5 and 7 cal/gm-mol. The gases hydrogen, nitrogen, and oxygen are diatomic.

Triatomic gas molecules may be thought of as shaped somewhat like Fig. 9.3. In order to specify completely the position and orientation of such a body, three coordinates are needed for the center of gravity, two angles for the direction of the axis perpendicular to the plane of the molecule, and one angle for the orientation of the molecule in this plane. A triatomic molecule thus has six degrees of freedom, so that using Boltzmann's rule the average energy of such a molecule is $6 \times (\frac{1}{2})kT = 3kT$. The energy of a gram-molecule at temperature T is

$$W = 3NkT = 3RT \quad (9\text{–}12)$$

The molecular heat at constant volume is

$$C_v = \frac{dW}{dT} = 3R \quad (9\text{–}13)$$

while the molecular heat at constant pressure is $4R$. In calories, the approximate values of the molecular heats at constant volume and constant pressure are 6 and 8 cal/gm-mol. The gases carbon dioxide and sulphur dioxide are triatomic gases.

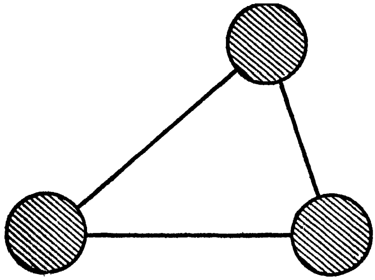


FIG. 9.3.—Representation of a triatomic molecule.

Gases such as NH_3 and CH_4 contain more than three atoms. However, except for vibration of the atoms within a molecule, such a molecule requires only three space and three angular coordinates to specify completely its position and orientation and therefore has six degrees of freedom. The molecular heats of polyatomic gases are therefore the same as those of triatomic gases. We shall discuss the effect of vibration of the atoms within a molecule in a later section.

Another way of considering degrees of freedom is as follows: A point having only position but no size can be given translational motion but not rotational motion. Further, a smooth sphere can be given translational but not rotational motion by collision with another smooth sphere. It is thought that when the molecules of a monatomic gas collide with each other their translational motions are mutually affected, but that they are unable to develop rotation in each other. When two bodies like that shown in Fig. 9.2 collide, not only are their translational motions mutually affected, but they set each other rotating. If the spheres are smooth, the body can be set rotating about either or both of two axes which are mutually perpendicular to each other and perpendicular to the line joining the centers of the spheres, but the body cannot be set rotating about the line joining the centers of the smooth spheres. Hence, bodies shaped like Fig. 9.2 can have translational motion which has components in three mutually perpendicular directions and rotational motion about two, but not three, perpendicular axes. In a gas consisting of diatomic molecules like Fig. 9.2 three-fifths of the energy is in translational kinetic energy and two-fifths is in rotational kinetic energy. When two bodies like that shown in Fig. 9.3 collide, they can set each other rotating about three mutually perpendicular axes. In a gas consisting of tri- or polyatomic molecules three-sixths or one-half of the energy is in translational kinetic energy and three-sixths or one-half in rotational kinetic energy.

Because of the rotational kinetic energy of diatomic and polyatomic molecules the phrase "rotational specific heat" has come into use. The rotational specific heat is the amount of energy required for the increase in the rotational kinetic energy of the molecules in a gram of gas when the temperature of the gas is raised through one degree.

9.6. Vibrational Energy.—If the atoms of the molecule depicted in Fig. 9.2 can vibrate back and forth along the line of centers, another degree of freedom is added because it is now necessary to add the distance apart of the two atoms in the molecule to the three coordinates of the center of gravity of the molecule and the two angles in order to specify completely the position, orientation and size of the molecule. Or, looking at the matter in another way, two such molecules can set each other vibrating when they collide. The two atoms in the molecule are thought to be in equilibrium when their centers are at a distance a_0 from each other. If the distance apart is a and not a_0, the atoms repel each other when $a < a_0$ but attract each other when $a > a_0$. The atoms if disturbed will vibrate along the line of centers. Now, if $a < a_0$, the potential energy of the system is greater than that for the equilibrium distance a_0; and, if $a > a_0$, the potential energy is again greater than that for a_0. A vibrating system has in general kinetic and potential energies. Because of this, the equipartition principle requires that the average energy associated with the vibrational kinetic energy of a diatomic molecule be $(\frac{1}{2})kT$ and the average potential energy be $(\frac{1}{2})kT$. Hence, although the vibration along the line of centers only adds one degree of freedom to the molecule, yet the average energy added is $2(\frac{1}{2})kT = kT$. Hence we may speak of the "vibrational specific heat" of a gas, just as we speak of the rotational specific heat.

If vibration of the atoms in a diatomic molecule is allowed, the total average energy of a molecule is $(3/2)kT$ on account of the translational motion plus $(2/2)kT$ on account of the rotational motion plus $(2/2)kT$

on account of the vibrational motion, the total being $(7/2)kT$. Hence, the energy of a gram-molecule at a temperature T is

$$W = (7/2)NkT = (7/2)RT \tag{9-14}$$

so that the molecular heat at constant volume is

$$C_v = \frac{dW}{dT} = (7/2)R \tag{9-15}$$

The molecular heats at constant volume and pressure for a diatomic gas whose molecules can vibrate are therefore respectively $(7/2)R$ and $(9/2)R$, or approximately 7 and 9 cal/gm-mol.

9.7. Ratio of the Specific Heats.—The ratio of the specific heat of a gas at constant pressure to that at constant volume occurs in the formula for the velocity of sound in a gas as we shall see in a later chapter. If $\gamma(= C_p/C_v)$ is the ratio of the specific heats, the velocity of sound in a gas is given by

$$v = \sqrt{\frac{\gamma p}{\rho}} \tag{9-16}$$

where p is the pressure and ρ the density of the gas. The values of γ for various kinds of gases are shown in Table 9.1.

TABLE 9.1

RATIO OF SPECIFIC HEATS

Kind of Gas	Cp/C_v
Monatomic	5/3 = 1.67
Diatomic without vibration	7/5 = 1.40
Diatomic with vibration	9/7 = 1.29
Triatomic without vibration	8/6 = 1.33

The values of γ as calculated from the measured velocity of sound in various gases agree very well with the values obtained from the theory described in Secs. 9.4 to 9.6.

9.8. Effect of Temperature.—The values of the specific heats of gases which we have discussed in the previous sections are for ordinary temperatures. However, the specific heat of a gas at constant volume is found to vary with the temperature. An interesting example is the gas hydrogen. At temperatures between 0° C and 300° C, the molecular heat C_v is practically constant at the value 4.95 in good agreement with $(5/2)R$ or 4.97 cal/gm-mol. Below 0° C (273° Abs), C_v for hydrogen falls until, between 75° and 30° Abs, it reaches another constant value 2.98, which is equal to $(3/2)R$. It seems that at low temperatures two degrees of freedom are suppressed. These are the degrees of freedom associated with the two-angle coordinates previously mentioned. So long as these two degrees of freedom are operative the molecule can have rotational as well as translational motion. The molecular heat which is due to these two degrees of freedom is known as the rotational molecular heat. At low temperatures of hydrogen there is only translational molecular (or specific) heat; while at ordinary temperatures the molecular heat is the sum of the translational and rotational molecular heats. At temperatures above 300° C (573° Abs), the molecular heat of hydrogen increases and seems to approach a value of about $(7/2)R$ at a temperature somewhere above 2000° C. It seems then that at very high temperatures all the molecules are set into vibration and the vibrational molecular heat becomes important.

As the temperature of a gas rises, the speeds of the molecules increase. It seems that the molecules of hydrogen cannot be set rotating unless the energy of one of the molecules which take part in a collision is above a certain amount. Also it seems that the atoms of the molecules cannot be set vibrating unless the energies are above a certain greater amount.

TABLE 9.2

Exceptions to the Dulong and Petit Law

Element	Temperature ° C	Atomic heat
H	−260.6	0.57
Be	+ 50	3.85
B	50	3.35
C	50	1.45
Si	50	4.95

9.9. Atomic Heats of Solids.—In 1819 Dulong and Petit announced that the atomic heat is the same for all elementary solids. The mean atomic heat of some thirty elementary solids is about 6.3 calories per gram-atom per degree. However, the atomic heats of beryllium, boron, and carbon were early found to depart considerably from the value 6.3, Later, the atomic heats of some other elements were also found to differ considerably from the Dulong and Petit value. These exceptions are shown in Table 9.2. The temperatures given in the table are the mean temperatures at which the atomic heats were obtained.

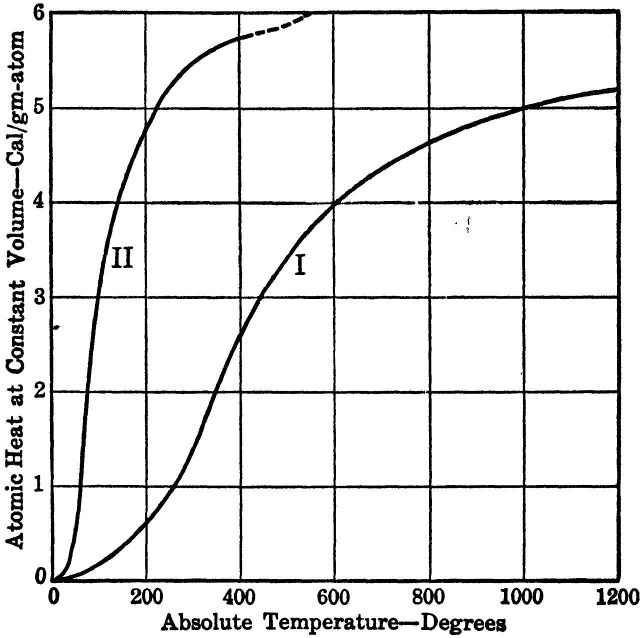


FIG. 9.4.—Variation of atomic heat with temperature—Curve I, diamond; Curve II, aluminum.

We have seen in Sec. 9.8 that the specific heat of hydrogen gas varies with the temperature. A similar effect is true for solids. In 1872 Weber observed that the specific heat of diamond (carbon) increases rapidly with rise of temperature above 0° C. In recent years specific heats have been studied at very low temperatures—even as low as that of solid hydrogen as shown in Table 9.2. It has been found that the atomic heats of all elements decrease below the Dulong and Petit value of 6.3 at low temperatures and that the atomic heats all approach 6.3 at high temperatures.

Curves I and II, Fig. 9.4, respectively show the variation of the atomic heats of diamond and aluminum with the temperature. It is seen that the two curves are similar in shape. It is a very remarkable fact that if the abscissae of the curve for diamond are divided by 4.6 and a new curve plotted, the curve so obtained very nearly coincides with the curve for aluminum. Similarly it has been found that the curve for any other element can be obtained by dividing the abscissae of the diamond curve by an appropriate number.

9.10. Classical Theory.—According to the kinetic theory a solid differs from a gas in that its atoms are almost held fixed in position relative to the solid as a whole, although the atoms can vibrate about their positions of equilibrium. Since an atom of a solid can vibrate in three mutually perpendicular directions, it has three degrees of freedom. But, as we have seen in Sec. 9.6, vibrational energy consists of both potential and kinetic energy so that the vibrational energy associated with the x-direction of vibration is kT. Since there are three mutually perpendicular directions, the average total vibrational energy of the atom is $3kT$. The total energy in a gram-atom at temperature T is therefore

$$W = 3NkT = 3RT \tag{9-17}$$

whence the atomic heat is

$$C_v = \frac{dW}{dT} = 3R \tag{9-18}$$

or 5.97 cal/gm-mol which is close to the experimental value of 6.3. It is to be noted that the experimental value is the atomic heat at constant pressure and not at constant volume. A solid is held together by internal attractions between its atoms or molecules and when it expands work is done against the attractive forces. It can be shown for a solid that

$$C_p - C_v = \beta^2 MT/K\rho \tag{9-19}$$

where β is the coefficient of volume expansion, M the molecular weight, K the compressibility and ρ the density of the solid. When C_v is calculated from the experimental value of C_p by means of (9-19), it is found that the experimental value of C_v is reasonably close to the theoretical value $3R$ for many solid elements.

However, although the above theory gives a result in agreement with the law of Dulong and Petit, the theory does not allow any variation of atomic heat with temperature of the kind shown in Fig. 9.4.

✓**9.11. Energy of Radiation.**—Consider the radiation in the space enclosed within a hollow steel ball which is maintained at a constant white heat. The radiation within such a space is called "black body" radiation because the radiation coming from each unit area of the walls of the enclosure is the same as that which comes from unit area of a perfectly black body maintained at the same temperature. The radiation from each unit area of the walls of the enclosure is determined only by the temperature of the walls and not by the substance of which the walls are composed. In such an enclosure the intensity of the radiation falling on a unit area of the walls equals the radiation coming from the unit area. The radiation within such an enclosure can be studied experimentally by making a very small hole in the wall. The radiation which otherwise would have fallen on the wall at the position where the hole now is will then pass through the hole. If the ratio of the area of the hole to the total area of the walls of the enclosure is small, the loss of radiation through the hole is small compared with the total radiation within the enclosure, so that the radiation in the enclosure remains practically unaltered. Although the radiation passing through the hole is only a small fraction of the radiation within the enclosure, yet it is a fair sample of the radiation within the enclosure and is therefore very nearly black body radiation. This radiation can be examined for the total energy of all frequencies and for the energy in the range of frequencies ν to $\nu + d\nu$ being emitted by unit area of the hole in unit time. The rate at which energy of all frequencies is emitted by unit area of the hole is given by Stefan's Law

$$R = \sigma T^4 \tag{9-20}$$

where σ is a constant and T the absolute temperature of the enclosure. The value of σ is 5.71×10^{-5} erg cm^{-2} deg^{-4} sec^{-1}.

We must now distinguish between the rate of emission of energy by unit area of the hole and the energy of the radiation contained in unit volume of the uniform temperature enclosure. The energy of the radiation within the enclosure is moving about with the speed of light c. At any instant of time there is a definite amount of energy in each unit volume of the enclosure. This energy per unit volume is known as the energy density of the radiation. The value of this energy density is

$$\rho = aT^4 \tag{9-21}$$

where $a = 4\,\sigma/c = 7.6 \times 10^{-15}$ erg cm^{-3} deg^{-4}.

We have seen in Sec. 3.8 that the number of vibrational modes in a unit cube when standing waves are set up is

$$dn = 8\pi\nu^2 d\nu/c^3 \tag{9-22}$$

for transverse waves of frequencies between ν and $\nu + d\nu$ which are traveling with velocity c. Assuming then that radiation in a hollow steel ball consists of standing waves and that each one of these harmonics or vibrational modes represents a degree of freedom and also that the average energy per vibrational degree of freedom is kT, we obtain

$$dW = 8\pi\nu^2 kT d\nu / c^3 \tag{9-23}$$

for the energy in the radiation between frequencies ν and $\nu + d\nu$. Now to obtain all the energy present at temperature T we must integrate (9–23) from $\nu = 0$ to $\nu = \infty$ since we suppose all frequencies to be present. The result is an infinite amount of energy which is completely at variance with the finite energy density given in (9–21).

Because of this complete break-down of the so-called "classical" theory, Planck in 1900 was led to another approach to the problem. It is to be remembered that radiation is really both absorbed and emitted by the walls of the enclosure, and by any gas which might be in the enclosure. Planck suggested the idea that when radiation is emitted or absorbed it is emitted or absorbed in multiples of a definite amount. This definite amount is known as a quantum. The energy of a quantum of radiation of frequency ν is proportional to ν and so is equal to $h\nu$, where h is a proportionality constant known as Planck's constant. As a result of this idea the energy of each vibrational mode in the enclosure can assume only the values $0, h\nu, 2h\nu, 3h\nu, \cdots$. Boltzmann's concentration law as given in Sec. 8.14 is applied as follows: The probabilities of a vibrational mode of frequency ν having the energies 0, $h\nu$, $2h\nu$, etc., are respectively proportional to e^0, $e^{-h\nu/kT}$, $e^{-2h\nu/kT}$, and so on. From (2–65) the average energy per vibrational mode in the frequency range ν to $\nu + d\nu$ is therefore

$$\frac{0e^0 + h\nu e^{-h\nu/kT} + 2h\nu e^{-2h\nu/kT} + \cdots}{e^0 + e^{-h\nu/kT} + e^{-2h\nu/kT} + \cdots}$$

$$= h\nu e^{-x} \frac{1 + 2e^{-x} + 3e^{-2x} + \cdots}{1 + e^{-x} + e^{-2x} + \cdots} \tag{9-24}$$

where $x = h\nu/kT$. The denominator is a geometric series whose sum is $1/(1 - e^{-x})$. In order to evaluate the numerator we remember that from the binomial theorem

$$1/(1 - y)^2 = (1 - y)^{-2} = 1 + 2y + 3y^2 + \cdots$$

Hence the value of the numerator is $1/(1 - e^{-x})^2$ and the average energy per vibrational mode of frequency ν is

$$w = h\nu/(e^{h\nu/kT} - 1) \tag{9–25}$$

But the number of modes in the frequency range $d\nu$ is given by (9–22), so that the energy of the modes in this range is

$$d\rho = \frac{8\pi\nu^2}{c^3} \cdot \frac{h\nu}{(e^{h\nu/kT} - 1)} \cdot d\nu \tag{9–26}$$

This can be written in the form

$$d\rho = \frac{8\pi k^4 T^4}{c^3 h^3} \cdot \frac{x^3 dx}{e^x - 1} \tag{9–27}$$

where again $x = h\nu/kT$. In order to integrate this, we make use of the definite integral

$$\int_0^\infty \frac{x^3 dx}{e^x - 1} = \pi^4/15 \tag{9–28}$$

which we mentioned in Sec. 2.4. Hence the energy density over the whole range of frequencies from 0 to ∞ is

$$\rho = \frac{8\pi^5 k^4}{15 h^3 c^3} \cdot T^4 \tag{9–29}$$

So from (9–21), $a = 8\pi^5 k^4/15h^3c^3$. Using the value of a given above, we can solve for Planck's constant h. The value so obtained is $h = 6.55 \times 10^{-27}$ erg sec.

9.12. Planck's Distribution Law.—From (9–26), we see that the energy density in the frequency range ν to $\nu + d\nu$ is given by $I(\nu)d\nu$, where the distribution function $I(\nu)$ is given by the coefficient of $d\nu$ on the right side of (9–26). It is usual to measure the distribution with respect to wave-length rather than frequency. We therefore make use of

$$\nu = c/\lambda \tag{9–30}$$

whence

$$d\nu = -c\,d\lambda/\lambda^2 \tag{9–31}$$

and $I(\nu)d\nu$ becomes $H(\lambda)d\lambda$, where

$$H(\lambda) = \frac{8\pi hc}{\lambda^5 \{\exp(hc/\lambda kT) - 1\}} \tag{9–32}$$

and $H(\lambda)$ is the distribution function with respect to wave-length. The graph of $H(\lambda)$ plotted against λ is shown in Fig. 9.5. This curve agrees very well with the experimental curve. If the average energy per vibrational mode is kT instead of the value given by (9–25), we obtain from (9–23) a distribution function $I(\nu) = 8\pi kT\nu^2/c^3$ with respect to frequency or, from (9–30) and (9–31), $H(\lambda) = 8\pi kT/\lambda^4$. This is the Rayleigh-Jeans distribution function and is shown as the broken curve in Fig. 9.5. At very long wave-lengths the Rayleigh-Jeans and the Planck curves become indistinguishable. This is because the average energy of a vibrational mode as given by (9–25) becomes kT for low frequencies. The full curve in Fig. 9.5 is sometimes called the spectrum of the continuous radiation given out by a hot black body.

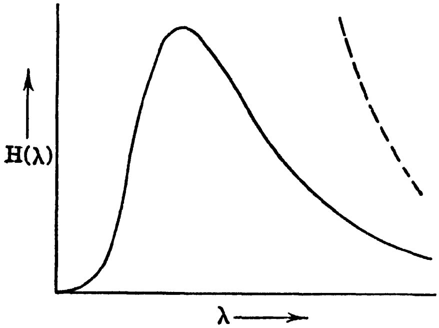


FIG. 9.5.—Radiation from a black body. Full curve—Planck distribution; broken curve—Rayleigh-Jeans distribution.

9.13. Wien's Displacement Law.—From Fig. 9.5 we see that the full curve for a given temperature T has a maximum at a certain value λ_m of the wave-length. To find λ_m we differentiate the right side of (9–32) and equate the result to zero. This equation is somewhat difficult to solve for λ but the solution is $\lambda = \lambda_m$, where

$$\lambda_m T = 0.2014hc/k \tag{9–33}$$

Experimentally it has been found that the wave-length of maximum intensity in the radiation coming from a hot body is inversely proportional to the absolute temperature of the black body emitting the radiation, so that

$$\lambda_m T = 0.288 \tag{9–34}$$

where 0.288 is the experimental value of the constant when λ_m is measured in centimeters. This is known as Wien's displacement law and agrees with the theoretical formula (9–33) because the right side of (9–33) is a constant. Equating the right sides of (9–33) and (9–34), we obtain

$$h = 1.43k/c \tag{9–35}$$

whence, on putting in values for k and c, we obtain $h = 6.62 \times 10^{-27}$ erg sec, the value obtained in Sec. 9.11.

9.14. Elastic Waves in a Solid.—In the case of elastic waves in a solid, there are both longitudinal and transverse waves. For both kinds of waves, the velocity is given by

$$\text{velocity} = \sqrt{\frac{\text{elasticity}}{\text{density}}} \tag{9–36}$$

The elasticity used in (9–36) is different for longitudinal waves from that used for transverse waves.

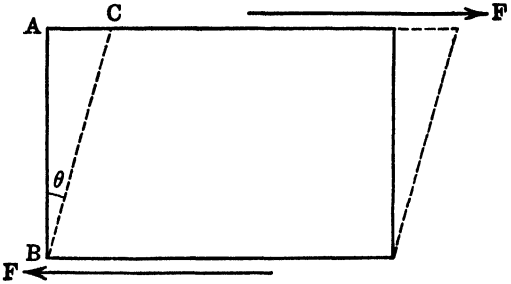


FIG. 9.6.—Block subjected to shearing stress.

Consider a block of material as shown in Fig. 9.6. Let a force F be applied to the upper surface and parallel to this surface. The base of the block is fixed so that the force F distorts the block into the shape shown by the broken lines. This kind of distortion is called a shear, and the angle θ, Fig. 9.6, is known as the angle of shear. If S is the area of the upper face of the block, the stress producing the strain θ is F/S, so that by Hooke's law

$$\frac{F}{S} = n\theta \tag{9–37}$$

where n is a constant known as the rigidity. (9–37) may be written

$$\text{rigidity} = \frac{F}{S\theta} \tag{9–38}$$

where the quantities on the right side are defined with reference to Fig. 9.6.

Consider a vertical wire hung from a support. If weights are attached to the lower end of the wire, the wire stretches. If a weight mg causes a stretch s in a wire of length L and cross section A, Young's modulus Y is given by

$$Y = \frac{\left(\frac{F}{A}\right)}{\left(\frac{s}{L}\right)} = \frac{FL}{sA} \tag{9–39}$$

Closely related to Young's modulus is another elastic constant known as Poisson's ratio. When a wire is stretched by a pull, the stretch is accompanied by a sideways contraction. If the diameter of the wire is D before, and $D - d$ after the application of the pull, and, further, if s is the stretch produced in the wire which is of length L by the pull, Poisson's ratio σ is given by

$$\sigma = \frac{\left(\frac{d}{D}\right)}{\left(\frac{s}{L}\right)} = \frac{dL}{sD} \tag{9–40}$$

It is shown in the theory of elasticity that the velocities of the longitudinal and transverse elastic waves in a solid are respectively

$$V_L = \sqrt{\frac{Y}{\rho} \cdot \frac{1 - \sigma}{(1 + \sigma)(1 - 2\sigma)}} \tag{9–41}$$

and

$$V_T = \sqrt{\frac{Y}{\rho} \cdot \frac{1}{2(1 + \sigma)}} \tag{9–42}$$

where ρ is the density of the solid.

9.15. Quantum Theory of Specific Heats.—The most satisfactory theory of the specific heats of solids is due to Debye. He considered the number of modes of vibration of a one centimeter cube. In this case there are longitudinal as well as transverse waves. As described in Sec. 9.11, Planck dealt with light and heat waves. However, when we are speaking of waves in a solid substance like copper, we are speaking of elastic waves or sound waves of high frequency. In Debye's theory we use quantum ideas even for these waves. From Sec. 3.8 the number of vibrational modes (or degrees of freedom) for the transverse waves in unit volume whose frequencies are in the range ν to $\nu + d\nu$ is

$$dn_T = \frac{8\pi\nu^2 d\nu}{V_T^3} \tag{9–43}$$

where V_T is the velocity of the transverse waves, while the number of modes of vibration for the longitudinal waves in the same frequency range is

$$dn_L = \frac{4\pi\nu^2 d\nu}{V_L^3} \tag{9–44}$$

where V_L is the velocity of the longitudinal waves. The total number of modes in the frequency range ν to $\nu + d\nu$ is therefore

$$dn = dn_T + dn_L = 4\pi\left(\frac{1}{V_L^3} + \frac{2}{V_T^3}\right)\nu^2 d\nu \tag{9–45}$$

The average energy of each mode of vibration is w as given by (9–25). Hence the energy in the frequency range ν to $\nu + d\nu$ is $w dn$. If we take the volume of the solid to be the volume v of a gram-atom of the element instead of 1 cm³, the number of modes is multiplied by v. Hence for a gram-atom the energy in the frequency range ν to $\nu + d\nu$ is $vw dn$ or

$$dW = 4\pi v\left(\frac{1}{V_L^3} + \frac{2}{V_T^3}\right)\frac{h\nu^3 d\nu}{e^{h\nu/kT} - 1} \tag{9–46}$$

Now whereas in Sec. 9.11 we integrated an expression similar to this from $\nu = 0$ to $\nu = \infty$, we must now remember that we are dealing with a solid made up of atoms instead of the space inside a uniform temperature enclosure. We can imagine light and heat waves to have exceedingly small wave-lengths and therefore exceedingly high frequencies but for elastic waves in a solid we cannot imagine a wave-length shorter than twice the distance between neighboring atoms, as is represented in Fig. 9.7. Since there is a minimum wave-length for longitudinal and also for transverse waves, there is a maximum frequency for each kind of waves. Debye made the assumption that the maximum frequency has the same value ν_m for each kind of waves. Furthermore, Debye assumed each vibrational mode to correspond to a degree of freedom and since the total number of degrees of freedom for N atoms cannot be greater than $3N$ it is reasonable to expect that the total number of vibrational modes of frequencies ranging from $\nu = 0$ to ν_m must be $3N$. Hence the integral of (9–45) between the limits 0 and ν_m must be $3N$, or

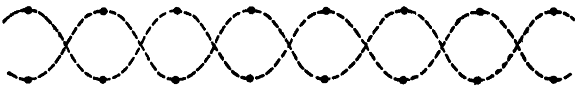


FIG. 9.7.—Shortest possible wave-length in a line of atoms.

$$4\pi\left(\frac{1}{V_L^3}+\frac{2}{V_T^3}\right)v\int_0^{\nu_m}\nu^2 d\nu = 3N \tag{9–47}$$

Carrying out the integration on the left side and rearranging, we obtain

$$\nu_m^3 = \frac{9N}{4\pi\left(\dfrac{1}{V_L^3}+\dfrac{2}{V_T^3}\right)}v \tag{9–48}$$

where N is the number of atoms in a gram-atom, or, in other words, Avogadro's number. To obtain the total energy at a temperature T in the total frequency range 0 to ν_m we integrate (9–46). We now introduce the characteristic temperature T_c, defined by

$$kT_c = h\nu_m \tag{9–49}$$

so that

$$T_c^3 = \frac{9h^3N}{4\pi k^3 v\left(\dfrac{1}{V_L^3}+\dfrac{2}{V_T^3}\right)} \tag{9–50}$$

To obtain the total energy at a temperature T in the total frequency range 0 to ν_m we integrate (9–46) between these limits. For brevity we put $x = h\nu/kT$ so that $d\nu = (kT/h)dx$. Then, in virtue of (9–49)

$$W = \frac{9NkT^4}{T_c^3}\int_0^{x_m}\frac{x^3}{e^x - 1}dx \tag{9–51}$$

where $x_m = h\nu_m/kT = T_c/T$. For small values of T_c/T or at high temperatures, the integral becomes

$$\int_0^{Tc/T} x^2 dx = \left[\frac{x^3}{3}\right]_0^{Tc/T} = \frac{T_c^3}{3T^3} \tag{9–52}$$

since $e^x - 1 = 1 + x - 1 = x$ for small values of x. Substituting in (9–51), we obtain

$$W = 3NkT = 3RT$$

and the atomic heat C_v is $dW/dT = 3R$, the Dulong and Petit value. For large values of T_c/T or at low temperatures, the value of the integral is a complicated function of x_m or T_c/T, so that W is a function of T_c/T. But T_c is a constant for a given element, so that the atomic heat, dW/dT, works out as another function of T_c/T, thus

$$C_v = \frac{dW}{dT} = 3RF\left(\frac{T_c}{T}\right) \tag{9–53}$$

The function $F(T_c/T)$ cannot be expressed in any simple algebraic form. Debye, however, has evaluated the function for the values of T_c/T shown in Table 9.3.

TABLE 9.3

T_c/T	$F(T_c/T)$	T_c/T	$F(T_c/T)$	T_c/T	$F(T_c/T)$
0.0	1.000	1.4	0.909	4.0	0.503
0.2	.998	1.8	.854	5.0	.369
0.4	.992	2.2	.797	6.0	.260
0.6	.983	2.6	.730	8.0	.138
0.8	.969	3.0	.663	10.0	.076
1.0	.952	3.5	.581	12.0	.045

The function $F(T_c/T)$ approaches unity at high temperatures. For $T_c/T = 1$, $F(T_c/T) = .952$ so that we may define the characteristic temperature as being that temperature at which the atomic heat equals .952 times that of the Dulong and Petit value. It is very interesting to note that the formula for T_c contains the velocities of longitudinal and transverse waves. These waves are sound waves in the solid. Heat vibrations in a solid are therefore nothing less than the vibrations of sound waves. The velocities V_T and V_L are given in terms of Young's modulus, Poisson's ratio and the density according to (9–41) and (9–42). The atomic (or specific) heat of a solid substance is thus a function of the elastic constants of the substance. The characteristic temperatures of some substances are shown in Table 9.4.

TABLE 9.4

Element	Characteristic Temperature
	deg. abs.
Al	413
Cu	340
Ag	220
Au	172
Pb	75

The excellent agreement between Debye's theory and experiment is shown for the case of aluminum in Table 9.5.

TABLE 9.5

ATOMIC HEAT OF ALUMINUM

Atomic Heat		
Temperature	Theory	Experiment
deg. abs.		
19.1	0.058	0.066
27.2	0.109	0.110
41.9	0.613	0.597
79.1	2.362	2.345
137.0	3.91	3.98
235.0	5.17	5.17
555.0	5.81	5.98

CHAPTER IX

PROBLEMS

1. Argon is a monatomic gas whose atomic weight is 40. Determine the specific heat of argon at ordinary temperatures (*a*) at constant volume and (*b*) at constant pressure.

2. Calculate the specific heats at ordinary temperatures of oxygen gas (O_2) at constant volume and at constant pressure.

3. Calculate the specific heats of carbon dioxide gas (CO_2) at constant pressure and at constant volume. Determine the velocity of sound in carbon dioxide at a temperature of 150° C.

4. Young's modulus and Poisson's ratio for copper are 12.3×10^{11} dynes/cm^2 and 0.337 respectively. The density of copper is 8.90 gm/cm^3. Find the respective velocities of longitudinal and transverse elastic waves in copper. Determine the characteristic temperature of copper. Find the specific heat of copper at each of the following temperatures: −250° C, −100° C, 0° C, 300° C. The atomic weight of copper is 63.6.

CHAPTER X

ELECTRONS IN METALS

10.1. Theory before 1926.—The difference between the electrical conductivity of conductors like metals and of insulators like sulphur is that some of the electrons in a metal are free while all or almost all of the electrons in an insulator are bound to atoms. If a difference of electrical potential is applied across a slab of an insulating material, the electrons bound to each atom may be displaced somewhat but nevertheless they still remain as part of the atom. According to modern theories this displacement explains the properties of dielectrics. However, if a difference of potential is applied to the two ends of a metal wire, a current immediately flows, and the free electrons immediately move under the action of the electric field. When no difference of potential is applied these free electrons are believed to be moving hither and thither, sometimes colliding with atoms, and sometimes traveling in the spaces between atoms. In Chapter IX we have described the atoms of a solid as being in a state of vibration. In the case of a gas, the average kinetic energy of the molecules according to (8–17) is $(3/2)kT$. In the case of an atom of a solid at a high temperature the average kinetic energy of vibration is likewise $(3/2)kT$ where k is Boltzmann's constant and T is the absolute temperature of the metal. The total energy for the atom is double this, for the vibrating atom in general has both kinetic and potential energy, the average potential energy being also $(3/2)kT$. If free electrons are intermingled with the vibrating atoms of a solid, the electrons will acquire an average kinetic energy of $(3/2)kT$. The electrons are traveling about inside the metal with motions which are similar to the motions of a gas, excepting that the velocities of the electrons are much higher than those of gas molecules at the same temperature. Since the mass of an electron is 1/1840 of that of a hydrogen atom we may say that the molecular weight of an electron gas is $1.0080/1840 = .00055$. Putting this in (8–40), we obtain $\alpha = 9.1 \times 10^6$ cm/sec at 0° C. Since the velocity of escape from the earth is 1.1×10^6 cm/sec, it is seen that the earth would immediately lose an electron atmosphere if the only forces holding the atmosphere down were gravitational forces.

At the surface of a metal there is a sudden drop of potential as the point of observation is moved from inside the metal through the surface to the outside. The potential inside is positive with respect to a point outside. This surface potential amounts usually to about 4 or 5 volts. It is this potential barrier or wall which prevents the electrons with their high speeds from escaping from the metal. If an electron has to overcome a potential V' esu in order to escape, its velocity must be greater than that given by

$$(\tfrac{1}{2})mv^2 = V'e' \tag{10-1}$$

Since the potential is usually measured in volts, it will be convenient to derive a formula connecting the velocity of the electron with that difference of potential measured in volts which will just stop the electron. In Sec. 5.6, we learned that

$$300 \text{ volts} = 1 \text{ esu of potential} \tag{10-2}$$

Hence, if a given difference of potential is measured in esu (V') and in practical units (V_P volts),

$$V' = \frac{V_P}{300} \tag{10-3}$$

Using the value of e'/m given in Chapter VII, we obtain

$$v = 5.94 \times 10^7 \times \sqrt{V_P} \tag{10-4}$$

Because the velocity of an electron can be calculated from the potential in volts just necessary to stop it, the velocity of an electron is very often spoken of as so many volts. For instance, when we say that an electron has a velocity of 4 volts, we do not mean what we say, but we mean that it has that velocity in cm/sec which is given by (10–4) when $V_P = 4$. Putting $V_P = 4$ in (10–4), we obtain a velocity of 1.19×10^8 cm/sec. If the surface drop of potential is 4 volts, the velocity of escape is 1.19×10^8 cm/sec or 750 mi/sec.

Consider a metal with a flat surface. The number of electrons per unit volume with velocity components perpendicular to the surface in the range u to $u + du$ is, according to (8–36),

$$nf(u)du = n(1/\alpha\sqrt{\pi}) \exp(-u^2/\alpha^2)du \tag{10-5}$$

where n is the total number of free electrons of all velocities per unit volume. The electrons moving up to unit area of the surface in unit time with component velocity u come from a volume of unit cross-section and depth u. The number of electrons in this volume with component velocities u in the range du is $nuf(u)du$. Hence the total number of electrons with component velocities greater than the velocity of escape u_3 is, following the method of Sec. 8.13,

$$n(\alpha/\sqrt{\pi})\int_{x_3}^{\infty} x \exp(-x^2)dx \tag{10–6}$$

where $x = u/\alpha$ and $x_3 = u_3/\alpha$. In this case, however, the integral can be evaluated in terms of known functions. Integrating (10–6), we obtain

$$(n\alpha/\sqrt{\pi})\left[-\frac{1}{2}\exp(-x^2)\right]_{x_3}^{\infty} = \frac{n\alpha}{2\sqrt{\pi}}\exp(-x_3^2) \tag{10–7}$$

Now, from (8–40), $\alpha = \sqrt{2RT/M}$, where M is the molecular weight of an electron gas. The relation between M and the mass m of the electron is $M = Nm$, so that

$$\alpha = \sqrt{2kT/m} \tag{10–8}$$

where k is Boltzmann's constant. According to most recent theory not all of the electrons given by (10–7) escape from the metal, but a definite fraction of this number do escape. Hence, in virtue of (10–8), the number of electrons escaping per second from unit surface of the metal is proportional to $n\sqrt{\frac{kT}{2m\pi}}\exp\left(-\frac{mu_3^2}{2kT}\right)$. Since each electron carries a charge, there is a current of negative electricity emitted by unit area of a metal surface maintained at temperature T given by

$$i = AT^{1/2}\exp(-b/T) \tag{10–9}$$

where

$$b = mu_3^2/2k \tag{10–10}$$

and A is a constant. Richardson has found that this formula fits the experimental results very well. In the derivation of (10–9) it is assumed that n, the number of free electrons per unit volume of the metal, is independent of the temperature T.

10.2. Richardson's Researches.—Until 1902 the experimental evidence on the escape of electricity from hot bodies was uncertain. Some observations had been made on the escape of electricity from hot bodies in a vacuum and others on the escape of electricity from hot bodies in air. Nearly two hundred years ago it was known that air became conducting in the vicinity of incandescent conductors. In 1873 Guthrie found that a red hot ball retained a negative but not a positive charge. At higher temperatures it was found that charges of both signs leaked rapidly. However, in 1902 Richardson began a series of investigations on the escape of electricity from hot bodies in a vacuum. In experiments of the type carried on by Richardson, the procedure is to heat a metal filament *F* in a vacuum as shown in Fig. 10.1. This filament passes down the axis of a hollow metal cylinder *PP*. The filament and cylinder are enclosed in an evacuated glass tube *T*. The filament is heated by means of a current from the battery *H*. The resistance of the filament can be determined from readings of the voltmeter *V* and ammeter *A*. Since resistance is a function of temperature, the temperature of the filament can be determined from its resistance, the wire thus acting as its own thermometer. Richardson found that at temperatures above a bright red heat most wires lose a negative charge to the surrounding cold metal cylinder *P*, Fig. 10.1. Richardson described the charge as being carried across the space between *F* and *P*, Fig. 10.1, by negative ions or carriers, and to the ions in this case he gave the name thermions. To the collection of phenomena which are connected with thermions Richardson gave the name thermionics.

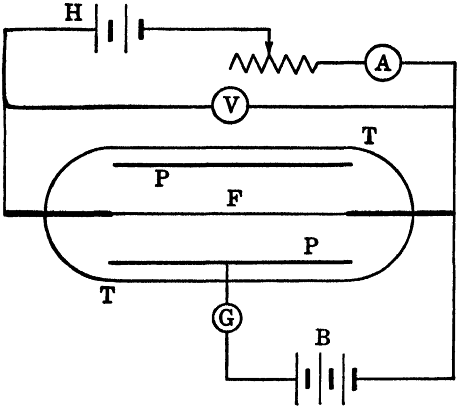


FIG. 10.1.—Apparatus for the study of the thermionic effect.

There are three variables in the experiment—the temperature *T* of the filament, the voltage V_B across the space between P and F, Fig. 10.1, and the current i_P registered on the galvanometer G. Richardson therefore plotted two curves: (1) i_P against V_B with T constant, (2) i_P against T with V_B constant. The curves for the two cases are shown in Figs. 10.2 and 10.3, respectively.

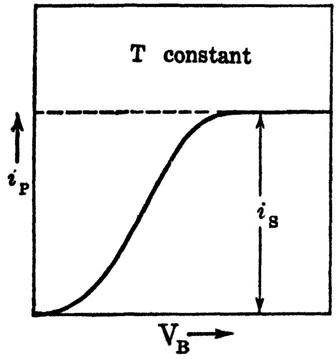


FIG. 10.2.

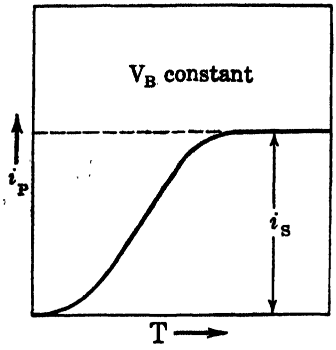


FIG. 10.3.

In Fig. 10.2 the current at first increases very slowly with the voltage, then more rapidly, and finally becomes constant. When the current has reached this constant value it is said to be saturated. The value of the saturation current is represented by i_S in Fig. 10.2. That value of the voltage beyond which an increase of voltage will not produce an increase of current is known as the saturation voltage.

In Fig. 10.3 the current at first increases very slowly, if at all, with the temperature. At higher temperatures the current rises rapidly, finally becoming constant at still higher temperatures. This constant value is known as the space charge limited current.

10.3. Nature of the Negative Thermions.—The nature of the thermions can best be determined by measuring e/m for the particles by means of an apparatus as shown in Fig. 10.4. The filament F is heated by current from the battery H. A difference of potential V_D is produced between the filament F and the metal diaphragm E by means of the battery D. The thermions in crossing this space attain a speed v given by

$$(\tfrac{1}{2})mv^2 = \frac{V_D e'}{300} \tag{10-11}$$

if V_D is considerably above the saturation potential and is measured in volts. Some of these thermions pass through the aperture in E into space G in which there is a uniform magnetic field H in a direction perpendicular to the plane of the diagram. The path of the thermions in this space is an arc of a circle of radius r. If the radius r is correct, the thermions pass through the aperture K into the metal collector L. As the charge on L increases, the electroscope M deflects. V_D and H are adjusted until the thermions enter L. Then

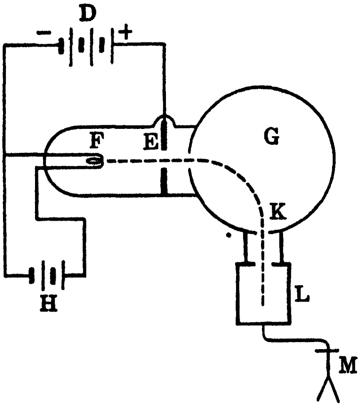


FIG. 10.4.—Apparatus for measuring e/m for thermions.

$$\frac{mv^2}{r} = \frac{He'v}{c} \tag{10-12}$$

as in Sec. 7.4. Eliminating v from (10–11) and (10–12), we obtain

$$\frac{e'}{m} = \frac{V_D c^2}{150H^2r^2} \tag{10-13}$$

where V_D is measured in volts and H in emu. When experimental values are substituted in the right side of (10–13), $e'/m = 5.28 \times 10^{17}$ esu/gm, which is the value for the electron. The negative thermions are therefore electrons.

10.4. Space Charge Effects.—Let us consider the peculiar shape of the curves in Figs. 10.2 and 10.3. As the thermions or electrons evaporate from the filament F in Fig. 10.1, an electron atmosphere forms between F and the cylinder P. This atmosphere is strongly charged negatively and tends to repel any electron from entering it. Hence, if this collection of negative electricity, or space charge as it is called, is very dense, an electron which is in the process of evaporating may be repelled back into the wire. If, however, the voltage between F and P in Fig 10.1 is increased, the rate at which electrons are removed from the space between F and P is increased, so that the density of the space charge becomes less and fewer electrons are repelled back into the filament. The electron current from F to P therefore increases with the voltage. However, if the voltage is great enough, then almost as soon as an electron has evaporated from the filament it is pulled over to the cylinder *P* and is not given enough time to form with other electrons a space charge between *F* and *P*. At this stage every electron which evaporates is pulled over to *P*, and, since the rate of evaporation is determined by the temperature of the filament and not by the voltage between *F* and *P*, the current reaches a maximum and is saturated. This explains the shape of the curve in Fig. 10.2.

In the case of Fig. 10.3 where the voltage is held constant the rate of evaporation at low temperatures is small, so that the voltage is able to remove the electrons as fast as they evaporate. However, as the temperature increases, the rate of evaporation also increases until the electrons cannot be removed fast enough by the field between *F* and *P*, Fig. 10.1. A space charge thus develops, and this repels some of the evaporating electrons back into the filament. With still further rise of temperature a situation develops in which no more than a certain number of electrons per second are allowed to leave the neighborhood of the filament, the remainder of the electrons being repelled back into the filament. We now have a constant current as in Fig. 10.3. The saturation current in this case is also known as the space charge limited current.

10.5. Effect of Temperature.—If the temperature of the filament is varied and at the same time a voltage greater than that needed for saturation is maintained between the filament *F* and the cylinder *P*, Fig. 10.1, it is found that the relation between the saturation current and the temperature is represented by a curve as shown in Fig. 10.5. This curve agrees very closely with that obtained when *i* as given by (10–9) is plotted against *T*. It therefore appears that the free electrons in a metal exist in the gaseous form, that an electron needs a certain velocity to escape from a metal, and that the free electrons in a metal have their velocities distributed according to the Maxwell law as we have described in Sec. 10.1.

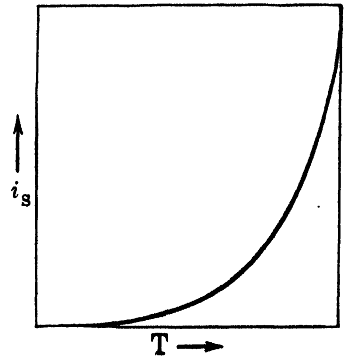


FIG. 10.5.—Graph of the saturation thermionic current against the temperature of the filament.

If our theory of thermionic emission is correct, it is the electrons with the most energy which leave the filament. The filament should then cool faster when the electrons are allowed to escape than when not. The electrons can be prevented from escaping by making the potential of P, Fig. 10.1, negative with respect to F. Richardson found that the rate of cooling is greater when the electrons are allowed to escape.

10.6. Thermionic Work Function.—The theory we have described in Sec. 10.1 is qualitative and approximate. In order to calculate the number of electrons escaping per second it is necessary to know the number of free electrons per unit volume in the metal and the mean free path of these electrons. Unfortunately, neither of these quantities is known. Richardson also found that the relation between saturation current and temperature is just as well expressed by

$$i = AT^2 \exp(-b/T) \tag{10-14}$$

as by (10–9). At present theoretical physicists favor (10–14) above (10–9). The form (10–14) can be obtained theoretically if it is assumed that n, the number of free electrons per unit volume, varies as $T^{3/2}$. In (10–6), u_3 $(=\alpha x_3)$ is the velocity of escape. The energy $(\frac{1}{2})mu_3^2$ is necessary to overcome the potential V' esu at the surface, so that

$$(\tfrac{1}{2})mu_3^2 = V'e' \tag{10-15}$$

Hence we may write $b = V'e'/k$; then (10–14) becomes

$$i = AT^2 \exp(-V'e'/kT) \tag{10-16}$$

V' may be expressed in volts and is usually of the order of 4 to 6 volts, its value depending on the particular metal used in the filament. The name "work function" is given to $V'e'$ because it represents the amount of work done by an electron in breaking through the surface of the metal. By 1946 it became the practice to express the work function in "electron-volts," although both terms, volts and electron-volts are used. In Chapter XI we shall have occasion to speak of another work function. The work function considered in this present chapter is known as the thermionic work function. The value of the work function for a given filament depends on the length of time that the filament has been heated. It seems that the superficial layers of the metal of a filament contain atoms or molecules of gas. A filament must be baked for hours at high temperature in the best vacuum obtainable before these gas atoms or molecules are driven out. This process is known as outgassing. When a filament has been outgassed it will give a constant value of the work-function. This value is the true value for the pure metal. Du Bridge after carefully outgassing a platinum filament obtained a thermionic work-function of 6.35 electron-volts. The electron-volt has been defined in Sec. 8.5. Its value is 1.6×10^{-12} erg.

10.7. Application to Vacuum Tubes.—The fact that a hot filament emits electrons is used in vacuum tubes. There are three general types of vacuum tubes—kenotrons, x-ray tubes, and radio tubes. As x-ray and radio tubes will be described in later chapters, we shall here confine ourselves to the description and use of kenotrons. A kenotron is also known as a diode since it contains two electrodes—the filament and the plate. A kenotron is a glass tube which has been evacuated to as high a degree as possible. The tube contains a filament *F* as shown in Fig. 10.6. The filament is heated by current from the battery *H*. Surrounding the filament and close to it is a metal cylinder *P*, which is called the plate. If a voltage *V* is applied across the kenotron, so that *F* is negative with respect to *P*, a current will flow through the ammeter *G*, but, if the voltage *V* is reversed, no current will flow through *G*. Hence a kenotron can be used as a valve to allow current to flow in one direction in a circuit but not in the opposite direction. A rectifier is a device for changing alternating current into direct current. If the voltage *V* in Fig. 10.6 is alternating, current will only flow through *G* during that half of each cycle when the voltage is in such a direction as to make *F* negative to *P*. The current through *G* will not be a steady direct current but a pulsating direct current.

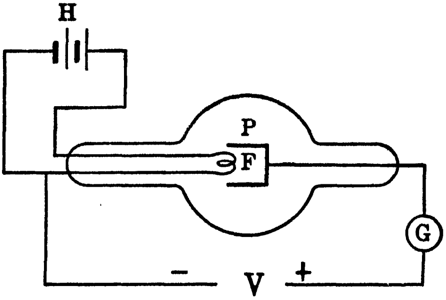


FIG. 10.6.—Diagram of a kenotron and connections.

With the aid of a condenser and a resistance or inductance it is possible by means of a kenotron to obtain a rectified current which is very nearly steady. We shall illustrate with a problem.

Problem 1. It is required to rectify an a.c. of 50,000 volts so as to give a current with a 2 per cent ripple through a resistance of 10 megohms. The frequency is 60 cycles.

The phrase "2 per cent ripple" means that the rectified current is not quite steady but varies from a maximum to a minimum, the difference between maximum and minimum being 2 per cent of the average current. When an alternating voltage is given as 50,000, this figure means the root mean square voltage, so that the maximum voltage is $50{,}000\sqrt{2}$ or 70,700 volts (see Chapter IV). A kenotron *K* and a condenser *BD* are connected as shown in Fig. 10.7. Excepting for the condenser *BD*, the current through the resistance *R* would be as shown by the curve in Fig. 10.8. There would be 60 pulses per second. Current will not flow through the kenotron *K* unless the plate *P* is more positive than *F*. When current does flow through *K* the condenser charges up so that the plate *B* becomes positive and *D* negative. The charge given to the condenser is determined by the maximum voltage and is $C \times 70{,}700$ coulombs where *C* is the capacitance of the condenser in

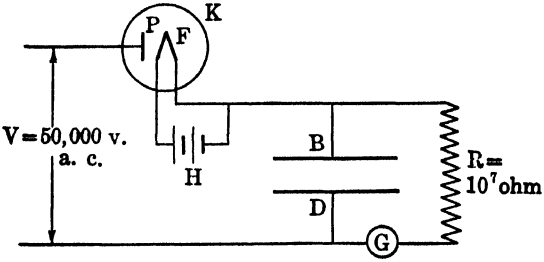


FIG. 10.7.—Rectification of an alternating current by means of a kenotron and a condenser.

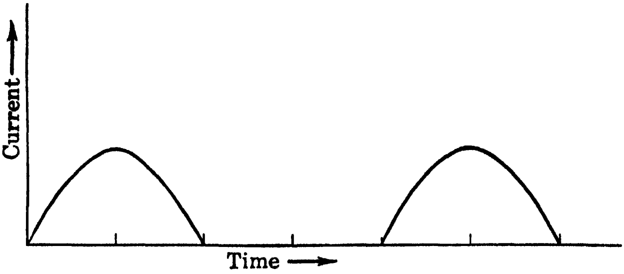


FIG. 10.8.—Graph of rectified current against time—no condenser in the circuit.

farads. The potential of P then falls relatively to F and becomes negative. During this time no charge is given the condenser, but rather the condenser is discharging through the resistance R. As the cycle proceeds the potential of P becomes positive and finally becomes more positive than that of F when current flows through K into the condenser and charges it again. The process is then repeated. When the condenser has been charged to 70,700 volts, the current through R is $70{,}700/(10 \times 10^6) = .00707$ amp. This is the maximum current, so that the minimum current is .00693 amp and the average current is .00700 amp since the ripple is 2 per cent. Since the ripple is small the voltage V rises almost to its maximum positive value before current flows through the kenotron. If, in order to simplify the problem, we suppose that current only flows through the kenotron into the condenser when the positive voltage is a maximum, the condenser discharges for 1/60 sec and at the end of each 1/60 sec it receives a sudden charge through the kenotron. Since the average current through the resistance R is .007 amp and this flows for 1/60 sec the quantity of charge is $.007 \times 1/60 = .000117$ coulomb. During this 1/60 sec the charge on the condenser has fallen 2 per cent, so that

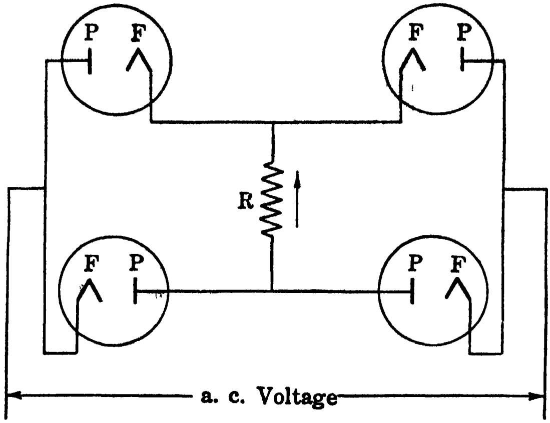


FIG. 10.9.—Hook-up of four kenotrons for full-wave rectification. The electron current flows upward through R.

$$.02 \times 70{,}700 \times C = .000117 \text{ coulomb}$$

whence $C = 8.2 \times 10^{-8}$ farad or .082 microfarad. A condenser of this capacitance will give the required ripple.

In a circuit as in Fig. 10.6, current is being supplied for only half the time if V is an a.c. voltage. During a half-cycle, current is flowing but during the next half-cycle no current is flowing, so that the apparatus is lying idle for half the time. This is not always desirable and a hook-up of four kenotrons has been devised to give full-wave rectification, as it is called. The hook-up is shown diagrammatically in Fig. 10.9. As is customary, the batteries for supplying the heating currents to the filaments F are not shown. The electron current will always flow upward through R. A graph of the current through R is shown in Fig. 10.10. It is seen that current flows during each half-cycle.

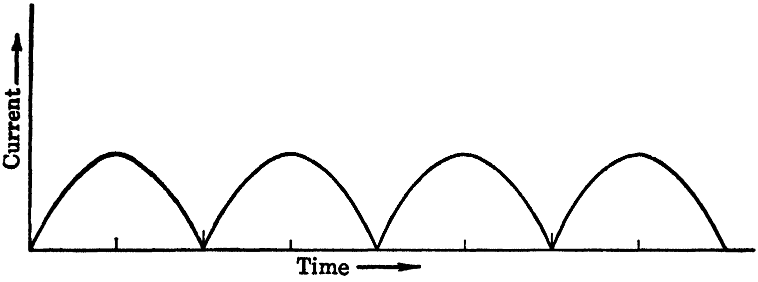


FIG. 10.10—Graph of the current against time, showing full-wave rectification— no condenser in the circuit.

10.8. Electron Theory of Electrical Resistance.—In Sec. 10.1 we have assumed that there are "free" electrons in metals which are moving about in the metal with random velocities like the molecules of a gas. Further, it was assumed that the average kinetic energy of an electron in a metal at temperature T is the same as the average translational kinetic energy of a molecule of gas at the same temperature T. These assumptions led at least to a qualitative explanation of the thermionic effect. In the electron theory of the conduction of electricity in metals the same two assumptions are used. In Chapter VIII on the Kinetic Theory of Gases, we have seen that the slowness of the diffusion of a gas into another gas, as when ammonia gas is released into the air of a room, is explained by the smallness of the mean free path of the gas molecules. It is supposed that the free electrons in a metal travel only short distances before they hit an atom, so that the mean free path of the free electrons in a metal is small. If now a battery is connected to the two ends of a wire, the free electrons in the metal begin to move under the action of the electric field produced by the difference of potential between the two ends of the wire. This motion is superimposed on the random motions of the electrons due to the temperature of the wire. Just as a difference of pressure produces a "wind" in gases, so a difference of electrical potential produces an electron "wind" in metals. This electron "wind" constitutes what we call electric current.

If V is the potential difference between the ends of a wire of length l, an electric field of strength V/l is produced in the wire. This field causes a force Ve/l to act on the electron, where e is the charge on the electron. This force produces an acceleration in the electron of amount Ve/lm, where m is the mass of the electron. If the mean free path of the electron is L and C is the root mean square velocity of the electron as given by the kinetic theory of gases, the average time interval between two consecutive collisions of an electron with atoms of the metal is L/C. Hence the distance moved by the electron in the direction of the field is $(\frac{1}{2}) \cdot (Ve/lm) \cdot (L/C)^2 = VeL^2/2lmC^2$. If we suppose that on the average a collision removes all wind velocity in the direction of the field and that the wind velocity must begin again from zero, the average distance moved in the direction of the field between the next two collisions is again $VeL^2/2lmC^2$. But the number of collisions per second for a given electron is C/L on the average, so that the distance which an electron moves in the direction of the field in one second is

$$u = \frac{VeL^2}{2lmC^2} \times \frac{C}{L} = \frac{VeL}{2mlC} \tag{10–17}$$

The quantity u may be considered as the "wind" velocity developed by the electric field V/l in the electrons. If the metal wire is of cross-section A, the number of electrons which in one second cross a plane taken at right angles to the direction of the electric field is nuA, where n is the number of electrons in unit volume of the metal. Since each electron carries a charge e, the charge crossing the plane in one second is $neuA$. This is the value of the current, so that the current

$$i = neuA = \frac{VAne^2L}{2mlC} \tag{10–18}$$

The quantity $Ane^2L/2mlC$ is a constant for a given temperature, so that i varies as V, and the theory we have just given thus leads to Ohm's law,

$$i = V/R \tag{10–19}$$

Comparing (10–18) and (10–19), we see that the resistance

$$R = \frac{2mC}{ne^2L} \cdot \frac{l}{A} \tag{10–20}$$

But, for a wire of length l and cross-section A,

$$R = \frac{\rho l}{A}$$

where ρ is the specific resistance of the material of the wire. Hence the specific resistance is given by

$$\rho = \frac{2mC}{ne^2L} \tag{10–21}$$

Now, C, the gas-kinetic velocity of the electrons, is related to the absolute temperature T according to

$$\left(\frac{1}{2}\right) mC^2 = \left(\frac{3}{2}\right) kT \tag{10–22}$$

so that

$$\rho = \left(\frac{2}{ne^2L}\right) \sqrt{3mkT} \tag{10–23}$$

where k is Boltzmann's constant. It is at this place that divergence of opinion arises. It is necessary to make some assumption concerning the relation of n and L to the temperature. If the product nL is assumed to be proportional to $1/C$ and therefore to $1/\sqrt{T}$, the specific resistance is proportional to $\sqrt{T} \times \sqrt{T}$ or to T. Making this assumption, we have

$$\rho = \text{const} \times T \tag{10–24}$$

This relation is in fairly good agreement with experiment, provided that the temperature is neither too high nor too low.

10.9. Electron Inertia in Metals.—If the electron theory of conduction of electricity in metals which we have described in Sec. 10.8 is correct, it should be possible to develop an electromotive force by suddenly stopping a wire which is moving with high speed in the direction of its length. Tolman and Stewart undertook to show the existence of this e.m.f. According to Sec. 10.8 the acceleration developed in an electron in a wire of length l by an electromotive force V is given by

$$a = \frac{Ve}{lm} \tag{10-25}$$

If a wire of length l which is moving with high velocity v is very quickly brought to rest, the electrons which are traveling with the wire suffer a negative acceleration a in being brought to rest. This acceleration of the electrons gives rise to a difference of potential V between the ends of the wire which is given by (10–25), so that

$$V = \frac{mla}{e} \tag{10-26}$$

This difference of potential continues while the wire is being brought to rest. The electrons, due to their inertia, which is measured by their mass m, tend to keep moving although the wire is being brought to rest. There thus tends to be an electron current in the wire in the direction of the original motion of the wire, while the wire is being brought to rest. The current i produced by the inertia of the electrons is

$$i = \frac{V}{R} = \frac{mla}{Re} \tag{10-27}$$

where R is the resistance of the wire. In the experiment of Tolman and Stewart, a flat coil of wire consisting of many turns of thin wire was set into fast rotation about an axis perpendicular to the plane of the coil and then suddenly stopped. Wires were connected to the two ends of the coil and taken along the axis of rotation to a sensitive balistic galvanometer. Upon suddenly stopping the rotation of the coil, a charge was found to flow through the galvanometer. If the coil takes a time t to come to rest, the total charge Q which passes through the galvanometer is it, so that

$$Q = \frac{mlat}{Re} = \frac{mlv}{Re} \tag{10-28}$$

where v is the original velocity of the rotating coil. The charge Q can be measured by the throw of the ballistic galvanometer.

The term "ballistic" refers to the way the galvanometer is used rather than to the construction of the galvanometer. If a current i flows for a time t, which is small compared with the vibrational period of the galvanometer, and then ceases, the galvanometer deflects over to a certain reading and then comes back to zero. This maximum deflection is known as the throw of the galvanometer. When a galvanometer is used this way it is said to be a ballistic galvanometer.

The torque acting upon the movable part of the galvanometer is proportional to the current i flowing through the galvanometer, so that the torque is equal to ki, where k is a proportionality constant. This torque produces an angular acceleration α such that

$$I\alpha = ki \tag{10-29}$$

where I is the moment of inertia of the movable part. If the torque acts for a time t, the angular velocity produced is

$$\omega = \alpha t = \frac{kit}{I} \tag{10-30}$$

At the end of the time t the movable part is moving with kinetic energy $(\frac{1}{2})I\omega^2 = k^2i^2t^2/2I$. However, as the deflection begins to increase the restoring torque due to the twist of the suspension begins to come into play. This restoring torque is equal to $K\theta$, where θ is the angle of twist of the suspension and K is a constant of proportionality. When the angle of twist is increased from θ to $\theta + d\theta$, the amount of work done against the restoring torque is $K\theta d\theta$. Hence the work done against the torque from $\theta = 0$ to $\theta = \theta_T$ is

$$W = \int_0^{\theta_T} K\theta d\theta = \frac{K\theta_T{}^2}{2} \tag{10-31}$$

In the case of a ballistic galvanometer the kinetic energy carries the movable part out to a deflection θ_T where all of the kinetic energy has been used up in doing work against the restoring torque. It is assumed that the interval t during which the movable part attains a kinetic energy $(\frac{1}{2})I\omega^2$ is small compared with the interval taken for the movable part to swing out to a deflection θ_T. Hence

$$\frac{k^2i^2t^2}{2I} = \frac{K\theta_T{}^2}{2} \tag{10-32}$$

Now it is the total charge Q flowing through the galvanometer, so that (10-31) becomes

$$Q = \theta_T\sqrt{\frac{KI}{k^2}} \tag{10-33}$$

The quantities under the square root sign are constant, so that Q is proportional to the angle of throw θ_T.

In Tolman's experiment the direction of the throw of the galvanometer depends upon the sign of e. Tolman found the direction to be that required by negatively charged electrons. Further, from the experimental values of Q, l, v and R, Tolman found the value of e/m to be 1.60×10^7 emu/gm for a copper wire. This is in reasonably good agreement with the value 1.76×10^7 emu/gm found for the electron by other methods. We must conclude therefore that the electric current in metals is carried by electrons.

10.10. Criticism of the Theory.—So far in this chapter we have applied the Maxwell distribution function for the component velocities of the molecules of a gas to the electrons in a metal. The treatment of the electrons in a metal as a gas having the Maxwell velocity distribution function has led to some formulas which are inconsistent with experiment. One of the great troubles of the theory has been to explain the value of the specific heat and that of the electrical resistance of a metal by means of one and the same set of assumptions. If we assume that the free electrons in a metal have the Maxwell distribution of velocities, each free electron must have an average kinetic energy $(3/2)kT$. The phenomena connected with electrical resistance seem to require that the number of free electrons be approximately equal to the number of atoms in the metal. Hence in a gram-atom of the metal the energy in the electrons is approximately $(3/2)NkT = (3/2)RT$. Differentiating with respect to T, we obtain $3R/2$ for the atomic heat due to the electrons. Adding this to the atomic heat due to the vibrations of the metal atoms we obtain $(3 + 3/2)R$ or 9 calories per gram-atom for the total atomic heat when the metal is above its Debye characteristic temperature. However, no such value of the atomic heat is found. Hence the assumption of a Maxwell distribution for the velocities of the electrons in a metal must be held in question. We shall return to this matter in a later chapter.

CHAPTER X

PROBLEMS

1. Find the velocities in cm/sec of electrons which have fallen through the following potential differences: 2, 10, 50, 250 and 10,000 volts.

2. Find the kinetic energy in ergs of an electron which has fallen through 1500 volts.

3. In the hook-up of Fig. 10.9 it is desired to obtain a current with a 5 per cent ripple through the resistance R which is 7 megohms, the alternating potential and frequency of the supply being respectively 40,000 volts and 500 cycles. This is accomplished by connecting a condenser in parallel with R. Determine the capacitance of the condenser.

4. In Richardson's formula $i = AT^2 \exp(-b/T)$ the values of A and b for tungsten are 60.2 amp/cm^2-deg^2 and 5.24×10^4 deg, when i is the current in amperes emmited by 1 cm^2 of the surface. Find the value of the current emitted by a tungsten filament of diameter 5 mils (1 mil = 0.001 inch) and length 3 cm when maintained at each of the following temperatures: 1000° C, 1500° C, 2000° C, 2500° C, the voltage between the filament and the surrounding cylinder in each case being sufficient to produce saturation.

5. The value of b for tungsten (see Problem 4) is 5.24×10^4 deg. Determine the value of the thermionic work function for tungsten in ev.

6. Oxide coated filaments are usually more efficient thermionic emitters than pure metals. The thermionic work function for a platinum filament coated with barium oxide is 1.68 ev while A in the Richardson equation is 2.88 amp/cm^2-deg^2. If the length of the filament is 2 cm and its diameter is 7 mils, find the current which is emitted by the wire when the temperature is maintained at 1000° C and the voltage between the filament and the cylinder is above the saturation value.

CHAPTER XI

THE PHOTOELECTRIC EFFECT

11.1. Experiments of Hertz.—In Sec. 5.9 we saw that the experiments of Hertz on electromagnetic waves gave final confirmation to the Maxwell electromagnetic theory of light. It is one of the ironies of physics that in these same experiments Hertz observed a phenomenon which cannot be explained by the electromagnetic wave theory of light. This phenomenon has led to the formulation of another theory of light—the quantum theory. In some of its aspects the quantum theory is a corpuscular theory. We may say then that the same experiments which gave final vindication to the wave-theory of light at the same time gave evidence for a kind of corpuscular theory.

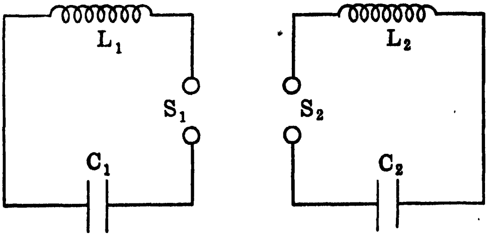


FIG. 11.1.—Experimental arrangement used by Hertz to show the existence of electromagnetic waves.

The diagram of the Hertz experiments described in Sec. 5.9, is reproduced in Fig. 11.1. Hertz not only found that the length of the sparks depends on how nearly L_2C_2 is made equal to L_1C_1 but also on whether the light from the spark at S_1 is allowed to fall upon the metal balls of the gap S_2 of the receiving set. Any opaque object between S_1 and S_2 diminishes the spark length at S_2. Glass shields S_2 but quartz does not. Quartz transmits ultraviolet light while glass does not. The effect is thus due to ultraviolet light. This discovery of the peculiar effect of ultraviolet light on metal surfaces was the forerunner of the discovery of the photoelectric effect.

11.2. Photoelectric Effect.—Hertz recognized the importance of his discovery of the peculiar effect of ultraviolet light. However, his principal interest was in his researches on electromagnetic waves, and he did no further work on this effect of light. It is to Hallwachs that the credit for the discovery of the photoelectric effect is usually given. In 1888 Hallwachs announced that a freshly polished zinc plate would lose its charge when charged negatively and illuminated by the ultraviolet light from an arc lamp but that the plate would not lose its charge when charged positively. In Secs. 7.1 and 7.2 we saw that the appearance of the discharge at the negative end of a low pressure discharge tube differs from that at the positive end, thus indicating a difference in the behavior of positive and negative electricity. We now see that there is a difference between the behavior of positively and negatively charged metal surfaces when illuminated by ultraviolet light.

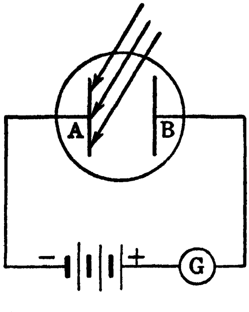


FIG. 11.2.—Apparatus for showing the photoelectric effect.

If two zinc plates A and B are placed in an evacuated quartz tube and are connected through a galvanometer to a battery as shown in Fig. 11.2, current will flow when A is illuminated by ultraviolet light, provided that A is connected to the negative terminal of the battery. If A is connected to the positive terminal, no current will flow unless the voltage of the battery is quite small. If A and B, instead of being connected to a battery, are connected to an electrometer, a difference of potential between A and B will be indicated by the electrometer. The potential will be such that A becomes positive with respect to B. The effect of the ultraviolet light is to expel negative but not positive electricity from the metal plate.

11.3. Photoelectrons.—In 1900 Lenard in Germany and Merritt and Stewart in America showed that the negative electricity which is expelled by ultraviolet light is in the form of electrons. To prove this, Lenard allowed ultraviolet light to pass through a quartz window Q, Fig. 11.3 and on to an aluminum plate P. Negative electricity is then liberated from P. If a fairly high difference of potential is applied between P and the metal diaphragm A, the negative charges speed up and pass through the hole in the diaphragm with high velocity. These negative charges then strike the electrode B, thus charging electrometer D. If now a magnetic field of the proper strength is applied to the space between A and B, the charges are deflected so as to strike the electrode

C, thus charging electrometer E. The electrodes P, B, and C are enclosed in an evacuated tube.

Between P and A the charges attain a velocity v given by

$$(\tfrac{1}{2})mv^2 = V'e' \tag{11-1}$$

where V' is the potential in esu between P and A and e' and m are the charge and mass of each particle respectively. When the magnetic field H is applied to the space between A and B in a direction perpendicular to the plane of the diagram, Fig. 11.3, the negative charges follow a circular path given by

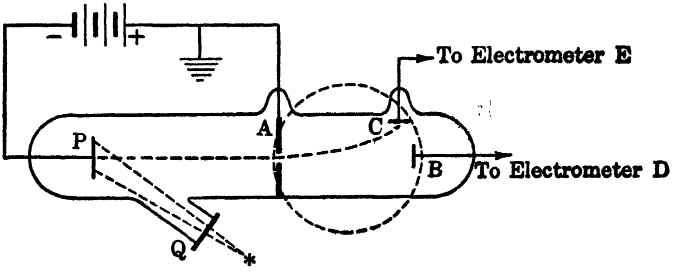


Fig. 11.3.—Apparatus for measuring e/m for photoelectrons.

$$\frac{He'v}{c} = \frac{mv^2}{r} \tag{11-2}$$

where r is the radius of the path determined by the geometry of the apparatus. From (11–1) and (11–2),

$$\frac{e'}{m} = \frac{2V'c^2}{H^2r^2} \tag{11-3}$$

Substituting experimental values in the right side of (11–3), the value of e'/m is found to be 5.28×10^{17} esu/gm. The particles are thus identified with Thomson's electrons. These particles are sometimes called photoelectrons not because their nature is any different from other electrons but merely to indicate their origin.

11.4. Stopping Potential.—If plate A, Fig. 11.2, is illuminated with light of given intensity and wave-length and if the voltage of the battery is varied, a curve similar to that shown in Fig. 11.4 is obtained. At small positive potentials of B a saturation current is obtained as is shown by the horizontal portion of the curve. At zero potential of B the current is a large fraction of the saturation value, the size of the fraction depending on the particular construction of the tube in Fig. 11.2. With a properly constructed apparatus, practically the full saturation current flows at zero potential of B relative to A. Further, it is seen that current flows for small negative potentials of B. This means that the electrons are expelled from plate A with a velocity great enough to overcome small negative potentials of B. As the negative potential of B is increased the current diminishes, finally becoming zero at a definite potential. This definite negative potential is known as the stopping potential because it is just sufficient to stop electrons which are expelled from plate A from passing over to plate B. Hence, if the fastest moving electrons are expelled from A with a velocity v, we have

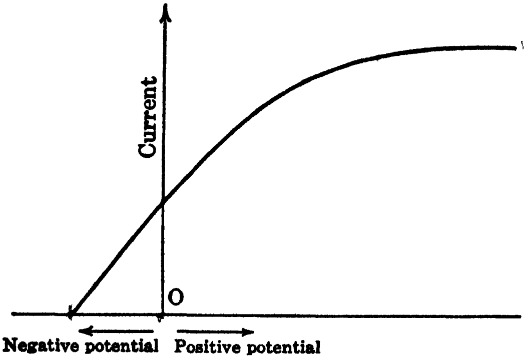


FIG. 11.4.—Variation of the photoelectric current with potential.

$$(1/2)mv^2 = V'e' \tag{11-4}$$

where V' is the stopping potential. If the stopping potential is measured in volts, the velocity v, as in (10–4), is given by

$$v = 5.94 \times 10^7 \times \sqrt{V_P} \tag{11-5}$$

where V_P is in volts. If the negative potential of plate B is increased above the stopping potential, no current is indicated by the galvanometer G, Fig. 11.2.

The shape of the curve for negative potentials of B is explained by supposing that the light penetrates a small way into the metal and that the photoelectric effect occurs at varying depths below the surface. Wherever it occurs the electrons are expelled with a certain definite velocity. However, in penetrating through the metal to the surface they lose energy and therefore velocity. Those photoelectrons having their origin in deeper layers emerge with a smaller velocity than those having their origin in the layers nearer the surface. The smaller the velocity the smaller the stopping potential, so that it is only the photoelectrons having their origin on the surface which give rise to the electrons of maximum velocity. As the negative potential of B is diminished, more and more electrons have sufficient velocity to get over to the plate B, so that in this portion of the curve of Fig. 11.4 the current increases almost linearly with decrease of negative voltage.

11.5. Effect of Light Intensity.—The first of the remarkable laws of the photoelectric effect is that, if the intensity of the light falling on plate A, Fig. 11.2, is varied while the wave length of the light remains constant, the maximum velocity of the electrons which are expelled from A is independent of the intensity of the light. On the classical electromagnetic theory of light, this expulsion of electrons must be due to the strength of the electric field which exists in a light wave. In Secs. 5.5 and 5.11 we have seen that the average energy per unit volume in an electromagnetic wave is $R'_A{}^2/4\pi$, where R'_A is the root mean square of the electric field strength in the wave. Hence, if the intensity of light is doubled, the field strength would be multiplied by $\sqrt{2}$, and it seems that the electrons should be expelled with a greater maximum velocity.

If an increase of the intensity of the light does not cause an increase of the maximum speed with which electrons are expelled, what effect, if any, does an increase of intensity have? It is found that the saturation current as registered by the galvanometer G, Fig. 11.2, is exactly proportional to the intensity of the light illuminating A, provided that the wave-length of the light is kept constant. Thus, doubling the intensity of the light doubles the rate at which electrons are expelled, but has no effect on the velocity with which they are expelled.

11.6. Effect of Wave-Length.—The second of the remarkable laws of the photoelectric effect is that, if the wave-length of the light illuminating plate A, Fig. 11.2, is diminished, the stopping potential and therefore the maximum velocity of expulsion of the photoelectrons is increased. It has been found more convenient to speak in terms of frequency rather than wave-length, and from now on we shall use the term frequency.

An increase of frequency produces an increase of the maximum velocity of expulsion. If, however, the frequency is decreased, the photoelectric effect disappears at a certain definite frequency known as the threshold frequency. For frequencies below the threshold frequency no electrons are expelled from the metal. It is found that clean and polished metal surfaces have different threshold frequencies depending on the particular metal used. Even for a particular metal the threshold frequency varies unless great care has been taken in the outgassing as described in Sec. 10.6.

For many metals the threshold frequency is in the ultraviolet so that these metals do not show the photoelectric effect for visible light. However, the threshold frequency for some metals, as for instance sodium and potassium, is in the infrared, and these metals show the photoelectric effect for visible light. It is for this reason that sodium and potassium are used in the construction of photoelectric cells which are used in television, some kinds of sound motion pictures, etc.

11.7. Quantum Theory.—The effects described in the previous two sections cannot be explained in terms of the classical electromagnetic wave theory of light. We are compelled to seek an explanation in another theory. We have described in Sec. 9.11 how Planck was able to explain the facts about black body radiation on the basis of absorption and emission of radiation in quanta. In Planck's explanation these quanta only appeared in the processes of absorption and emission; the radiation itself did not consist of light quanta. In Sec. 9.11 we have spoken of the number of vibrational modes or standing waves in a uniform temperature enclosure with frequency ν in the range $d\nu$. We were evidently speaking in terms of a wave theory of light and heat radiation. However, in 1905 Einstein proposed that the light or heat radiation itself consisted of quanta. A quantum of light on this view is a kind of corpuscle of light. More recently it has come to be known as a photon.

The difficulties which appear when it is attempted to explain the photoelectric effect in terms of the classical wave theory vanish when a corpuscular type of theory is used. Following Newton, we think of a beam of light as a rain of corpuscles. When this rain of corpuscles falls upon a metal surface, some of the metal atoms are hit by the corpuscles and some not. Both the atoms and the corpuscles being small, only a small proportion of the atoms are hit. According to Einstein's form of the quantum theory, each corpuscle or quantum of light carries with it an energy $h\nu$ where ν is the frequency and h is a proportionality constant known as Planck's constant. Right here we observe an inconsistency. The quantum theory is a kind of corpuscular theory, yet the energy of a corpuscle is measured in terms of the frequency of the light. Frequency is a term which is used in the wave theory. The frequency of light is obtained by dividing the velocity of light by its wave-length and the wave-length is measured by means of a diffraction grating. The action of a diffraction grating can only be interpreted in terms of a wave theory. Hence, right at the start, the quantum theory depends upon a measurement which can only be interpreted in terms of a wave theory. We pass over this inconsistency and simply assume that the energy of a quantum is $h\nu$. When a quantum hits an atom it completely gives up its energy to the atom. In the photoelectric effect this energy is used up in expelling the electron from the atom. According to Einstein a certain amount of energy W is required to remove an electron from an atom or from a metal surface, so that the energy available for giving kinetic energy to the electron is $h\nu - W$. We thus obtain Einstein's photoelectric equation

$$(\tfrac{1}{2})mv^2 = h\nu - W \qquad (11\text{–}6)$$

where m and v are the mass and velocity of the photoelectron respectively. The velocity measured is the maximum velocity. Since the maximum kinetic energy is proportional to the stopping potential, Einstein's equation shows that, if a metal surface is illuminated with light of varying frequency, the stopping potential should vary linearly with the frequency. Experiments to test this prediction were performed in 1911 and 1912 by Hughes at the Cavendish Laboratory, England and by Richardson and K. T. Compton at Princeton. On plotting the stopping potential V against the frequency, they obtained a straight line as shown in Fig. 11.5.

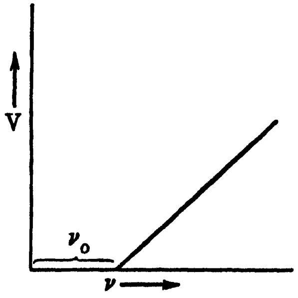


FIG. 11.5.—Graph of the stopping potential against the frequency of the light causing the photoelectric effect.

In terms of stopping potential (11–6) may be written

$$V = \left(\frac{h}{e}\right)\nu - \frac{W}{e} \qquad (11\text{–}7)$$

If V is plotted against ν, the coefficient h/e is the slope of the line shown in Fig. 11.5. Hence

$$h = e \tan \theta \tag{11-8}$$

where θ is the angle between the line and the axis of ν in Fig. 11.5. We already know the value of e the charge on the electron, so that by measuring $\tan \theta$ we can obtain a value of h. Using this method Millikan in 1916 determined h with great accuracy and obtained the value of 6.62×10^{-27} erg · sec. This agrees with the value found from Stefan's constant as described in Sec. 9.11 and with that found from the constant in Wien's displacement law as described in Sec. 9.13.

11.8. Photoelectric Work Function.—By referring to Fig. 11.5 it is seen that the threshold frequency is measured by the length of the intercept on the frequency axis. If we call the length of this intercept ν_0, (11–7) becomes

$$Ve = h(\nu - \nu_0) \tag{11-9}$$

Now $h\nu_0$ or W represents the work necessary to remove an electron from a metal and is known as the photoelectric work function. It is convenient to measure this energy in terms of electron-volts by means of the relation

$$h\nu_0 = (V_0/300)e' \tag{11-10}$$

Du Bridge after carefully outgassing platinum obtained 6.30 electron-volts for the photoelectric work function of platinum. This agrees very well with the value given in Sec. 10.6 for the thermionic work function of platinum.

11.9. A Relation between Energy and Wave-Length.—The wave-length of light is usually given in angstroms, the relation between an angstrom and a centimeter being

$$1 \text{ angstrom} = 10^{-8} \text{ centimeter.}$$

The abbreviation for angstrom is A. For instance, the wave-lengths of the D lines in the spectrum of sodium are respectively 5890 A and 5896 A. Now, if the threshold frequency for a hypothetical metal is zero, (11–9) reduces to

$$Ve = h\nu \tag{11-11}$$

But, since the wave-length λ is related to the frequency ν according to

$$\lambda = \frac{c}{\nu}$$

(11–11) becomes

$$\lambda = \frac{hc}{Ve} \tag{11–12}$$

where λ is in centimeters and V and e in emu. If, however, V is expressed in volts and λ in angstroms, (11–12) reduces to

$$\lambda \text{ (in angstroms)} = \frac{12345}{V \text{ (in volts)}} \tag{11–13}$$

The value of the constant 12345 is somewhat uncertain but making it 12345 makes the constant easier to remember, so that the relation between wave-length and voltage is usually remembered as in (11–13). We shall have occasion to make use of (11–13) later on. As we shall see later, it has become the practice to speak of a *photon* of light. The energy of a photon of wave-length λ is given by (11–13), where V is measured in electron-volts rather than volts.

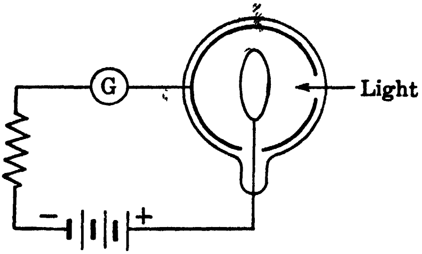


Fig. 11.6.—Diagram of a photoelectric cell and connections.

11.10. The Photoelectric Cell.—An evacuated glass or quartz tube as in Fig. 11.6 has its inner surface coated with sodium or potassium. A window is left through which light can enter the tube and a wire loop for collecting the electrons is situated as shown. Such a tube is called a photoelectric cell. When light enters the cell the galvanometer G shows a current, and this current is proportional to the intensity of the light provided that more than the saturation voltage is applied to the cell. Recently, it has become the practice not to evacuate the tube completely but to place an inert gas (helium or argon) in the tube and then evacuate to a pressure of a few millimeters. A potential of about 100 volts is then applied across the tube. In this case when a photoelectron is expelled from the sodium or potassium surface the electron is accelerated by the voltage to such a velocity that it is able to knock other electrons off the atoms of the gas in the cell. The atoms then become charged positively while the electrons which are knocked off carry a negative charge. The gas is then said to be ionized. The positive ions proceed to the sodium or potassium surface and the negative ions or electrons proceed to the wire loop. The ionization of the gas produced in this way is called ionization by collision. It multiplies the effect of the light on the sodium or potassium surface, so that a photoelectric cell containing gas gives several times as much current for the same intensity of light as a cell containing no gas.

11.11. Effect in Non-Metals.—We have so far spoken only of the photoelectric effect in metals. However, the effect is shown, not only by metals but also by non-metals and by liquids and gases. Thus far we have spoken only of metals because the photoelectric effect is not so easy to demonstrate for insulators. When a photoelectron is ejected from an insulator, the surface of the insulator becomes charged positively and we cannot discharge the insulator by connecting to ground as we would a metal. The effect of the positively charged surface is to prevent any more electrons from leaving the surface. Also for most insulators the threshold frequency is very far in the ultraviolet and this again makes it more difficult to observe the effect. It is only a matter of obtaining light of short enough wave-length and any substance will show the photoelectric effect. X-rays are of the same nature as light, differing from light only in their very short wave-length. All substances exhibit the photoelectric effect when illuminated by x-rays—in particular, the effect for gases is very noticeable.

In Sec. 7.7 we stated that Millikan charged his oil drops by illuminating them with ultraviolet light or x-rays. We see now that the effect of this illumination is to expel one or more electrons from the drops, leaving them charged positively. When x-rays expel electrons from air molecules, the remaining molecule or atom is charged positively. The expelled electron may attach itself to an oil drop, so that the drops may be charged either positively or negatively.

11.12. Contact Difference of Potential.—In measurements of the stopping potential, the plates *A* and *B*, Fig. 11.2 must be of the same metal to avoid complications arising from another effect—that of contact difference of potential. This phenomenon is usually described in first-year textbooks of physics. For our present purpose we shall describe the phenomenon in the following way: Consider two metal plates *A* and *B* as in Fig. 11.7 where the plates *A* and *B* are made of different metals, say, copper and zinc respectively. The plates may be in a vacuum or may be in air. Plate A is connected to a sensitive electroscope E and plate B to a potentiometer. The resistance CD is grounded as shown at O, so that by sliding the contact F along CD the plate B can be connected to either a positive or negative potential with respect to ground. K represents a grounding key which when closed is in contact with the wire connected to the plate A and to the electroscope E. We begin with F at O and the key K closed, so that A and B are both connected to ground. The key K is now opened, so that the plate A and the electroscope E are disconnected from ground. If next the plate B, which remains connected to ground, is moved toward or away from the plate A, it is found that the electroscope shows a charge, the sign of which depends upon the direction of the movement of B. B is said to be electro-positive toward A, if, upon moving B towards A, the electroscope shows a positive charge. The experiment is now repeated with the sliding contact F at some other point than O, so that the plate B is not connected to ground but to some potential which is positive or negative with respect to ground. It is found that, for a particular potential of F, the plate B may be moved toward or away from the plate A without the electroscope E showing a charge when the grounding key K is open. If B is electro-positive to A, this particular potential is negative, say, $-V_c$. The contact difference of potential (for brevity, contact p.d.) of the metal of plate B relative to the metal of plate A is then $+V_c$.

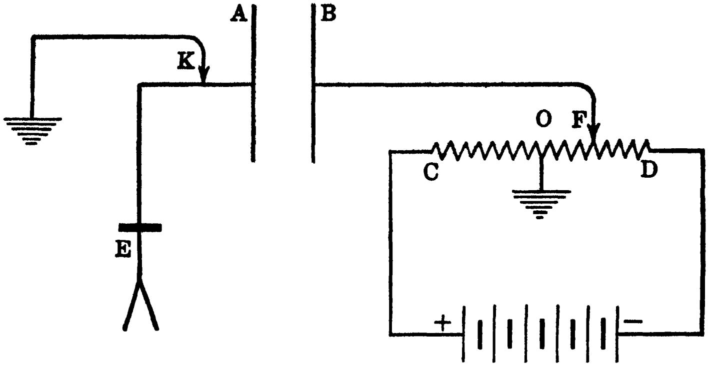


Fig. 11.7.—Apparatus for showing contact difference of potential.

If in Fig. 11.2 plates A and B consist of two different metals, and a curve similar to Fig. 11.4 is obtained, it is found that the measured stopping potential is less or greater than the true stopping potential (i.e., for the case when A and B are of the same metal) by the contact p.d. between the two metals. The contact p.d. is very susceptible to changes in the surfaces of the plates A and B, Fig. 11.7. Even when the plates are of the same metal there is a contact p.d. if the surfaces of of the two plates have undergone different treatments. It is on account of this surface effect that prolonged outgassing of the plates A and B, Fig. 11.2 is necessary in order to obtain an accurate value of the stopping potential, even though both plates are of the same metal.

The contact p.d. between two metals is numerically equal to the difference of the work functions (photoelectric or thermionic) of the two metals when these are measured in electron-volts, the metal with the greater work function being electro-negative with respect to the metal with the smaller work function.

11.13. Bose-Einstein Theory.—We have seen that the photoelectric effect requires a kind of corpuscular theory of light for its explanation, but that Planck in his derivation of the distribution law for the heat radiation inside a uniform temperature enclosure made no use of the idea of corpuscles of light or heat radiation. Accordingly in 1924–25 Bose and Einstein attacked the problem of the distribution function for the energy density of the radiation in a uniform temperature enclosure from a new point of view. They treat the radiation inside the enclosure as a sort of a gas made up of quanta (or "photons" as quanta are now called). These photons, or quanta, are supposed to be rushing about in all directions within the enclosure with all sorts of energies. However, there is an important difference between an ordinary gas like oxygen and this photon "gas" which we are now describing. Whereas the molecules of a gas rush about with different velocities and with different energies, the photons all travel with the same velocity c but with different energies $h\nu$. Bose and Einstein have shown that the number of photons per unit volume whose frequencies are between ν and $\nu + d\nu$ is $G(\nu)d\nu$, where

$$G(\nu) = \frac{8\pi\nu^2}{c^3\left\{\exp\left(\frac{h\nu}{kT}\right) - 1\right\}} \qquad (11\text{–}14)$$

where k is Boltzmann's constant, and T is the absolute temperature. Since the energy of each photon in the frequency range ν to $\nu + d\nu$ may be taken as $h\nu$, the total energy of the $G(\nu)d\nu$ photons is $h\nu G(\nu)d\nu$. But this energy is also represented by $I(\nu)d\nu$, so that

$$I(\nu) = \frac{8\pi h\nu^3}{c^3\left\{\exp\left(\frac{h\nu}{kT}\right) - 1\right\}} \tag{11-15}$$

which is in agreement with Planck's formula (9–26) and so also with the experimental curve.

CHAPTER XI

PROBLEMS

1. The photoelectric thresholds of particular samples of certain metals are as follows: aluminum, 4770 A; copper, 3000 A; potassium, 6000 A; sodium, 6800 A; tungsten, 2300 A. Determine the photoelectric work function for each of these samples.

2. Determine the stopping potential in volts for each of the metals in Problem 1 when the metal is illuminated (*a*) with light of wave-length 1849 A and (*b*) with light of wave-length 3500 A.

3. The thermionic work functions for particular samples of certain metals are as follows: silver, 3.64 ev; nickel, 4.05 ev; lithium, 2.13 ev. Determine the respective threshold wave-lengths and also the threshold frequencies.

4. Determine the maximum velocity in cm/sec of the electrons ejected from each of the metals given in Problem 3 when the metal is illuminated with light of wave-length 2536 A.

CHAPTER XII

ELECTRONICS AND RADIO

12.1. Current-Voltage Characteristics.—For a metal wire maintained at a constant temperature, the current i through the wire is proportional to the potential V across the wire. In such a case we have Ohm's law

$$i = V/R \tag{12-1}$$

where $1/R$ is a constant of proportionality and R is known as the resistance. If i is plotted against V, we obtain a straight line as shown in Fig. 12.1. The angle of slope $\theta = \tan^{-1}(1/R)$. Any part of a circuit, usually external to the battery, is known as a *circuit element*. When the graph of i vs V for a circuit element is a straight line passing through the origin the circuit element is said to be *linear*. The idea of linear circuit elements is extended to alternating currents. We have noted in Chapter IV that the complex current a is proportional to the complex potential v across a resistance, an inductance, or a capacitance, although there may be a difference of phase between a and v. Air-core inductances and capacitances as well as resistances are called linear circuit elements in the theory of alternating currents.

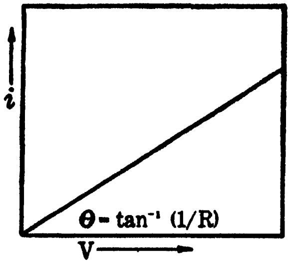


FIG. 12.1.—Ohm's law.

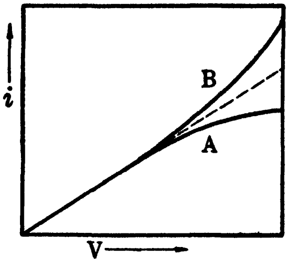


FIG. 12.2.—Departure from Ohm's law.

But there are many circuit elements for which i is not proportional to V. Such circuit elements are called *non-linear* or *non-ohmic* circuit elements. For these Ohm's law does not hold. It does not hold for pure metals when these are caused to rise in temperature by the current passing through them. The graph of i vs V for a tungsten wire which is allowed to rise in temperature is shown in Curve A of Fig. 12.2. This is the case of a tungsten filament in an ordinary lamp. Curve B of Fig. 12.2 is for a carbon filament lamp. We say that the resistance of a tungsten filament increases whereas that of a carbon filament decreases with rise of temperature. It is only for very small temperature rises that tungsten and carbon filaments obey Ohm's law as shown by the tangency of Curves A and B at the origin to the dotted straight line in Fig. 12.2.

The graphs appearing in Figs. 12.1 and 12.2 are known as *current-voltage characteristics*, or, briefly, *characteristics*. We have shown the characteristics of certain circuit elements in Chapters X and XI: the characteristic of a kenotron or diode in Fig. 10.2 and that of a photoelectric cell in Fig. 11.4. In addition to these we shall now describe the characteristic of a black ceramic material called *thyrite*. The equation of the characteristic of this material is very nearly given by

$$i = AV^{3.5} \tag{12-2}$$

where A is a constant. If we say that V/i is the resistance R of a thyrite disc, then

$$R = 1/(AV^{2.5}) \tag{12-3}$$

and the resistance of this non-ohmic disc diminishes very rapidly with increase of the voltage across it. For a thyrite disc whose diameter is 6 in and thickness 0.75 in, the resistance at 300 volts is 3300 ohms whereas the resistance at 2100 volts is 25.5 ohms. Such a disc connected in parallel with a part of a circuit protects this part from an extreme surge of voltage such as that produced by a lightning stroke.

It is seen that Ohm's law does not hold for many circuit elements. Instead of i being proportional to V, about the best we can say is that i is a function of V.

12.2. Circuit Elements in Series.—Consider the circuit shown in Fig. 12.3. It is desired to find the readings of the ammeter A and of the voltmeters V_C and V_R. In order to solve this problem we must know the values of E_B and R, and also the characteristic of the non-ohmic circuit element. Let us suppose that the circuit element is a kenotron or diode whose characteristic is given by Fig. 10.2. The graph of this figure is reproduced as the curve OFG of Fig. 12.4. Let the graph of

Fig. 12.1 be the characteristic of the resistance. We place the origin of this graph at *H* in Fig. 12.4, where *OH* represents E_B, the e.m.f. of the battery in Fig. 12.3. The plane of Fig. 12.1 is then rotated through 180° about the axis of *i*. This gives us the straight line *HF* in Fig. 12.4. The line *HF* and the curve *OFG* intersect at *F*. The ordinate *DF* of the point of intersection gives the reading of the ammeter *A* in Fig. 12.3

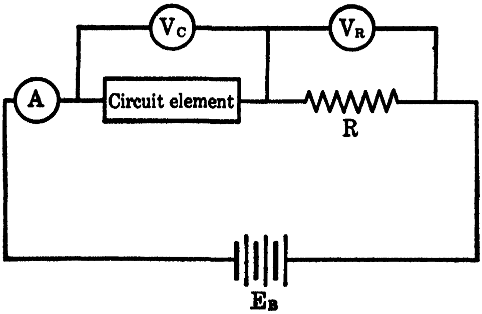


FIG. 12.3.

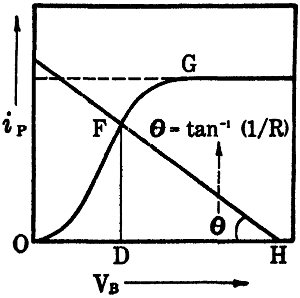


Fig. 12.4.

and the lengths *OD* and *DH* represent the readings of the voltmeters V_C and V_R, respectively. When a resistance is in series with a non-ohmic circuit element as in Fig. 12.3 and when Fig. 12.4 is obtained by the method just explained, the straight line *HF* which makes an angle $\theta = \tan^{-1}(1/R)$ with *HO* is frequently known as a *load line.*

The method just described for a non-ohmic and an ohmic circuit element in series can be applied when the ohmic element is replaced by a second non-ohmic element. If the characteristics of the two circuit elements can be expressed by algebraic equations, a solution can be obtained by the methods of algebra. Usually, however, the characteristic of a non-ohmic element cannot be so expressed and a graphical solution is the only one possible. Of course, an experimental solution can be obtained.

12.3. Three-Electrode Tube or Triode.—In order to make modern radio possible it was necessary to discover the properties of various kinds of vacuum tubes. We have described the construction and use of the diode or kenotron, which is a form of vacuum tube, in Sec. 10.7. A diode contains a filament, which is heated by an "*A*" battery, and a plate. A voltage is applied between the filament and the plate by means of a "*B*" battery. In 1907 it occurred to De Forest, an American inventor, to place a grid (or wire gauze) between the filament and the plate and to determine the effect of raising the grid to different potentials. A tube which contains a filament, a grid and a plate is known as a three-electrode tube or a *triode*. The connections for a three-electrode tube are shown diagrammatically in Fig. 12.5. The *A* battery is used to heat the filament *F* and is of low voltage (about 2 to 10 volts); the *B* battery is used to raise the plate *P* to a high potential (50 to 500 volts); and the *C* battery is used to put various voltages on the grid. The positive terminal of the *B* battery is connected to the plate while *P*, either the positive or negative terminal of the *C* battery is connected to the grid. We recognize three circuits in Fig. 12.5: the filament circuit, the plate circuit, and the grid circuit. Usually there is no current in the grid circuit.

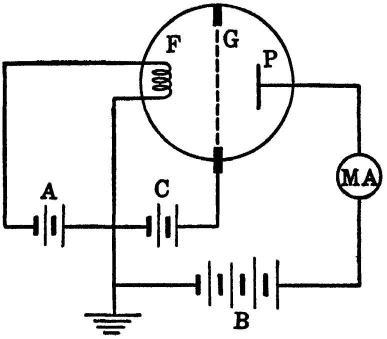


Fig. 12.5.—Connections for the three-electrode tube.

In Chapter X, when we were studying thermionics, we considered the current between *F* and *P*. It was noted that as the temperature of *F* is raised with a given difference of potential between *F* and *P* (*G* being absent) the current increases until a constant value is reached. This constant value is brought about by the action of the space charge which accumulates between *F* and *P*. De Forest believed that this space charge could be partially or wholly removed by placing a "grid" between *F* and *P*. The action of the grid depends on the potential of the grid and so on the voltage of the *C* battery.

12.4. Characteristics of a Triode.—There are two types of characteristics: *plate characteristics* and *transfer characteristics*. The grid potential V_G is the difference of potential between the grid and the negative end of the filament; the plate potential V_P is the difference of potential between the plate and the negative end of the filament.

To obtain a plate characteristic of a triode, V_G is kept constant and V_P is varied. A graph of the plate current i_P vs V_P is plotted, and this is the plate characteristic for the particular constant value of V_G. A tube manual such as that published at intervals by the Radio Corporation of America gives a family of plate characteristics for different values of V_G for a given type of tube. A family of such characteristics is shown in Fig. 12.6 for a type 6L6 tube when it is connected as a triode. Each curve is for the value of E_C shown on the curve. The grid circuit is also known as the C circuit since it contains the C battery. Hence E_C, the e.m.f. of the grid battery, as well as V_G is written for grid volts, when there is no grid current.

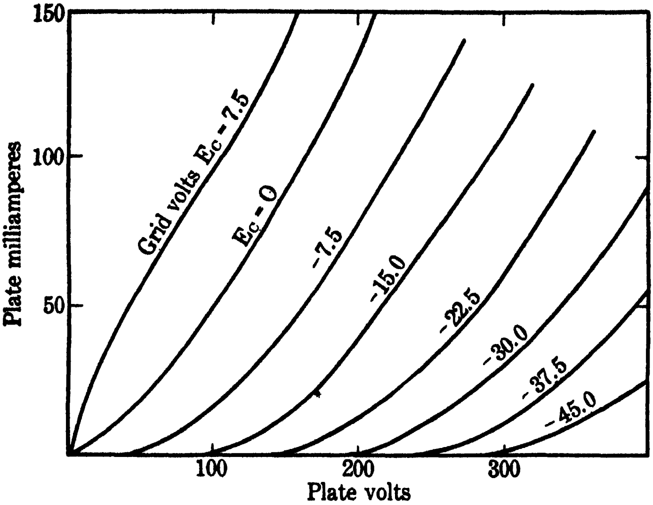


FIG. 12.6.—Family of plate characteristics for a 6L6 tube connected as a triode.

To obtain a transfer characteristic, the plate potential V_P is kept constant and V_G (or E_C) is varied. A graph of plate current i_P vs V_G is plotted, and this is the transfer characteristic for a certain constant value of V_P. It should be noted that a transfer characteristic gives the relation between the current in one circuit (the plate circuit) and the potential across a part of a second circuit (the grid circuit). If a family of plate characteristics for a given type of tube is available, a family of transfer characteristics can be constructed. Referring to Fig. 12.6, we can read off the values of i_P for various values of E_C for, say, $V_P = 200$ volts. These values are shown in Table 12.1. Plotting i_P against V_G, we obtain the transfer characteristic shown in Fig. 12.7. We note that the transfer characteristic cuts the axis of grid volts at -31 volts. This grid potential is called the *cutoff* potential. No current flows in the plate circuit when the grid potential is more negative than the *cutoff* potential.

TABLE 12.1

V_P = 200 volts

i_P / ma	V_G / volts
1.2	−30.0
12.0	−22.5
40.0	−15.0
80.0	− 7.5
136.0	0 0

Another transfer characteristic could have been constructed for $V_P = 150$ volts. The cutoff would then have occurred at $V_G = -23$ volts. We note that for a 6L6 tube used as a triode the quotient

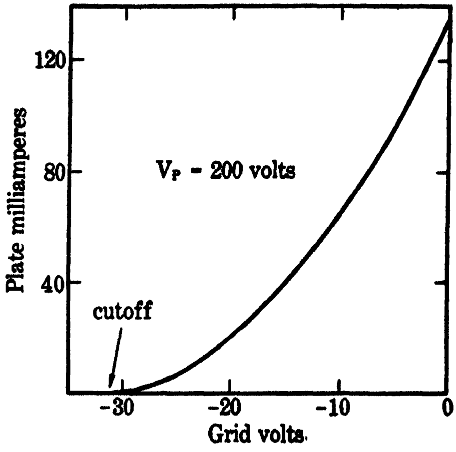


FIG. 12.7.—Transfer characteristic for a 6L6 tube connected as a triode.

$-V_P/V_{\text{cutoff}}$ is nearly constant, namely 6.5. Different types of tubes have different kinds of cutoff: sharp, medium, and remote cutoffs. The cutoff in Fig. 12.7 may be described as *medium*. The transfer characteristic of Fig. 12.7 tends to become tangent to the axis of grid volts. For many purposes it is desirable to have a straight rather than a curved transfer characteristic. If the characteristic is a straight line, it cuts the axis of grid volts at a definite point and the cutoff is *sharp*. If the characteristic approaches the axis of grid volts tangentially, the cutoff is said to be *remote*.

12.5. Heater Type of Triode.—It is impossible to obtain a sharp cutoff when the cathode of a triode is a filament heated by a current passing through it. Figs. 12.6 and 12.7 apply to a 6L6 type tube. For this tube the potential drop across the filament is 6.3 volts. The meaning of $V_G = -22.5$ volts is that it is the potential of the grid relative to the negative end of the filament. The potential of the grid relative to the positive end of the filament is $-22.5 - 6.3 = -28.8$ volts so that there is no one definite difference of potential between the grid and the filament. It is desirable for many purposes to have all points of the cathode at the same potential. Such a cathode must be indirectly heated. It usually consists of an oxide-coated cylindrical sleeve of metal within which is a heating filament which is not in electrical contact with the sleeve. The hot filament heats the cathode to a temperature sufficient to produce an adequate supply of electrons by thermionic emission. An oxide-coated metal has a much lower thermionic work function than a pure metal. The work function of an oxide-coated cathode is less than 1 ev, whereas that of pure tungsten is 4.52 ev.

12.6. Plate Resistance.—The resistance of a non-ohmic circuit element may be defined as V/i where i is the current through the element and V is the potential across the element. However, it frequently happens that a voltage consisting of a d.c. (direct current) component and an a.c. component is applied to a circuit element. We shall call the d.c. component the operating voltage and represent this by V_O. If the amplitude of the a.c. component is small compared with V_O, we may represent the instantaneous value of the a.c. component by dV. This increment dV causes an increment di in the current. It has become the custom in many problems to call dV/di the resistance of the non-ohmic circuit element at the operating potential V_O. Graphically we may think of a point P on the characteristic of a circuit element as shown in Fig. 12.8. This point P under the influence of an alternating potential of small amplitude oscillates along a small arc DF of the characteristic. This small arc corresponds to the tangent to the characteristic at the operating point E, whose abscissa is OB, representing V_O. The reciprocal of the slope of this tangent is dV/di, the resistance to the a.c. component of the current.

A triode is a non-ohmic circuit element. The reciprocal of the slope of one of the plate characteristics of Fig. 12.6 at a certain point (the operating point) is the plate resistance. Hence, the plate resistance R_P is given by

$$R_P = dV_P/di_P, \qquad V_G = \text{constant} \tag{12-4}$$

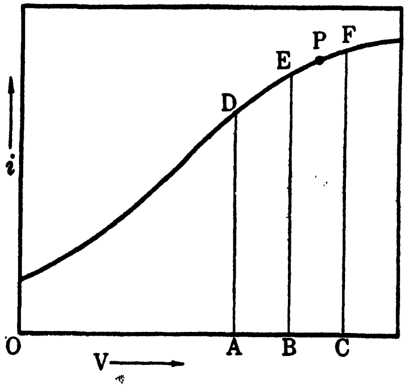


Fig. 12.8.

From Fig. 12.6 we see that, using the curve for $E_C = -7.5$ volts, the reciprocal of the slope at $V_P = 200$ volts is about 1.32 volts/ma = 1320 volts/amp = 1320 ohms. Again using the curve for $E_C = -7.5$ volts we note that the slope is less at $V_P = 100$ volts, and so the plate resistance at an operating plate potential of 100 volts is greater than that at an operating plate potential of 200 volts. Moreover, the plate resistance depends on the particular E_C curve used.

12.7. Transconductance.—This is also known as mutual conductance. The transconductance g_m of a triode is defined as

$$g_m = di_P/dV_G, \qquad V_P = \text{constant} \tag{12-5}$$

It is the slope at the operating point of a transfer characteristic such as that shown in Fig. 12.7. This transfer characteristic is for a 6L6 tube used as a triode with $V_P = 200$ volts. At an operating potential of $V_G = -15$ volts, $g_m = 4.53$ ma/volt = 0.0453 amp/volt. The unit, ampere per volt, is called either a reciprocal ohm (ohm^{-1}) or a mho ("ohm" spelled backwards). So $g_m = 0.0453$ ohm^{-1} or mho. However, since g_m is frequently much smaller than unity, a unit called the millimho (0.001 mho) is used. Then g_m in the foregoing is 45.3 millimhos.

12.8. Amplification Factor.—Consider a plate current i_P produced by particular operating values of V_P and V_G. Now imagine V_P increased by dV_P. In order to keep i_P at its original value it is necessary to change V_G by an increment dV_G. The amplification factor μ is defined as

$$\mu = -dV_P/dV_G, \qquad i_P = \text{constant} \tag{12–6}$$

The negative sign appears because if dV_P is an increment, dV_G is found to be a decrement if i_P is to be kept constant. So dV_G itself is negative. The negative sign on the right of (12–6) causes μ to be a positive number.

In many cases the plate current of a triode is approximately given by

$$i_P = (V_P + \mu V_G)/R_P \tag{12–7}$$

Differentiating, we have

$$di_P = (dV_P + \mu dV_G)/R_P$$

From this we have

$$di_P = \mu dV_G/R_P \quad \text{when} \quad dV_P = 0.$$

Hence, from (12–5),

$$g_m = \mu/R_P \tag{12–8}$$

This is a useful relation.

12.9. Discharge of a Condenser.—Before we consider the oscillating circuits used in radio, we shall first study the discharge of a condenser of capacitance C through an inductance L and a resistance R, the hook-up being as shown in Fig. 12.9. During the discharge under certain circumstances an oscillating current flows in the circuit. We shall represent this oscillating current by the complex quantity a. From Sec. 4.6, it is seen that the sum of the back e.m.f.'s is given by the left side of (12–9)

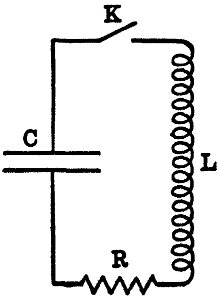


FIG. 12.9.—Oscillatory circuit.

$$(R + jL\omega - j/\omega C)a = 0 \tag{12–9}$$

where $\omega = 2\pi f$ and f is the frequency. Since there is no applied e.m.f., the right side of (12–9) is zero. From (12–9), either

$$a = 0 \tag{12–10}$$

or

$$R + jL\omega - j/\omega C = 0 \tag{12–11}$$

The meaning of (12–10) is that no current flows when the switch K has been closed for a long time and no external alternating e.m.f. is connected to the circuit. This solution does not interest us. Eq. (12–11) applies to the case of a condenser carrying a charge Q coulombs until time $t = 0$. At $t = 0$ the switch K is closed and the condenser begins to discharge through L and R. Multiplying (12–11) throughout by $\omega C/j$, we obtain

$$LC\omega^2 - jRC\omega - 1 = 0 \tag{12–12}$$

This is a quadratic in ω of which the solution is

$$\omega = \pm \frac{1}{\sqrt{LC}} \cdot \sqrt{1 - \frac{R^2C}{4L}} + j\frac{R}{2L} \tag{12–13}$$

We take the plus sign of the ambiguity, so that (12–13) has the form

$$\omega = \omega_0 + j\alpha \tag{12–14}$$

12.10. Complex Frequency.—In the last section we obtained a frequency $f = \omega/2\pi$ which is complex since ω is complex. We now ask what is the meaning of a complex frequency. In Secs. 4.3 to 4.5 we have used the symbol p to denote the rotating vector defined by

$$p = \sqrt{2}\,(\sin \omega t - j \cos \omega t) \tag{12–15}$$

Now, putting $\omega = \omega_0 + j\alpha$, we obtain

$$p = \sqrt{2}\,(\cos j\alpha t + j \sin j\alpha t)\,(\sin \omega_0 t - j \cos \omega_0 t) \tag{12–16}$$

From trigonometry

$$\cos \theta + j \sin \theta = e^{j\theta}$$

so that

$$\cos j\alpha t + j \sin j\alpha t = e^{-\alpha t}$$

Hence (12–16) becomes

$$p = \sqrt{2}\, e^{-\alpha t}\,(\sin \omega_0 t - j \cos \omega_0 t) \tag{12–17}$$

This represents a vector rotating with angular velocity ω_0 but whose magnitude is continually decreasing with the time according to the formula $\sqrt{2}\, e^{-\alpha t}$. The real part of the rotating vector is

$$P_{\text{real}} = \sqrt{2}\, e^{-\alpha t} \sin \omega_0 t \tag{12-18}$$

a graph of which is shown in Fig. 12.10. The amplitude of the oscillations thus grows smaller as time goes on. If α in (12-17) and (12-18) is negative, the amplitude of the oscillations grows larger as time goes on, as shown in Fig. (12.11). Hence, if α in a complex frequency is negative, oscillations can be built up; but, if α is positive, the oscillations die out. In a later section we shall discuss the oscillations of current in a vacuum tube circuit. We shall make use of the rule that, in order for oscillations to be built up, α must be negative.

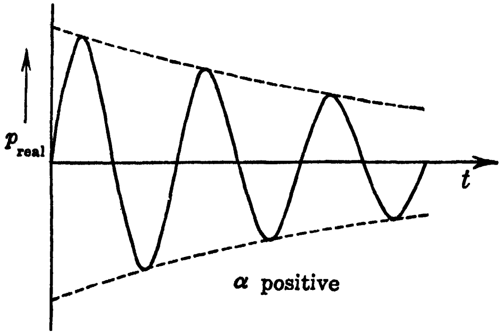


FIG. 12.10.—Oscillations with decreasing amplitude.

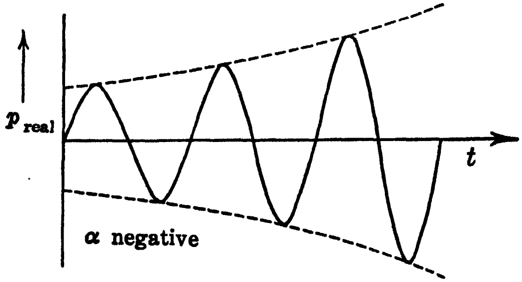


FIG. 12.11.—Oscillations with increasing amplitude.

When we were studying alternating currents in Chapter IV, the end P of the rotating vector described a circle as shown in Fig. 12.12a.

When a complex frequency is involved, the end of the rotating vector describes an inward logarithmic spiral as shown in Fig. 12.12b for oscillations of decreasing amplitude. The relation of the circle in Fig. 12.12a to simple harmonic motion is similar to the relations of the spirals in Figs. 12.12b and 12.12c to damped harmonic motion and expanding harmonic motion, respectively.

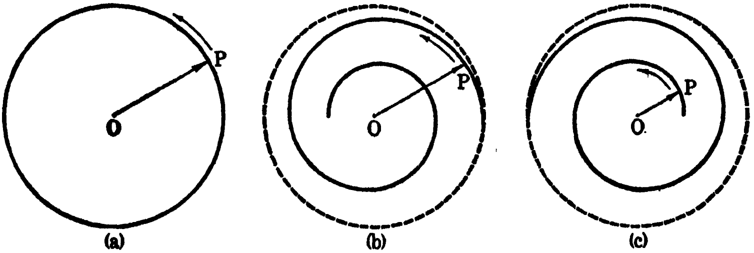


FIG. 12.12.—(*a*) Sustained or constant amplitude, (*b*) decreasing amplitude, and (*c*) increasing amplitude.

12.11. Time Constant.—Referring to (12–13) and (12–14), we see that, when $R^2C/4L > 1$, ω and ω_0 become imaginary, so that p can no longer be represented graphically by a rotating vector. We therefore drop the method of complex quantities and, instead, apply the method of first principles to the circuit of Fig. 12.9. We consider the case when $L = 0$. At time $t = 0$, let the charge on the condenser C be Q_0. The switch K is closed at $t = 0$ and the condenser begins to discharge through the resistance R. At $t = 0$, let the charge on the condenser be Q. The potential across the condenser is then Q/C. This potential is also applied to the resistance and from Ohm's law forces a current whose instantaneous value is

$$i = Q/RC \tag{12–19}$$

through it. But

$$i = -dQ/dt \tag{12–20}$$

where $-dQ$ is the decrease of charge on the condenser during a small interval of time dt. Hence we may write (12–19) in the form

$$dQ/Q = -dt/RC$$

Integrating both sides, we obtain

$$\log Q = -t/RC + \text{const} \tag{12–21}$$

Since $Q = Q_0$ at $t = 0$, const = log Q_0 and (12–21) may be written

$$\log (Q/Q_0) = -t/RC$$

or

$$Q = Q_0 e^{-t/RC} \tag{12–22}$$

The graph of Q against t is shown in Fig. 12.13. The value of Q drops to $1/e$ of the original value Q_0 when $t/RC = 1$, or when

$$t = RC \tag{12–23}$$

This value of t is called the *time constant* of a circuit consisting of a capacitance C and a resistance R.

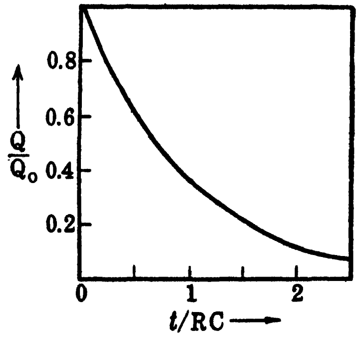


FIG. 12.13.—Discharge of a condenser through a resistance.

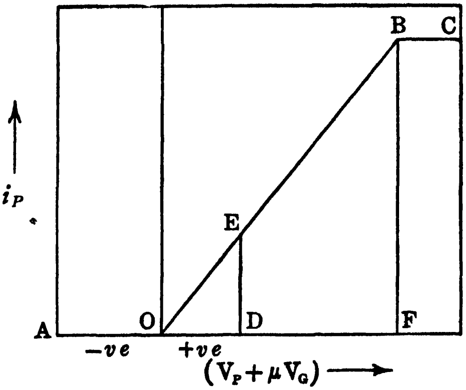


FIG. 12.14.—Idealized characteristic curve.

12.12. Idealized Oscillating Circuit.—The problem of radio circuits is simplified if we suppose that the relation

$$i_P = (V_P + \mu V_G)/R_P \tag{12–24}$$

is exact and that μ and R_P are constants irrespective of the values of V_P and V_G. We can then plot a combined plate and transfer characteristic such as shown in Fig. 12.14. A plate current represented by ED can be obtained by a combination of V_P and V_G such that $V_P + \mu V_G = OD$, provided that μV_G is negative. The horizontal section BC of the graph is obtained when $V_P + \mu V_G$ is great enough to produce the saturation value of i_P.

Consider the circuit shown in Fig. 12.15. The plate, grid, and heated cathode of the triode are represented by P, G, and K (originally "cathode" was spelled with a "k"). The heating circuit for K is not shown nor will it be shown in later diagrams. If there are no oscillations, $i_P = (E_B + \mu E_G)/(R + R_P)$, where R_P is the plate resistance of the tube. This is the value of the d.c. component of the plate current. For a while we shall forget this d.c. component and fix our attention on a possible a.c. component of the plate current. We shall represent this by the complex quantity a_P. We shall suppose that a_P consists of a_C through the capacitance C and a_R through the resistance R and inductance L_1 in series. From Sec. 4.12, the complex current a_R produces a complex e.m.f. $v_G = j\omega Ma_R$ in L_2 and this complex e.m.f. is applied to the grid circuit of the triode in Fig. 12.15. In order to simplify the following analysis, we shall assume that the potential E_C of the grid battery is such that the grid *bias* is always negative even though v_G is added to E_C. Consequently, no current flows through L_2 in the grid circuit, and the potential of the grid is v_G. Moreover, there is no induced e.m.f. in L_1. The effect of v_G on the complex plate current a_P is the same as if a generator giving a complex e.m.f. μv_G and having an internal resistance R_P were connected to an external circuit consisting of the capacitance C in parallel with the inductance L_1 and the resistance R as shown in Fig 12.15. The sum of the back e.m.f.'s around the

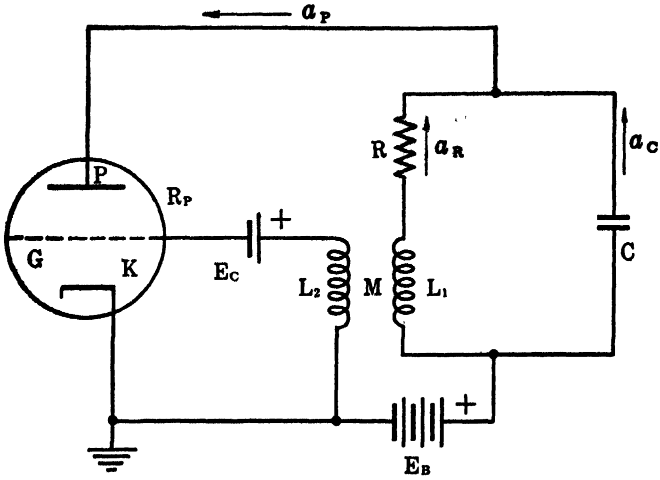


Fig. 12.15.—Tuned-plate oscillator.

circuit consisting of the triode and the capacitance C is given on the left side of (12–25).

$$R_P a_P - ja_C/\omega C = j\mu\omega M a_R \tag{12–25}$$

The generator voltage is shown on the right side. Next, the sum of the back e.m.f.'s around the circuit consisting of R, C, and L_1 is given on the left side of (12–26).

$$R a_R + ja_C/\omega C + j\omega L_1 a_R = 0 \tag{12–26}$$

Since there is no generator in this circuit, the right side of (12–26) is zero. Lastly, Kirchhoff's first law as applied to alternating currents states that the sum of the complex currents flowing away from (or into) a point in a circuit diagram is zero, so that

$$a_P - a_C - a_R = 0 \tag{12–27}$$

Eliminating a_P from (12–25) by means of (12–27) and rearranging (12–25) and (12–26), we obtain

$$(R_P - j\mu\omega M)a_R + (R_P - j/\omega C)a_C = 0 \tag{12–28}$$

and

$$(R + j\omega L_1)a_R + (j/\omega C)a_C = 0 \tag{12–29}$$

These are homogeneous linear equations. Such equations have a solution, other than the trivial solution $a_R = a_C = 0$, only when the solutions for a_R/a_C obtained from each equation are equal. Hence we must have

$$\frac{(R_P - j/\omega C)}{R_P - j\mu\omega M} = \frac{j/\omega C}{R + j\omega L_1} \tag{12–30}$$

or, after cancelling ωC and rearranging,

$$jL_1CR_P\omega^2 + (CRR_P + L_1 - \mu M)^{\omega} - j(R + R_P) = 0 \tag{12–31}$$

This is a quadratic in ω and has a solution of the form $\omega = \omega_0 + j\alpha$. The interpretation of a complex frequency has been discussed in Sec. 12.10. In the present case

$$\alpha = (RR_PC + L_1 - \mu M)/2L_1R_PC \tag{12–32}$$

Oscillations will build up if α is negative. The condition for this is that

$$\mu k\sqrt{L_1L_2} > RR_PC + L_1 \tag{12–33}$$

where $k\sqrt{L_1L_2}$ has been written for the mutal inductance M and k is the coupling coefficient for the inductances L_1 and L_2. The threshold condition is obtained from (12–33) by replacing the "greater than" sign with an "equals" sign. At the threshold

$$\omega_k^2 = (1 + R/R_P)/L_1C \qquad (12\text{–}34)$$

We shall call $\omega_k/2\pi$ the *oscillating frequency.*

The R, C, L_1 circuit of Fig. 12.15 is often called the *tank* circuit. It should be noted that any oscillating current which appears in the plate circuit of Fig. 12.15 causes an oscillating current in the tank circuit which in turn causes an oscillating voltage in L_2 which is applied to the grid. This voltage is called a *feedback* voltage. We recognize two types of feedback voltage—*positive* or *regenerative* feedback and *negative* or *degenerative* feedback. Whether we have positive or negative feedback depends on which terminal of the coil L_2 in Fig. 12.15 is connected to the grid. In (12–25) the right side has a positive or negative sign depending upon the sign of the feedback. If the sign is positive, we finally have a negative sign in front of μM in (12–32) and oscillations build up if (12–33) is satisfied. This is the case of positive feedback. On the other hand, if the sign is negative, we finally have a positive sign in front of μM in (12–32). In such a case α is always positive no matter what the values of L_1, L_2 and the coupling coefficient k may be. Any oscillation which starts for one reason or another is rapidly damped out. This is the case of negative feedback. When it is desired to prevent oscillations from building up, negative feedback is used.

12.13. Useful Circuit Elements.—The R, C, L_1 circuit in Fig. 12.15 is a part of the plate circuit and so may be considered as a circuit element of this circuit. This particular element is called a *parallel resonant* circuit element. By the methods of Chapter IV, the complex impedance of this circuit element is found to be

$$z = \frac{R + j\omega(L_1 - \omega^2L_1^2C - R^2C)}{(1 - \omega^2L_1C)^2 + \omega^2R^2C^2} \qquad (12.35)$$

For this to be entirely real, $\omega = \omega_r$, where ω_r is given by

$$L_1 - \omega_r^2L_1^2C - R^2C = 0 \qquad (12.36)$$

or

$$\omega_r = \sqrt{\frac{1 - R^2C/L_1}{L_1C}} \qquad (12.37)$$

The frequency $\omega_r/2\pi$ may be called the *resonant frequency*. For this frequency

$$z = Z = L_1/RC \tag{12-38}$$

where Z is the magnitude of z. In this case Z is entirely resistive, the reactance being zero.

From (12–38), we see that the resistance of the R, C, L_1 element of Fig. 12.15 increases with decrease of the resistance R. The case of $R = 0$ is shown in Fig. 12.16. When an oscillating voltage of the resonant frequency ω_r is applied at the points A and B, no current flows from A to B. At a frequency other than the resonant frequency current flows from A to B. Furthermore, if the voltage applied to A and B consists of a d.c. component and an a.c. component of the resonant frequency, the circuit element passes the d.c. component but blocks the a.c. component. Since the circuit element shown in Fig. 12.16 will not pass an alternating current of the resonant frequency, the element is sometimes called an *anti-resonant* circuit element.

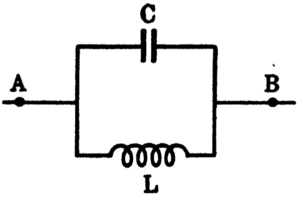


FIG. 12.16.—Parallel-resonant circuit element.

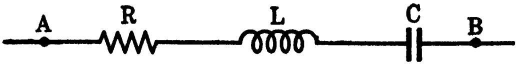


FIG. 12.17.—Series-resonant circuit element.

Consider the *series resonant* circuit element shown in Fig. 12.17. By the methods of Chapter IV, the current as read by an a.c. ammeter in the circuit is given by

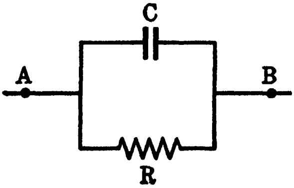
$$A = \frac{V}{\sqrt{R^2 + (L\omega - 1/C\omega)^2}} \tag{12-39}$$

This is a maximum when $\omega = \omega_n$, where

$$\omega_n = 1/\sqrt{LC} \tag{12-40}$$

We shall call the frequency $\omega_n/2\pi$ the *natural* frequency. If R is made very small, the current becomes very large at the natural frequency. For d.c. the frequency is zero and $1/C\omega = \infty$, so that $A = 0$. A series resonant circuit element will not pass a direct current but will pass an alternating current, most preferably if the latter has the natural frequency.

Another useful circuit element consists of a resistance and a capacitance in parallel as shown in Fig. 12.18. The r.m.s. value of an alternating current through this circuit element is



$$A = V\frac{\sqrt{1+\omega^2R^2C^2}}{R} \tag{12-41}$$

FIG. 12.18.—By-pass Condenser.

where V is the r.m.s. value of the a.c. voltage. For $\omega^2R^2C^2 \gg 1$, (12–41) gives $A = VC\omega$ and this current flows through the condenser. When a d.c. voltage E is applied to this circuit element, we have $I = E/R$ and this current passes through the resistance. In the case of a.c. and d.c. components, the condenser passes the a.c. component and the resistance the d.c. component. The condenser in Fig. 12.18 is called a *by-pass* condenser.

In radio circuits we recognize two orders of magnitude of frequency. Audio-frequencies are in the range of about 50 to 5000 cycles and broadcast radio-frequencies in the range 550 to 1500 kc (1 kc = 1000 cycles). An inductance has a reactance $L\omega$. A coil with a large value of $L\omega$ for a frequency $\omega/2\pi$ is called a *choke*. For audio-frequencies ω is relatively small and L must be large to give the choking effect. An audio-frequency choke therefore has an iron core. For radio-frequencies ω is relatively large and L may be relatively small. A radio-frequency choke therefore has an air core.

For $R/R_P \ll 1$ in (12–34) and $R^2C/L_1 \ll 1$ in (12–37), we have $\omega_k = \omega_r = \omega_n$ very nearly. The approximate equality of the oscillating, resonant, and natural frequencies is often very useful.

12.14. Q of a Circuit.—Eliminating a_C from (12–27) and (12–29), we obtain

$$\frac{a_P}{a_R} = 1 - \omega^2L_1C + j\omega RC \tag{12-42}$$

At the threshold we have the oscillating frequency given by (12–34). Putting $\omega = \omega_k$ in (12–42), we have for the r.m.s. values of the currents

$$\frac{A_P}{A_R} = \frac{R}{R_P}\sqrt{1 + \frac{R_P{}^2C}{L_1}\left(1 + \frac{R}{R_P}\right)} \tag{12-43}$$

We see that, when the resistance R of the tank circuit is small with respect to the plate resistance R_P, the ratio A_P/A_R is small. Consequently an oscillating current of large amplitude can be built up in the tank circuit by an oscillating current of small amplitude in the plate circuit.

The oscillating current in the tank circuit is similar to the oscillating movement of a pendulum. A frictionless pendulum, when once set going, would oscillate indefinitely. But no pendulum is frictionless and a small amount of energy must be supplied to the pendulum to keep it oscillating. Similarly, if $R = 0$ in the tank circuit of Fig. 12.15, oscillations, once set going, would continue indefinitely. In such a case no energy would need to be supplied by the plate circuit. However, $R \neq 0$, so a small amount of energy must be supplied by the plate circuit. Hence, a ratio describing the smallness of the friction or resistance of an oscillating circuit or circuit element has been devised. This ratio is known as the *quality* or Q of the circuit or circuit element and is defined by

$$Q = L\omega/R \tag{12–44}$$

where L is the inductance and R the resistance of the circuit. There is some uncertainty as to the value of ω. It might be either ω_k, ω_r or ω_n. However, these are usually nearly equal, so that we shall take $\omega/2\pi$ as being the natural frequency of the circuit or circuit element.

From (12–44), it is seen that Q is large when R is small. The greater Q the smaller the energy needed to keep oscillations of a certain amplitude going. For ordinary commercial radio, Q is in the range 50 to 300. However, in recent work at frequencies above those of commercial radio, values of Q in the thousands are known. As defined by (12–44), Q is a pure number without a unit.

If the tank circuit of Fig. 12.15 is suddenly disconnected from the triode and driving battery E_B after sustained oscillations have been excited, it becomes similar to the circuit of Fig. 12.9 when the switch K is closed. From (12–18) we note that the amplitude of the oscillating current decreases with the factor $e^{-\alpha t}$, where, from (12–13) and (12–14) $\alpha = R/2L_1$. If t is the period $2\pi/\omega$ of the oscillations in the tank circuit, $\alpha t = \pi/Q$, so that $e^{-\pi/Q}$ is the ratio of one maximum to the previous maximum of current. The number n of to and fro unsustained oscillations for a maximum of current to be reduced to $1/e$ of the initial maximum is given by $n\pi/Q = 1$, whence $n = Q/\pi$. From Fig. 12.12b we

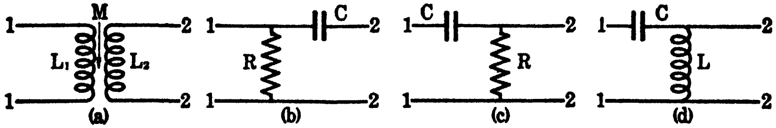


FIG. 12.19.—Different kinds of coupling.

note that Q is the number of radians turned through for the rotating vector to be reduced to $1/e$ of a given value.

12.15. Coupled Circuits.—It is often desirable to transfer electrical energy from one circuit to another. We have discussed the transfer of electrical energy from the primary to the secondary circuit of a transformer in Sec. 4.12. The two circuits are said to be *coupled* and the coupling is described as *transformer coupling*. This is shown in Fig. 12.19a. *Direct*, *resistance-capacitance*, and *impedance-capacitance* coupling are shown in Figs. 12.19b, c, and d, respectively. The primary and secondary terminals are indicated by 1 and 2, respectively. In coupling two triodes the plate of one triode is sometimes connected through a condenser to the grid of the other triode. This is properly called "resistance-capacitance coupling" but is frequently referred to simply as "resistance coupling."

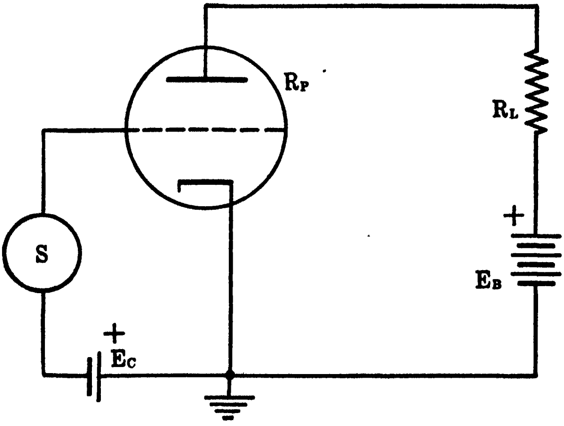


FIG. 12.20.

12.16. Triode Amplifiers.—These are of two kinds: voltage amplifiers and power amplifiers. Each kind is divided into three classes: *A*, *B*, and *C*. In voltage amplification an alternating potential is connected in series with the biasing battery in the grid circuit. This alternating potential is known as the *signal* and is represented by S in Fig. 12.20. It produces an alternating or pulsating current in the plate circuit. This state of affairs is represented graphically in Fig. 12.21. The line OX is the axis of time, OY is the axis of instantaneous plate current, OY^1 is a second axis of time and OX^1 is the axis of instantaneous grid potential, the negative values of which are measured to the left of O. The e.m.f. of the biasing battery is represented by OB. Curve I is the graph of the signal voltage versus time. The point P on Curve I indicates that at time t the total voltage on the grid is represented by OD. The

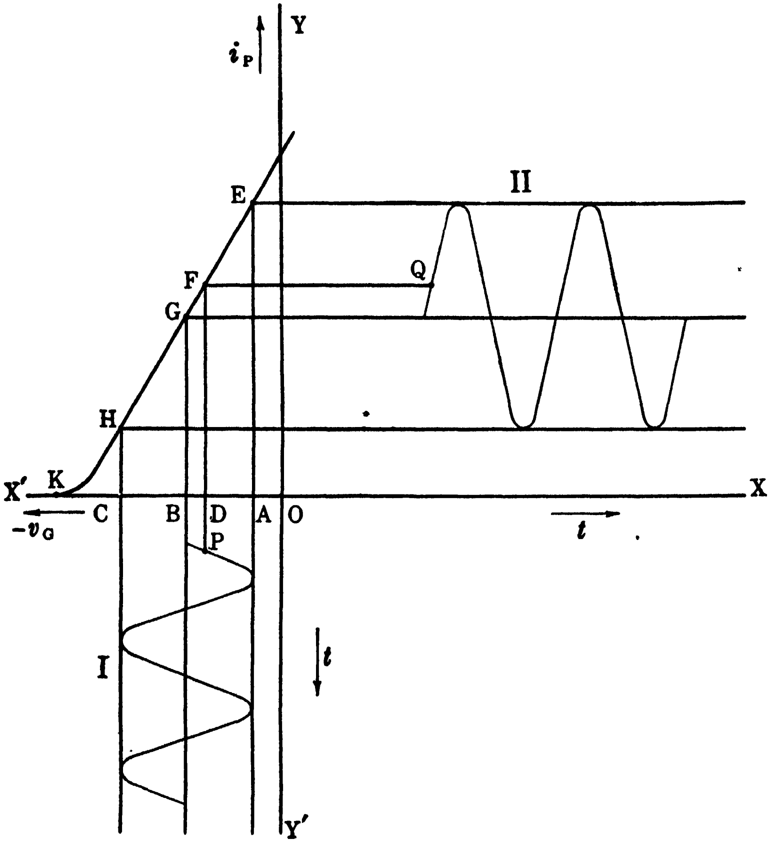


FIG. 12.21.

curve *EFGHK* is the transfer characteristic (plate current versus grid volts) for a certain plate potential of the triode. At the instant t the point on the characteristic corresponding to v_G is F. This in turn gives rise to the point Q on Curve II, which is the graph for the plate current versus time. As the point P moves along Curve I, the point Q moves along Curve II. If the transfer characteristic is a straight line between H and E, Curve II has the same shape as Curve I. For instance, if Curve I is a sine curve, Curve II will be a sine curve. Usually, however, the operating portion of the transfer characteristic departs from a straight line, so that the shape of Curve II departs from the shape of Curve I and we have *distortion.*

When plate current flows at every instant during a cycle of the signal voltage, we have *Class A* amplification. This is the case shown in Fig. 12.21. *Class B* amplification occurs when the plate current flows for approximately half a cycle of the signal voltage. This case is shown in Fig. 12.22a. The biasing e.m.f. of the battery is approximately equal to the cutoff OK. *Class C* amplification occurs when the plate current flows for a small fraction (much less than one half) of a cycle of the signal voltage.

The amplitude of the signal voltage may be so great that for a part of a cycle the grid is driven positive and current flows from the grid to the cathode of the triode. This is indicated by adding the subscript 2 to the letter indicating the class of amplification. Thus, Fig. 12.22b represents Class C_2 amplification. To indicate the case of no current from the grid to the cathode the subscript 1 is added to the letter indicating the class of amplification. Thus, Fig. 12.22a represents Class B_1 amplification.

A grid biasing battery is shown at E_C in Fig. 12.20. However, it is more usual to obtain grid bias by means of a resistance and by-pass condenser as shown in Fig. 12.23. For simplicity we shall limit ourselves to Class A_1 amplification. Before the signal voltage is applied to the grid, the plate current I_P flows through R_L, R_P, and R_C in series. It flows downward through R_C in the diagram, so that the cathode is at a positive potential R_CI_P with respect to ground. However, if $S = 0$ in Fig. 12.23, the grid is at ground potential and so has a negative bias of R_CI_P with respect to the cathode. If the transfer characteristic is a straight line and the amplification factor is μ, we have

$$I_P = \frac{E_B - \mu R_C I_P}{R_L + R_P + R_C} \tag{12-45}$$

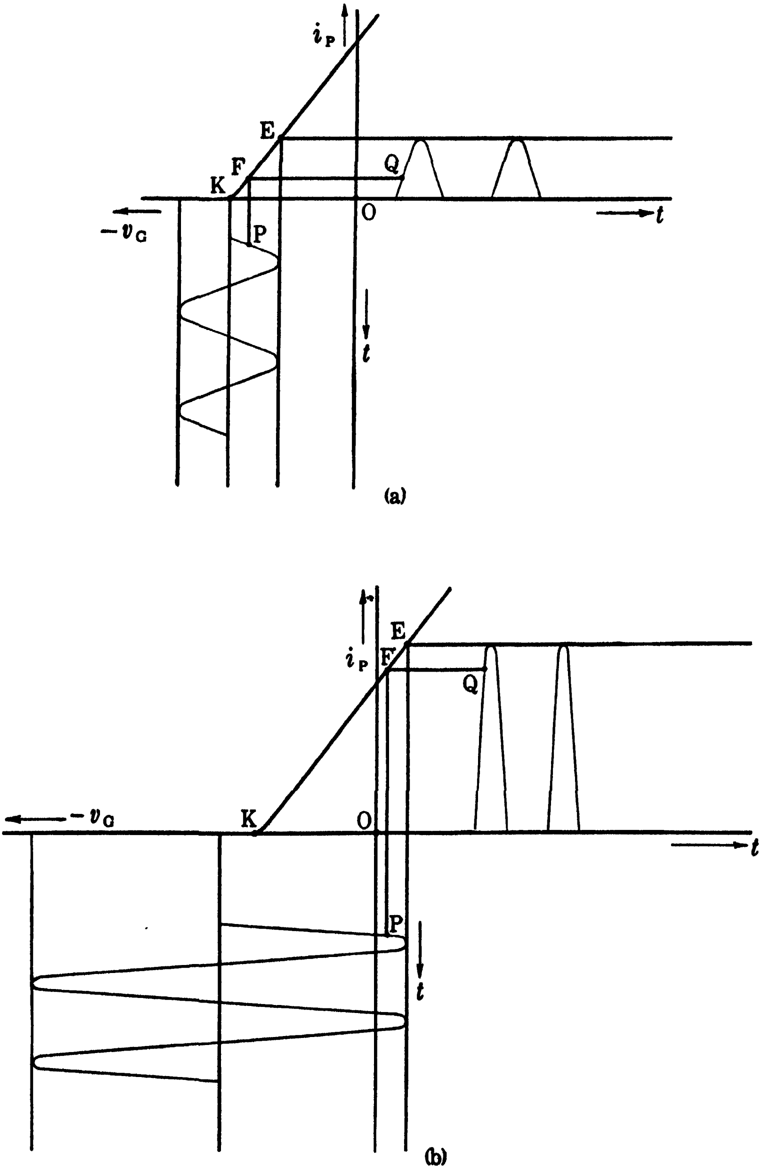


FIG. 12.22.

whence

$$\text{Grid bias} = -R_C I_P = -\frac{R_C E_B}{R_L + R_P + (\mu + 1)R_C} \tag{12–46}$$

If now the signal voltage is applied, we have an a.c. component of the plate current as shown in Fig. 12.21. This a.c. component passes through the condenser C in Fig. 12.23 and so by-passes the resistance if the proper value of C is chosen.

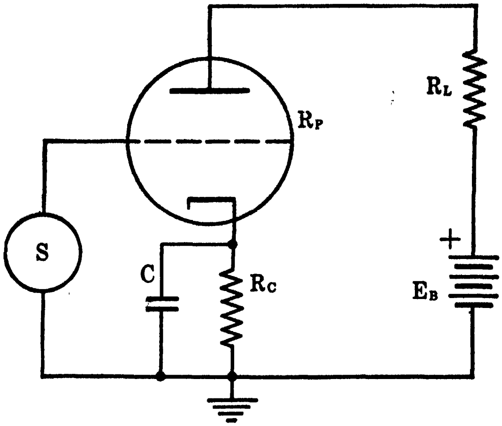


Fig. 12.23.

The a.c. component of the voltage across R_L in Figs. 12.20 and 12.23 is known as the output voltage. The ratio of the amplitude of the output voltage to the amplitude of the signal or input voltage is called *voltage amplification*. In Fig. 12.20 the a.c. component of plate current gives rise to an a.c. component of voltage across R of amplitude $\mu E_S R_L/(R_L + R_P)$, where E_S is the amplitude of the signal voltage. Hence

$$\text{Voltage amplification} = \mu/(1 + R_P/R_L) \tag{12–47}$$

and approaches the amplification factor μ when R_L is much greater than R_P.

In two stage amplification the amplified voltage is applied to the grid of a second triode. This case is shown in Fig. 12.24. Since there is a d.c. component of voltage across R in the first stage, it is necessary to protect the grid of the second tube from this voltage. This is done by means of the *blocking condenser* C_1. This passes this a.c. component but blocks the d.c. component of the output voltage of the first stage. In Fig. 12.24 we have resistance-capacitance coupling between the two stages.

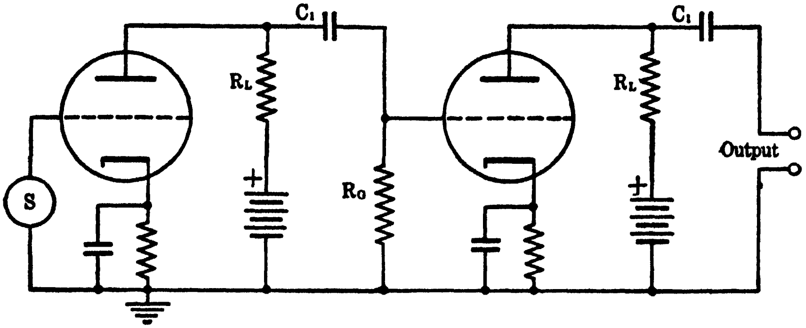


FIG. 12.24.—Two-stage voltage amplification.

In power amplification we consider the average a.c. power dissipated in the *load* resistance R_L of Fig. 12.20. This is R_LA^2 where A is the r.m.s. value of the a.c. component of the plate current. But $A = (\mu E_S/\sqrt{2})/(R_L + R_P)$, where E_S is the amplitude of the signal voltage. Hence the power dissipated is given by

$$P = \frac{\mu^2 E_S{}^2 R_L}{2(R_L + R_P)^2} \tag{12-48}$$

We consider the case when E_S is constant. The power P is a maximum when $dP/dR_L = 0$. After differentiation, it is found that for this condition

$$R_L = R_P \tag{12-49}$$

Hence, with a constant E_S, the maximum output of power is obtained when the load resistance *matches* the plate resistance. The power output in this ideal case is

$$P = \mu^2 E_S{}^2/8R_P \tag{12-50}$$

For a voltage amplifier R_P/R is made as small as feasible while for a power amplifier R_P/R is made as near unity as feasible. In multistage power amplification it is usual for all stages before the last to be voltage amplifiers and the last stage to be a power amplifier.

12.17. Practical Oscillating Circuit.—It would seem that the amplitude of the a.c. component of the plate current in Fig. 12.15 would build up indefinitely. However, the conventional plate current can only flow from the plate to the cathode (the electron current flows from the cathode to the plate). The resistance of a triode to a current flowing from the plate to the cathode is R_P, whereas the resistance when the plate is negative with respect to the cathode is infinite. This is described as the *valve action* of the triode. The circuit of Fig. 12.15 is idealized. Actually the grid battery E_G is replaced by a condenser and a resistance in parallel as shown in Fig. 12.25. The purpose is to give a *grid-leak* negative bias to the grid. When a switch in the plate circuit of Fig. 12.15 is closed the potential of the grid is zero and the initial value of the plate current is $E_B/(R_P + R)$. Oscillations of potential thereafter appear in the grid and plate circuits if the condition (12–33) is fulfilled. For there to be oscillations the alternating potential between the grid and the cathode must be 180° out of phase with that between the plate and the cathode.

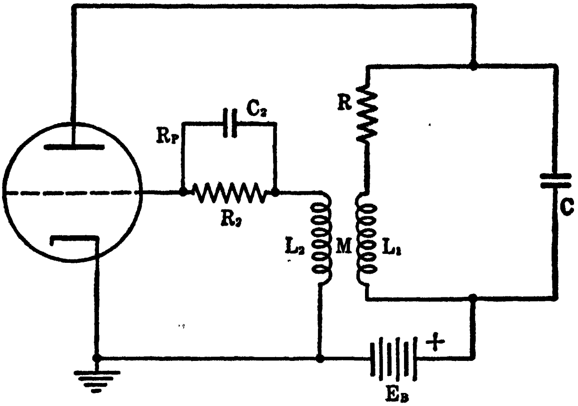


FIG. 12.25.—Grid-leak negative bias.

Let e_P and e_G be the instantaneous values of the a.c. components of the plate and grid potentials at a time t after the switch is closed. The variations of e_P and e_G with t are shown approximately in Fig. 12.26. We see that the grid is driven positive for half of each cycle. Current flows through R_2 of Fig. 12.25 for alternate half cycles and we have a series of pulses of current always in the same direction but increasing in magnitude with the time. A voltage drop occurs across R_2 and this potential difference charges the condenser C_2. The condenser discharges through R_2 during those alternate half cycles when no current flows from the grid to the cathode. First the grid current through R_2 and then the discharge of the condenser through R_2 keep the grid at a negative potential. This negative bias increases as the magnitude of the pulses increases. The values of R_2 and C_2 are so chosen as to make the time constant R_2C_2 long compared with the period of a cycle At a certain stage this bias equals cutoff and the triode is operating as a Class B_2 amplifier. But the pulses increase still further in magnitude until the triode operates as a Class C_2 amplifier. However there must be some plate current in order to maintain the oscillations in the tank circuit since, although Q may be large for the tank circuit, it is not infinity. Finally a state of equilibrium comes about where the bias of the grid is sufficiently negative to give pulses of plate current of just

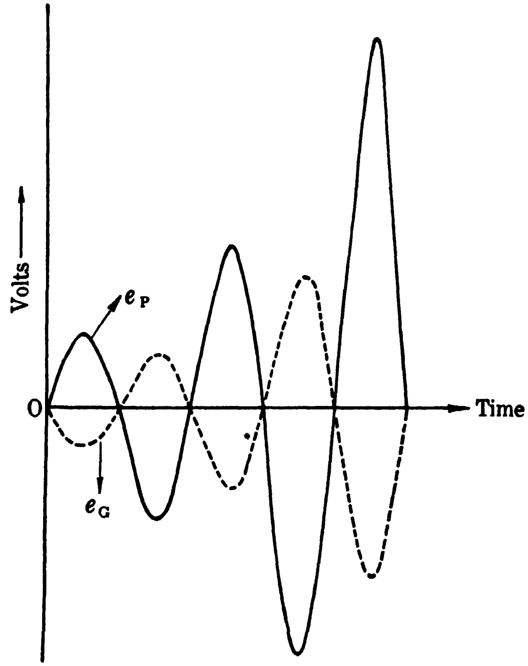


FIG. 12.26.—Variations of e_P and e_G with time during the building up of oscillations.

sufficient length of time to produce *sustained* oscillations in the tank circuit of constant amplitude.

The state of sustained oscillations is reached in a few microseconds (1 microsecond = 1×10^{-6} sec) in a tank circuit whose frequency is in the range 550 to 1500 kc (1 kc = 1 kilocycle = 1000 cycles), which is the range of commercial broadcast frequencies. When the condition of equilibrium has been attained, the state of affairs is shown by Figs. 12.27 and 12.28.

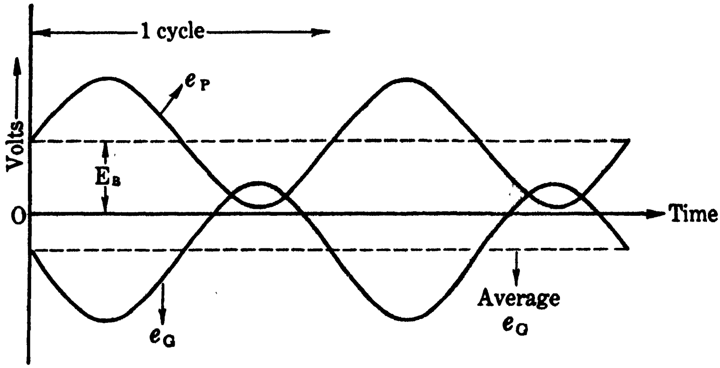


FIG. 12.27.—Condition of sustained oscillations. (The scales of e_P and e_G are different in the diagram.)

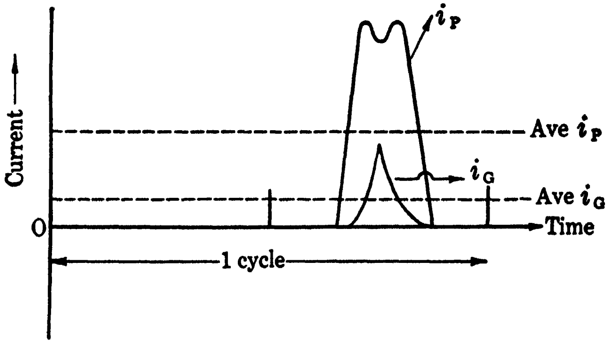


FIG. 12.28.—Condition when the oscillations are sustained. (The scales of i_P and i_G are not the same. The averages are taken over a whole cycle and are not drawn to scale. Each current flows for less than one-half cycle.)

12.18. Modulation.—The amplitude of the sustained oscillations in the tank circuit increases with the average plate current. Since the average plate current can be increased by increasing E_B, the amplitude of the sustained oscillations can be controlled by the e.m.f. of the plate battery. It should be remembered that the energy to sustain the oscillations of the tank circuit comes from the plate battery.

Consider Fig. 12.29. On the extreme left of the diagram is the microphone F which is connected to the primary of an iron core transformer and a battery. The air vibrations hitting the diaphragm of the microphone produce vibrations of the diaphragm and these in turn change the pressure between the carbon granules in the microphone. The electrical resistance in the microphone circuit changes and so if a certain voice- or audio-frequency enters the microphone an a.c. component of current of this frequency flows through the primary of the transformer T. This a.c. component gives rise to an audio-frequency a.c. component of potential on the grid of triode T_1, which acts as a power amplifier and is known as the modulator tube. The resulting a.c. component of plate current of T_1 is prevented from passing through the battery E_B by the audio-frequency choke L_4, across which is developed an audio-frequency a.c. component of voltage. This a.c. voltage is passed through the radio-frequency choke L_3 to the plate of triode T_2. The effect is the same as if the plate battery e.m.f. E_B were varied slowly—that is, slowly in terms of the period of the sustained oscillations of the tank circuit. Consequently, the amplitude of the sustained oscillations of the tank circuit varies slowly—that is, with audio-frequency $\omega_a/2\pi$. The oscillations in the tank circuit are represented by Fig. 12.30. The oscillations are said to be *modulated* and the process

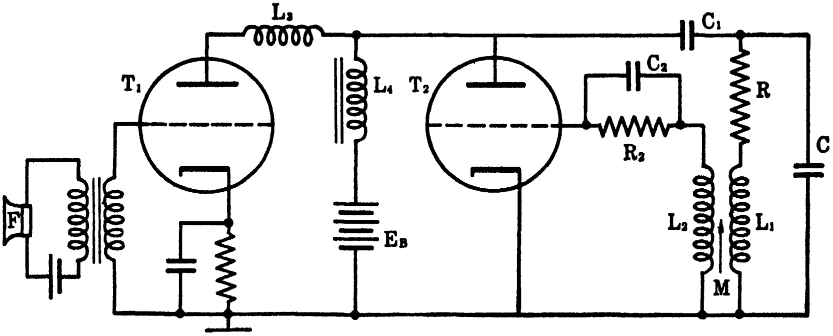


FIG. 12.29.—Circuit for modulation.

is known as modulation. The broken curves in Fig. 12.30 constitute the envelope of the radio-frequency oscillations.

The purpose of the radio-frequency choke L_3 is to prevent the radio-frequency oscillations in the plate circuit of triode T_2 from reaching the plate of triode T_1. The condenser C_1 prevents the flow of direct current to the tank circuit and also if its capacitance is small shields the tank circuit from the audio-frequency voltage supplied by the modulator triode T_1 to the plate of the oscillator triode T_2. Moreover, an analysis similar to that of Sec. 12.12 shows that oscillations cannot be built up in the tank circuit unless $L_3 > L_1$ and that the frequency of these oscillations only approaches $1/(2\pi\sqrt{L_1C})$ when $L_3 \gg L_1$.

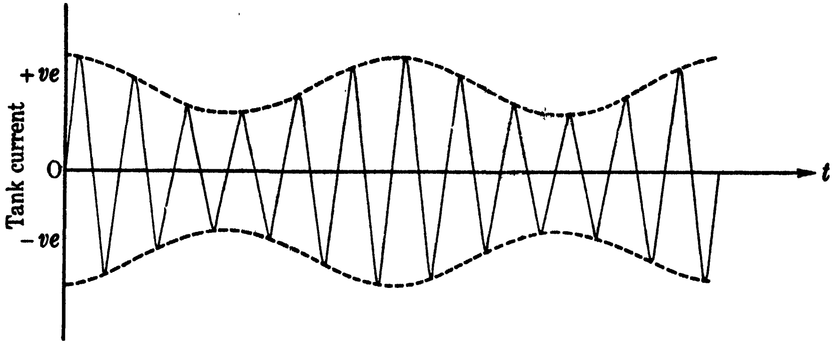


FIG. 12.30.

12.19. Side Frequencies and Bands.—If a musical note of a single frequency enters the microphone F of Fig. 12.29, the upper and lower parts of the envelope in Fig. 12.30 are sine curves. The amplitude A of the radio-frequency oscillations of the tank circuit then varies according to

$$A = K(1 + m \sin \omega_m t) \tag{12-51}$$

where $\omega_m/2\pi$ is the frequency of the musical note (the modulation frequency) and K and m are constants. The constant m is called the modulation factor and is usually measured in per cent. It is never greater than 100 per cent. From (12-51) the instantaneous value of the alternating current in the tank circuit is then

$$i = K(1 + m \sin \omega_m t) \sin \omega_k t \tag{12-52}$$

where $\omega_k/2\pi$ is the oscillation frequency in the tank circuit when there is no modulation. From trigonometry we may write (12–52) in the form

$$i = K\sin\omega_k t - (1/2)Km[\cos(\omega_k + \omega_m)t - \cos(\omega_k - \omega_m)t] \quad (12\text{–}53)$$

The oscillating current in the tank circuit consists of three components whose frequencies are $\omega_k/2\pi$, $(\omega_k + \omega_m)/2\pi$ and $(\omega_k - \omega_m)/2\pi$. The first frequency is the oscillating frequency of the tank circuit; the other two frequencies are known as *side frequencies.* We shall from now on call $\omega_k/2\pi$ the *carrier* frequency. When the carrier frequency is 1000 kc and the modulation frequency is 2 kc, the two side frequencies are 998 and 1002 kc. The voice contains various frequencies, so that, instead of there being two side frequencies, there may be several such, and these are known collectively as *side band* frequencies.

The type of modulation we have discussed in this chapter is called amplitude modulation. Frequency and phase modulation are two other types of modulation. These are beyond the scope of this book and will not be discussed.

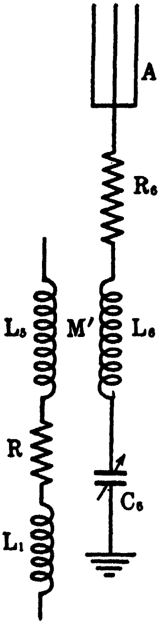


FIG. 12.31.

12.20. A Radio Transmitter.—We replace the R, L_1 arm of the tank circuit of Fig. 12.15 by the arm containing L_1, R, and L_5 as shown in Fig. 12.31. The inductance L_5 together with the inductance L_6 of the antenna circuit forms a transformer with mutual inductance M'. Under proper conditions standing or stationary waves of current and potential are set up in the antenna A of Fig. 12.31. The transformer coupling of the antenna circuit to the tank circuits changes somewhat the threshold condition for oscillations in the tank circuit and the frequency of these oscillations. We shall call this somewhat changed frequency the *carrier* frequency.

Under proper conditions standing or stationary waves of current and potential are set up in the antenna. The discussion of the radiation of electric waves from an antenna is beyond the scope of this book. Suffice it to say that the frequency of these waves is that of the carrier frequency and that the amplitude of the electric waves is proportional to the amplitude of the oscillations in the tank circuit. The electric field above the earth's surface at a point some distance from the antenna varies with the time in a manner similar to that shown in Fig. 12.30.

12.21. A Radio Receiver.—Let us first consider the receiving circuit shown in Fig. 12.32. Electric waves falling upon the antenna produce radio frequency oscillations of current through L_1. This coil is transformer coupled to L_2. By means of the variable capacitance C_2 the L_2, C_2 circuit is tuned to the carrier frequency of the desired transmitting station. The alternating potential across C_2 is imposed on the grid of the triode together with the negative bias produced by the self-biasing combination of C_3 and R_3 between the cathode and the negative terminal of the plate battery. A triode with a very nearly straight or linear transfer characteristic is used. If the negative grid bias is at cutoff, we have the state of affairs represented in Fig. 12.33 and we have Class B amplification.

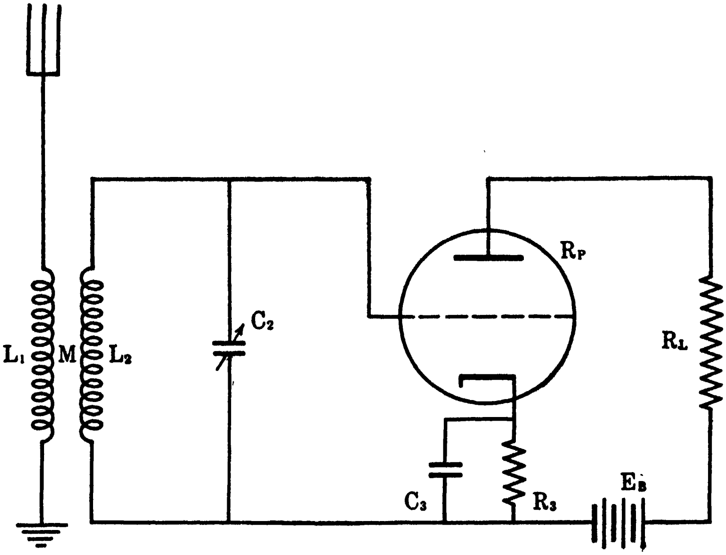


Fig. 12.32.

12.22. Harmonics.—There is a very useful theorem in mathematics known as *Fourier's theorem*. Application of this theorem to a problem is known as *Fourier analysis*. For our purposes the theorem states that when a function of the time, such as $f(t)$, repeats itself in a period of time $2\pi/\omega_k$, $f(t)$ may be expressed as a series of sines, thus

$$f(t) = a_0 + a_1 \sin (\omega_k t + \phi_1) + a_2 \sin (2\omega_k t + \phi_2) + a_3 \sin (3\omega_k t + \phi_3) + \cdots \qquad (12\text{–}54)$$

where, in alternating current theory, a_0 is the d.c. component, and a_1, a_2, a_3, etc., are the amplitudes of the a.c. components of frequencies $\omega_k/2\pi$, $2\omega_k/2\pi$, $3\omega_k/2\pi$, etc., respectively, and the ϕ's are the corresponding phases of these different a.c. components at $t = 0$. The a.c. components whose frequencies are $\omega_k/2\pi$, $2\omega_k/2\pi$, $3\omega_k/2\pi$, etc., are known as the 1st, 2nd, 3rd, etc., harmonics, respectively. In Fig. 12.33 it is seen that, if the transfer characteristic is a straight line with a sharp cutoff and there is no modulation, the plate current is a series of pulses, each having the shape of half a sine curve. We can then say that the instan-

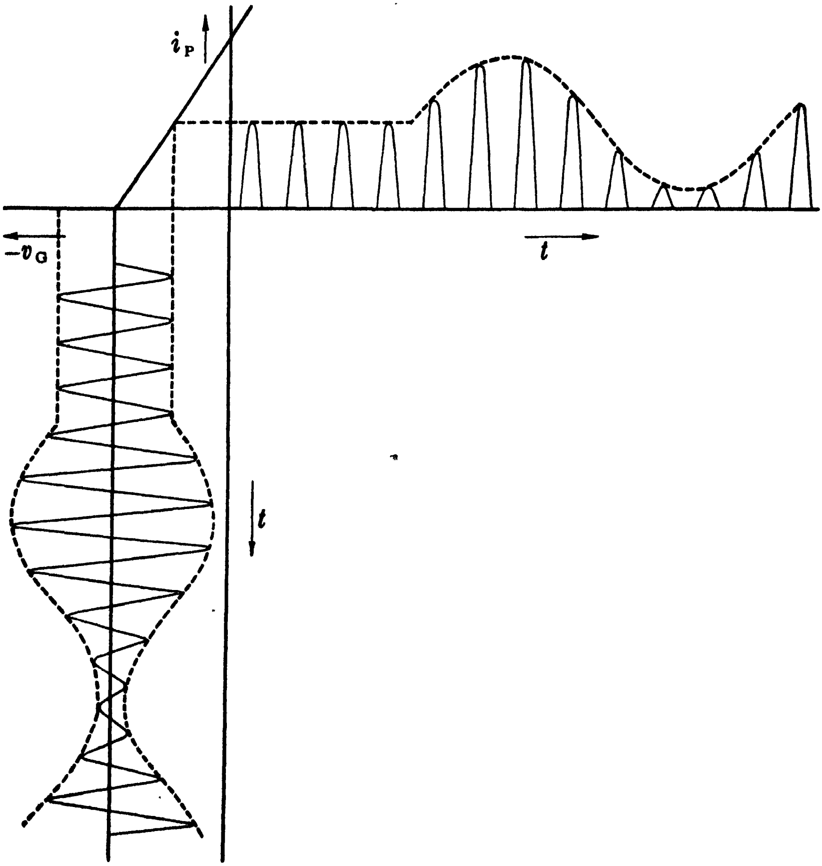


FIG. 12.33.—Process of demodulation.

taneous value of the plate current is $I_0 f(t)$, where I_0 is the maximum current of the pulse, $f(t) = \sin \omega_k t$, for the first half of each cycle, and $f(t) = 0$ for the second half of each cycle. In such a case, Fourier analysis shows that

$$f(t) = \frac{1}{\pi} + \frac{\sin \omega_k t}{2} + \frac{2 \sin (2\omega_k t - \pi/2)}{1 \times 3\pi} + \frac{2 \sin (4\omega_k t - \pi/2)}{3 \times 5\pi} + \frac{2 \sin (6\omega_k t - \pi/2)}{5 \times 7\pi} + \cdots \quad (12\text{–}55)$$

We note that the pulses consist of a d.c. component I_0/π, a first harmonic or fundamental of amplitude $I_0/2$, a second harmonic of amplitude $2I_0/3\pi$, a fourth harmonic of amplitude $2I_0/(3 \times 5\pi)$ and so on for the even harmonics. There are no odd harmonics beyond the first. However, the even harmonics all have a phase lag of $\pi/2$ at $t = 0$.

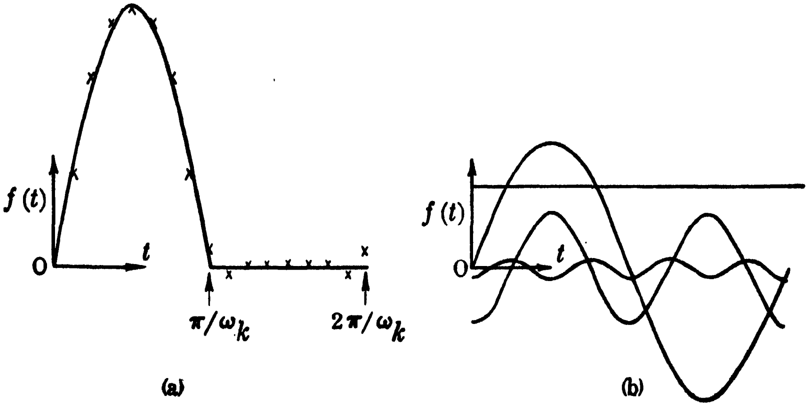


FIG. 12.34.—Solid curve in (*a*) shows $f(t)$ versus t. Curves in (*b*) show d.c. component and 1st, 2nd, 4th harmonics. The crosses in (*a*) show the sum of the components in (*b*) at steps of $t = \pi/(8\omega_k)$

That (12–55) represents $f(t)$ is shown in Fig. 12.34, where the sum of the d.c. component and the 1st, 2nd and 4th harmonics is seen to give the original $f(t)$ very closely. With the inclusion of still higher harmonics a closer approximation is obtained.

In alternating current theory, inductances, capacitances, and ohmic resistances behave towards each harmonic and the d.c. component (which may be called the zero harmonic) as though all the other harmonics were absent. Hence the actual current flowing through one of these circuit elements equals the sum of all the various harmonic currents. This is called the *principle of superposition*. For R_3 in Fig. 12.32 to develop a negative grid bias it is necessary for a d.c. component to be present in the plate current. The grid bias is then R_3I_0/π in the case for which (12–55) applies. The d.c. component is the average of the pulse current over a cycle as can be seen from the application of (2–63) to $I_0f(t)$, thus

$$\text{Ave}\, I_0f(t) = \frac{\omega_k}{2\pi}\int_0^{\pi/\omega_k} I_0 \sin \omega_k t\, dt = I_0/\pi \qquad (12\text{–}56)$$

If the modulation begins at $t = 0$ in (12–51) and (12–52), the plate current pulses of varying magnitude also begin at $t = 0$. Next, let us suppose that $\omega_k/2\pi = 1000$ kc and $\omega_m/2\pi = 1$ kc, the heights of the pulses will then increase as t increases from 0 to $(1/4) \cdot 2\pi/\omega_m$. During this interval there will be 250 pulses. If the modulation factor is m, the height of the 250th pulse will be $(1 + m)I_0$ and the average current during the radio-frequency cycle containing the 250th pulse will be $(1 + m)I_0/\pi$. The average current during successive radio-frequency cycles thus varies with the time according to $(I_0/\pi) \cdot (1 + m \sin \omega_m t)$. This may be considered as made up of a steady component I_0/π and an audio-frequency component $(mI_0/\pi) \sin \omega_m t$. The pulses of varying magnitude thus give rise to an audio-frequency component on the right side of (12–55). This process of regaining the audio-frequency component is known as *demodulation* or *detection*. Fourier analysis of the varying pulses also gives an audio-frequency component together with radio-frequency components. Consequently the experimental problem is to *filter out* the radio-frequency components. This may be done by means of the circuit shown in Fig. 12.35. The coil L_1 is a radio-frequency choke and C_1 is a radio-frequency by-pass condenser. The choke L_1 allows the audio-frequency component of the plate current to pass through to the transformer T, in the secondary circuit of which is developed an audio-frequency alternating potential. This may be applied to a Class A amplifier, which in turn operates a power amplifier and a loudspeaker. It is to be noted that the by-pass condenser C_2 must pass the audio-frequency component of the plate current. The d.c. component of the plate current passes through the primary coil of T, the coil L_1 and the resistance R_2.

12.23. The Resonant Transformer.—We must now consider *selectivity* and *fidelity* of reception. If the resistance in the L-C circuit of Fig.

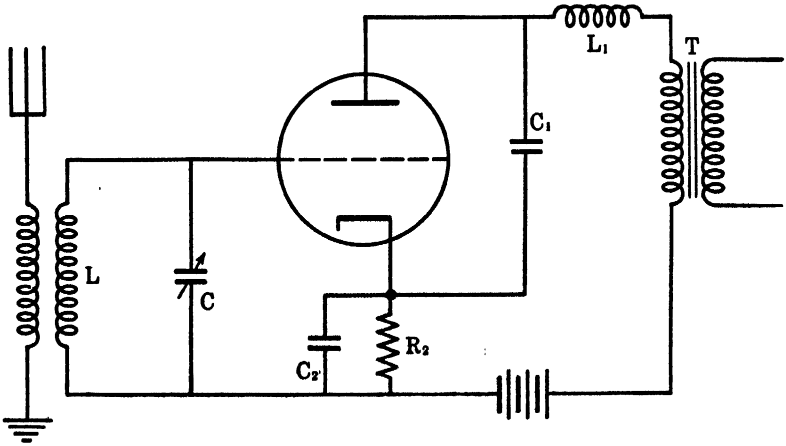


FIG. 12.35.

12.35 is very small, the tuning is very sharp. However, the electric waves falling on the antenna contain according to our description the radio frequencies $(\omega_k - \omega_m)/2\pi$, $\omega_k/2\pi$, and $(\omega_k + \omega_m)/2\pi$. If the receiving set is tuned to accept the carrier frequency $\omega_k/2\pi$ and if the tuning is very sharp, the side frequencies may not be accepted. Furthermore, in the case of side bands of frequencies, one frequency may not produce the same *response* as another. This produces *distortion* so that the vibrations of the diaphragm of the loudspeaker of the receiver do not correspond to the vibrations of the diaphragm of the microphone of the transmitter. This causes lack of fidelity. In order to give high fidelity, the receiver must accept equally well electric waves over a range of some 5 to 10 kc centering about the carrier frequency. On the other hand, in the United States, the different broadcasting stations are allotted carrier frequencies over a range of 550 to 1500 kc in steps of 10 kc. If the fidelity is too high, electric waves from a station other than the one desired may cause a response which is not negligible. So a fair degree of selectivity is desired. A compromise is necessary. We now describe a coupling and a tuning device which brings this about.

Consider the circuit shown in Fig. 12.36. The mutual inductance is shown as M. The resistance, inductance, and capacitance of the primary circuit are indicated by the subscript 1, while those of the secondary circuit are indicated by the subscript 2. A complex voltage v_1 is supplied by a generator to the primary circuit. This produces the complex currents a_1 and a_2 as shown. Following the method of Sec. 4.12, we have, for the primary circuit,

$$(R_1 - j/\omega C_1 + j\omega L_1)a_1 + j\omega M a_2 = v_1 \tag{12-57}$$

where $\omega/2\pi$ is the frequency of v_1, while, for the secondary circuit,

$$j\omega M a_1 + (R_2 - j/\omega C_2 + j\omega L_2)a_2 = 0 \tag{12-58}$$

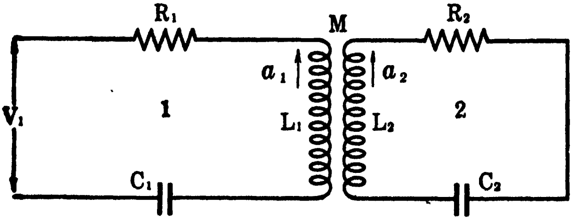


FIG. 12.36.—The resonant transformer.

We can solve these two equations for a_1 and a_2. However, the algebra is cumbersome. We therefore simplify the problem by making $R_1 = R_2 = R$, $L_1 = L_2 = L$, and $C_1 = C_2 = C$. We now obtain

$$a_2 = \frac{v_1}{R' + jX'} \tag{12-59}$$

where

$$R' = \frac{2R(1 - \omega^2 LC)}{\omega^2 MC} \tag{12-60}$$

and

$$X' = \frac{\omega^4 C^2(M^2 - L^2) + \omega^2(2LC - R^2C^2) - 1}{\omega^2 MC} \tag{12-61}$$

The ammeter reading in the secondary circuit is

$$A_2 = \frac{V_1}{\sqrt{R'^2 + X'^2}} \tag{12-62}$$

where V_1 is the open circuit voltage of the generator as read on a voltmeter. We seek the condition under which A_2 is a maximum or a minimum. We have defined the natural frequency $\omega_n/2\pi$ of an L, C circuit in Sec. 12.13. However, we shall assume that $\omega/2\pi$, the frequency of v_1, although nearly equal to $\omega_n/2\pi$, is not quite equal to it. Hence, we write $\omega_n{}^2 - \omega^2 = (\omega_n + \omega)(\omega_n - \omega) = 2\omega_n\omega'$ approximately, where $\omega' = \omega_n - \omega$. We now make use of Q, which we defined in Sec. 12.14, and also of k, the coupling coefficient of a transformer. In the present case, since $L_1 = L_2 = L$, we have $M = kL$. We can therefore write (12–62) in the form

$$A_2 = \frac{kQV_1}{R\sqrt{(1 + k^2Q^2)^2 + x^2(x^2 + 2 - 2k^2Q^2)}} \qquad (12\text{–}63)$$

where

$$x = 2\omega'/(RC\omega^2) = 2Q\omega'/\omega \qquad (12\text{–}64)$$

It is to be noted that x is a function of ω as well as ω' and that $Q = L\omega/R$ is a function of ω. Since $\omega = \omega_n - \omega'$, a change of ω' from 0.01 ω_n to 0.02 ω_n produces a 1 per cent change in ω and consequently a 1 per cent change in Q. However, such a change in ω' produces nearly a 100 per cent change in x. Hence Q is a slowly varying function of w' whereas x is a rapidly varying function of ω'. As a result we can for the present purpose treat kQ in (12–63) as a constant. To obtain the maximum-minimum values of A_2, we differentiate (12–63) with respect to x and equate the result to zero. From this process, we obtain

$$x(x^2 + 1 - k^2Q^2) = 0 \qquad (12\text{–}65)$$

whence

$$x = 0 \qquad (12\text{–}66)$$

or

$$x^2 = k^2Q^2 - 1 \qquad (12\text{–}67)$$

Since x^2 must be positive, (12–67) is only possible when $kQ > 1$. For $kQ > 1$, the graph of A_2 versus x has two maxima at $x = \pm\sqrt{k^2Q^2 - 1}$ respectively and a minimum at $x = 0$. For $kQ \not> 1$, the graph has a maximum at $x = 0$. Graphs are given in Fig. 12.37 for the values of kQ shown adjacent to the respective curves.

Let us assume $Q = 200$, a not unusual value in radio receivers, so that, with $\omega/2\pi = 1000$ kc and $\omega' = 2.5$ kc, $x = 2 \times 200 \times 2.5/1000 = 1$. We note from Fig. 12.37 that for this value of x and for $kQ = 0.5$, the ratio of the response at $x = 1$ to that at $x = 0$ is $0.5/0.8 = 0.625$. Hence $kQ = 0.5$ gives good selectivity but poor fidelity. Using $kQ = 3$, the ratio is $0.65/0.60 = 1.08$ and the fidelity is good but the selectivity is poor. For $kQ = 3$ and $Q = 200$, we have $k = 0.015$. The coupling coefficient k can be changed by varying the position of the coil L_2 with respect to the coil L_1 in Fig. 12.36. The condition described by the curve for $kQ = 1$ is called *critical coupling*.

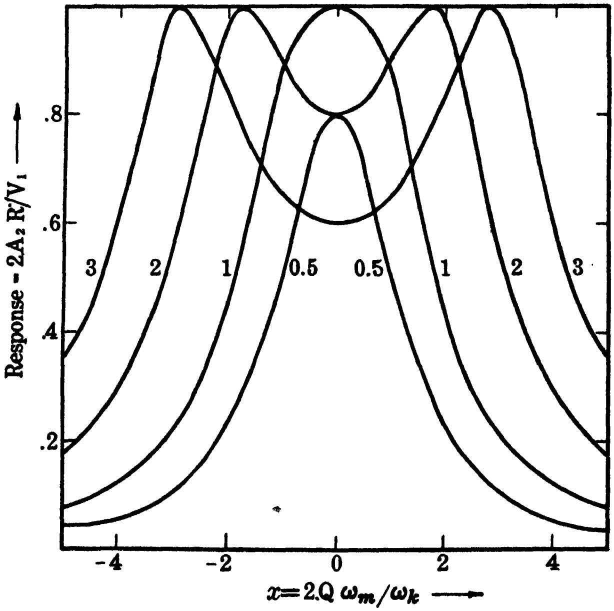


FIG. 12.37.—Response curves for the resonant transformer.

12.24. Superheterodyne Receiver.—This depends on the phenomenon of beats. In sound the student has learned that when sounded together two tuning forks of frequencies 512 and 516 cycles produce beats of frequency $516 - 512 = 4$ cycles. In this case the beats are audible to the ear as such. However, if the tuning forks have frequencies of 500 and 600 cycles the beat frequency is 100 cycles and the ear merges the beats into a continuous note of 100 c. This is known as the difference tone. There is also the summation tone of frequency $500 + 600 = 1100$ c. Modulation is a phenomenon akin to beats. From (12–53), we see that summation and difference frequencies, $(\omega_k + \omega_m)/2\pi$ and $(\omega_k - \omega_m)/2\pi$, are produced. In this case $\omega_k/2\pi$ is a radio frequency and $\omega_m/2\pi$ an audio frequency. However, both frequencies may be in the range of radio frequencies.

Suppose that the carrier frequency of an incoming radio signal is

1000 kc and that this is caused to modulate an oscillating frequency of 1455 kc developed locally in the receiver. The result is a difference or beat frequency of $1455 - 1000 = 455$ kc and a summation frequency of $1455 + 1000 = 2455$ kc. The summation frequency can be filtered out and the beat frequency of 455 kc remains. The incoming and oscillating frequencies are said to be *mixed* or *heterodyned.* If the incoming frequency is modulated by an audio frequency, so that the band width extends from 995 to 1005 kc the beat frequency is also modulated with the same band width of 10 kc but extending from 450 to 460 kc. This band contains the same audio frequency "intelligence" components as the original incoming radio frequency received by the antenna. The essential idea of the superheterodyne receiver is to change the frequency of the carrier wave to the new beat or *intermediate* frequency without affecting the nature of the modulation. The intermediate frequency usually has a fixed value of 455 kc for commercial broadcast frequencies. Consequently, the demodulating and amplifying circuits can be permanently designed to have great stability, proper selectivity and good fidelity. In the superheterodyne type of radio frequency amplification a change of the frequency of the carrier to fit the amplifier is made rather than an adjustment of the amplifier to fit the incoming carrier frequency. Most receivers of recent design are of the superheterodyne type. A "block" diagram of such a receiver is shown in Fig. 12.38.

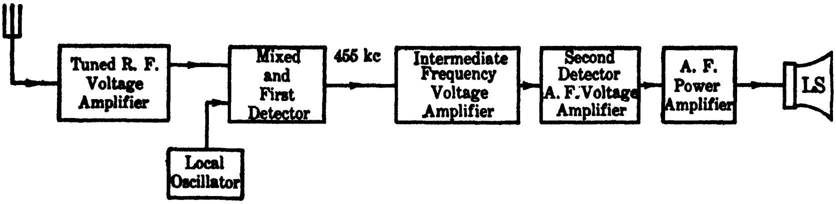


FIG. 12.38.—"Block" diagram for superheterodyne receiver. All "blocks" to the right of the local oscillator are unchangeable when once assembled.

It should be remembered that demodulation and detection describe the same process. The first detector in Fig. 12.38 is more properly described as a mixer since it mixes the 1000 kc frequency of the carrier with the 1455 kc frequency of the local oscillator to give the 455 kc intermediate frequency. In tuning to a transmitting station, the radio-frequency amplifier is tuned to the frequency of the station and the oscillator is adjusted to give a frequency always 455 kc above that of the transmitting station. The tuning and adjusting is accomplished by variable condensers which are connected mechanically in such a way that the two are controlled by turning one and the same dial. This results in great simplicity of operation.

12.25. Cathode-Ray Oscilloscope.—This is an application of the method used by J. J. Thomson in measuring the electrostatic deflection of a stream of cathode rays as described in Sec. 7.3. The modern cathode ray oscilloscope is shown diagrammatically in Fig. 12.39. It consists of a filament or heated cathode *F*, which is surrounded by a cylinder *GG* which is usually negative with respect to *F*. This cylinder focuses the electrons from *F* into a beam that is attracted to the circular plate *PP* which is maintained at positive potential of 500 to 2500 volts with respect to *F*. Some of the electrons pass through the small hole in *PP* and form a concentrated beam that passes on to the fluorescent screen *SS*, where a spot of light is produced. The beam of electrons passes between the deflecting plates *HH* and *VV*. When the voltage applied to each set of plates is zero the spot of light is on the center line of the screen *SS*. When a voltage is applied between *H* and *H* the spot is deflected horizontally and when a voltage is applied between *V* and *V* the spot is deflected vertically. If now a voltage which gradually increases from zero to E_H is applied to *HH* while the voltage between *V* and *V* is kept zero the spot of light traces a horizontal line on the screen. If, after reaching E_H, the voltage across *HH* suddenly drops to zero and the same procedure occurs again, and if the time for this procedure is less than the time of persistence of vision, a continuous horizontal line

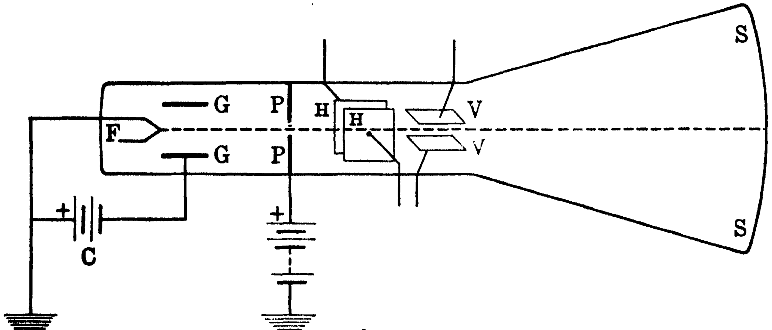


FIG. 12.39.—Cathode-ray oscilloscope.

will be seen on the screen. Now suppose that in addition a varying voltage is applied to VV. This will cause deflection of the spot in the vertical direction. If the period P_H of the voltage applied to HH is equal to or a multiple of that of the varying voltage applied to VV, the "wave form" of the voltage applied to VV is obtained. This wave form is a graph of the instantaneous voltage applied to VV plotted against the time.

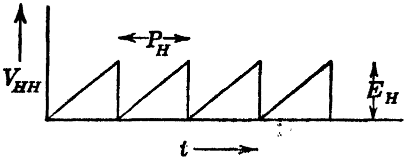


FIG. 12.40.—Graph of a saw-tooth timing wave.

In order for the graph to be correct, the *timing wave* or *sweep* applied to the deflecting plates HH must be of the *saw-tooth* variety shown in Fig. 12.40. One way of obtaining a voltage with this wave form is by means of a condenser C, a neon lamp N, and a high resistance R connected as shown in Fig. 12.41. The neon lamp will not glow (i.e., it will not carry an appreciable current) until the voltage across it has reached the *striking voltage* V_S. The lamp then glows and the condenser is rapidly discharged through the lamp. The discharge continues until the voltage across the lamp falls to the *extinction voltage* V_X. The lamp will then no longer carry current and the condenser will begin to charge up again, with the voltage gradually rising to V_S. The charging period is given by

$$P_H = \frac{CR(V_S - V_X)}{V_B - (V_S + V_X)/2} \tag{12-68}$$

where V_B is the voltage of the battery in Fig. 12.39. The discharge of the condenser through the neon lamp is very rapid and the *fly back* is for practical purposes instantaneous. If P_V is the period of the voltage ($1/P_V$ is the frequency) applied to VV, then, in order to *lock* or *synchronize* the wave form of this voltage in the *frame* of the picture as seen on the fluorescent screen (i.e., to prevent the graph from moving sideways), it is necessary to make $P_H = nP_V$ where n is an integer. Since P_V is not under control we must control P_H by means of a variable capacitance C in Fig. 12.41. This is varied until the graph locks into place.

The cathode ray oscilloscope has become one of the most useful instruments to the physicist. One of its principal purposes is the study and measurement of various wave forms of voltage and current. Not only may sinusoidal waves be studied but also square waves and pulses of very short duration. A common size of oscilloscope has a circular "face" (fluorescent screen) of 5 in. diameter, the face being divided into squares similar to those of graph paper, each smallest square having a side of 0.1 in. In the horizontal direction the sweep frequency is about 4000 cycles. The usable length is 4 inches, so that the time base is about 62.5 μ sec/in. (1 μ sec = 1 microsecond). The oscilloscope is used in radar and television.

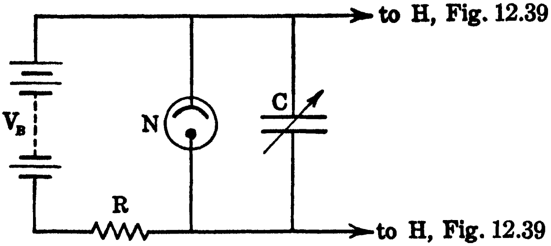


FIG. 12.41.—Circuit to give a saw-tooth timing wave.

In a typical radar set the time base is of the order of 500 μ sec extending over 5 in. An electromagnetic pulse of duration about 1 μ sec is sent out by the transmitting antenna. This pulse is registered on the face of the oscilloscope. It is reflected at any boundary between media of different dielectric constants. The reflected pulse or *echo* is received on the same antenna as that used for transmission. The echo pulse is shown on the face of the oscilloscope. Since electromagnetic pulses travel with a speed very close to that of light in a vacuum, the interval between the outgoing pulse and the echo pulse reflected from a flight of birds or an airplane at a distance of 46.5 mi is 500 μ sec. Hence the two pulses would appear on the face of the oscilloscope at a distance 5 in. apart.

In modern radar the oscillating frequency of the tank circuit is of the order of 1000 mc (1 mc = 1 megacycle = 1×10^6 cycles) and this is the frequency of the outgoing electromagnetic waves. These have a wave-length of the order of 30 cm. By means of a biasing voltage the grid of the triode in the oscillating circuit is for most of the time held at a sufficiently negative potential to stop the plate current from flowing. However, for an interval of about 1 μ sec the grid is driven to a potential sufficiently positive to cause saturation of the plate current snd oscillations are set going. In the foregoing case there are 1000 oscillations in each pulse sent out with an interval of at least 500 μ sec between pulses.

12.26. Television.—The cathode ray oscilloscope is used for reception in television. If, in addition to the saw-tooth voltage of period P_H, a saw-tooth voltage of much longer period P_V, where $P_V = kP_H$ and k is an integer, is applied to the deflection plates VV, the spot of light moves along the *scanning* lines shown in Fig. 12.42. The dotted lines represent the various *flybacks*. The scanning from A to H must be completed in an interval which is less than the period of persistence of vision (about 1/30 sec). The carrier wave falls upon the antenna of a receiving circuit such as that shown in Fig. 12.32. The average plate current through R_L varies with the modulations of the incoming waves, so that the voltage drop across R_L also varies with the modulations. This voltage drop is applied to the focusing cylinder GG of Fig. 12.39. The more negative the voltage of GG the brighter the spot on the face of the oscilloscope. This spot appears at a certain point on the path of the scanning spot of light shown in Fig. 12.42. The scanning at the receiver must be in step or in phase with the scanning at the transmitter.

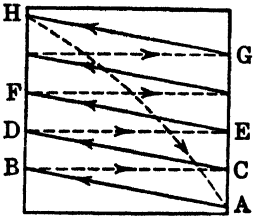


Fig. 12.42.—Path of the scanning spot of light.

An *iconoscope* is used for transmission. This is a type of cathode ray oscilloscope in which the fluorescent screen of Fig. 12.39 is replaced by a thin mica plate in contact on the back side with a conducting plate. The forward side of the mica plate in one type of iconoscope is prepared by exposing it to a mixture of silver and caesium vapors. Globules of silver mixed with caesium are laid down and are insulated from one another on the mica as shown at P, Fig. 12.43. Each globule together with the metal plate at the back of the mica forms a condenser. So the mosaic of globules forms a screen consisting of many condensers. The image of the scene to be transmitted is projected by means of a lens system upon this screen. Electrons are photoelectrically ejected from the caesium in a globule and the number n ejected is proportional to the intensity of the light falling on the globule. Thus a positive charge ne is produced on the globule and a negative charge $-ne$ on the plate P

in Fig. 12.43. The electron beam *EE* from the cathode ray oscilloscope part of the iconoscope scans the mosaic side of the screen in Fig. 12.43. As the beam passes over each globule the condenser consisting of the globule and plate is discharged with the result that a charge $-ne$ flows through R to ground, producing a momentary potential at A which is negative to ground. The point A in Fig. 12.43 is connected to the grid of a triode whose plate circuit contains a resistance R_L. Let us suppose that a certain globule is scanned by the electron beam at time t. This same globule was scanned at times $t - P_V, t - 2P_V, t - 3P_V$, etc. The negative potential of the grid of the triode at time t is proportional to the charge ne accumulated on the globule between time $t - P_V$ and time t. This accumulated charge is proportional to the average brightness of the part of the image falling upon this globule between times $t - P_V$ and t. Let us suppose that this average brightness is greater than that at surrounding globules. Then at time t the grid potential will be a negative maximum (positive minimum). Hence the plate current in the resistance R_L will be a minimum and the voltage drop across R_L will also be a minimum. The terminal of R_L remote from the plate battery is connected to the grid of a triode of an oscillating circuit. As a result the amplitude of the electromagnetic oscillations sent out by the transmitting antenna is a maximum at time t. Consequently, the average plate current in the receiver is a maximum at time t. The voltage drop in R_L of the receiving circuit is a maximum and the cylinder *GG* of Fig. 12.39 is made more negative. This in turn makes the spot on the face of the receiving oscilloscope brighter than near-by spots at time t. This bright spot corresponds to the bright spot on the iconoscope screen. In order that the picture on the fluorescent screen may not be a meaningless jumble of bright and dark spots, it is necessary that

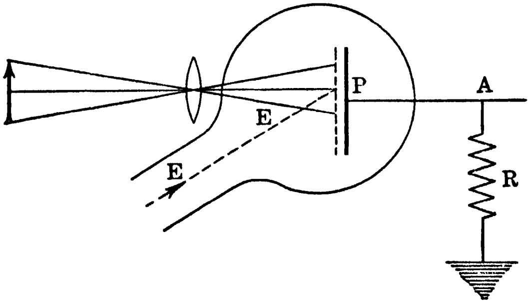


Fig. 12.43.—Iconoscope.

P_V and P_H at the transmitting end must be synchronized respectively with P_V and P_H at the receiving end. Also the cathode ray beam at the transmitter must be in step with that at the receiver.

12.27. The Ionosphere.—In his experiments of 1887, Hertz of Germany showed that electromagnetic waves could be reflected from a sheet of metal. Maxwell's theory shows that these waves are reflected from any boundary between two media of different refractive index. It was early found that the intensity of the electric waves falling on a receiving antenna was greater than that which could be expected of direct or *ground waves* from the transmitting antenna when the transmitting and receiving antennas are a considerable distance apart. This increased intensity is caused by the *sky wave* which is reflected from a conducting layer in the atmosphere. This conducting layer was proposed by Kennelly of Harvard and Heaviside of England. It is now believed that instead of one layer there are several layers or regions. These regions constitute the ionosphere. The lowest layer occurs at an altitude of about 70 miles. Since electromagnetic waves travel with the velocity of light, the time taken upwards and downwards along a vertical line is about 750 μ sec. As in radar a sharp pulse a few microseconds wide is sent out and is shown on a cathode-ray oscilloscope. The returning pulse is observed on the oscilloscope and so the altitude is found. Another conducting region is found at an altitude of about 175 miles. The conductivity of the air in these layers is believed to be caused partly by the ultraviolet rays of sunlight. The ultraviolet rays are absorbed in the atmosphere and the intensity of these rays arriving at sea level is very small. These ultraviolet rays are more intense high in the atmosphere. They give rise to the photoelectric effect in which electrons are ejected from the atoms of the air, leaving these atoms positively charged. Such charged atoms are called *ions*. The conducting layer consists of electrons and ions.

The conductivity of the layer or layers changes from day to night and it is in some way connected with the earth's magnetic field. During a magnetic storm the layers appear to have irregular and ill-defined boundaries so that interference effects occur and radio transmission is poor. Fading is caused by varying interference between the ground and reflected sky waves.

CHAPTER XII

Problems

1. Determine the constant *A* in (12–3) for the thyrite disc described in Sec. 12.1. Determine the resistance of the disc for (*a*) 600, and (*b*) 1500 volts.

2. A circuit element whose characteristic is $i = AV^2$, where $A = 2 \times 10^{-6}$ ohm^{-1} volt^{-1}, is connected in series with an ohmic resistance R. Find the value of R to give equal voltage drops across the circuit element and the resistance when the potential across the whole circuit is 400 volts. Find the voltage drop across the circuit element when the potential across the whole circuit is 100 volts.

3. A battery of E_B volts is connected to a resistance of 2000 ohms in series with a 6*L*6 tube operated as a triode (see Fig. 12.6). Determine the value of E_B to give a plate current of 50 ma when the grid bias is -22.5 volts.

4. A circuit element consists of a resistance of 2 megohms (1 megohm = 10^6 ohms) and a capacitance C. Determine the value of C to make the time constant 5×10^{-4} sec. Find the time for the charge on the condenser to be reduced to one tenth its initial value.

5. Find the value of L_1 in the circuit of Fig. 12.15 in order to make the oscillating frequency 1000 kc when $R_P = 1320$ ohms, $R = 2.64$ ohms, and $C = 50$ $\mu\mu$f.

6. In a parallel resonant circuit element, $R = 0$, $L = 636$ μh, $C = 159$ $\mu\mu$f. Determine the resonant frequency. Find the impedance at (*a*) the resonant frequency, (*b*) 10 kc above the resonant frequency, (*c*) 20 kc below the resonant frequency.

7. Find (*a*) the Q of the tank circuit in Problem 5 and (*b*) the number of oscillations for the amplitude of the current to be reduced to $1/e$ of the amplitude at the time when the tank circuit was disconnected from the driving battery and triode.

8. In a circuit similar to that of Fig. 12.23, the e.m.f. of the battery is 250 volts, the plate and load resistances are 1500 and 1200 ohms, respectively. Determine the resistance in parallel with the by-pass condenser to give a grid bias of 10 negative volts when the amplification factor of the triode is 6.0.

9. A signal voltage of frequency 750 cycles is applied to the circuit of Problem 8. Find the capacitance of the by-pass condenser which with its shunt resistance has a time constant ten times the period of the signal voltage. Using this capacitance, find (*a*) the complex impedance of the by-pass condenser and shunt resistance, (*b*) the total impedance Z of the plate circuit to the nearest ohm, (*c*) the amplitude of the a.c. component of the plate current and the voltage drop across the load resistance R_L when the amplitude of the signal voltage is 5 volts. Determine the a.c. and d.c. components of the plate current which pass through the shunt resistance.

10. Determine the a.c. power dissipated in the load resistance R_L of Problem 9 when the amplitude of the signal voltage is 7 volts. Find the d.c. power dissipated.

11. Find the value of $f(t)$ out to the 12th harmonic in (12–55) when $t = \pi/(2\omega_k), \pi/(\omega_k), 3\pi/(2\omega_k)$.

12. The course of a projectile is followed by a radar set placed near a gun. The transmitted waves have a frequency of 100 megacycles. The echo waves from the projectile are heterodyned with the transmitted waves and a note of audio frequency 320 cycles is heard. Find the speed of the projectile.

CHAPTER XIII

RELATIVITY—SPECIAL THEORY

13.1. Velocity of Light.—The discovery that light takes time to travel from one point to another was made by the Danish astronomer Roemer in 1676. He found that the intervals between successive eclipses of one of the moons of Jupiter were longer when the earth was receding from Jupiter and shorter when approaching. Fizeau of France made the first determination of the velocity of light between points on the earth's surface in 1849. The best modern results for the velocity of light in a vacuum give a value of 2.99776×10^{10} cm/sec, or 3×10^{10} cm/sec to within less than one part in a thousand. The velocity of light is often remembered as 186,000 mi/sec.

Owing to the rotation of the earth about its axis the speed of a point on the equator of the earth is 25,000 mi/day = 0.29 mi/sec. The speed of the earth in its orbit around the sun is $2\pi \times 93{,}000{,}000$ mi/year = 18.5 mi/sec. We may therefore neglect the speed due to the rotation of the earth upon its axis. In 1726 Bradley, the English Astronomer Royal, observed the aberration of the fixed stars in or near the plane. When the earth is moving at right angles to the direction of a line joining a star to the earth, a telescope must be pointed in a direction making an angle θ with this line in order for the rays of light from the star to pass straight down the tube of the telescope. If the earth is moving to the right of the line, θ must be to the right of the line. Six months later, θ is found to be to the left of the line. If light travels through space with a velocity c and the earth is moving sideways with velocity v, $\tan \theta = v/c$. From the measured value of θ, v was found to be 18.5 mi/sec, which agrees with the speed calculated for the earth in its orbit around the sun.

The Maxwell electromagnetic wave theory of light, described in Chapter V, seems to require the presence of a medium—the ether—in order to support the electromagnetic vibrations from the place of emission to the place of detection. The theory, which was developed in 1861–64, predicted the existence of electromagnetic waves. Hertz of Germany verified the existence of these electromagnetic waves in 1887.

The electromagnetic waves which are emitted by an antenna obey the laws of the Maxwell theory. There are no phenomena in the range of ordinary radio waves—from 500 to 1500 kc—which require the quantum theory for their explanation. Perhaps the lack of quantum phenomena in radio waves is due to the small energy of a radio quantum or photon. Assuming a frequency of 1000 kc, the energy of a radio photon is $h\nu = 6.62 \times 10^{-27} \times 10^6 = 6.62 \times 10^{-21}$ erg $= 6.62 \times 10^{-21}/1.6 \times 10^{-12} = 4.14 \times 10^{-9}$ ev. The energy of a photon of radio frequency is so far below the photoelectric or thermionic work function of any known substance that quantum phenomena cannot be observed.

Hertz's discovery of electromagnetic waves seemed to prove the existence of the ether. But if the ether exists, the velocity of light must be relative to the ether. Bradley's discovery of aberration seems to show that the earth is moving through the ether with a velocity of 18.5 mi/sec.

Let us now consider the case of a swimmer in a river. The speed of the swimmer is 3 ft/sec in still water and the speed of flow of the river is 1.8 ft/sec. To reach a point 300 ft distant requires $300/(3 - 1.8) = 250$ sec if the swimmer is swimming against the current and $300/(3 + 1.8) = 62.5$ sec if he is swimming with the current. If he swims to a point 300 ft distant and returns, the total time will be $(250 + 62.5) = 312.5$ sec. Next, suppose that he swims to a point 300 ft distant but across the stream and then swims back. In order to reach the point he must swim in a direction making an angle $\sin^{-1}(3/1.8)$ with the direction of straight across. His relative velocity in the direction of straight across is $\sqrt{3^2 - 1.8^2} = 2.4$ ft/sec. The total time in this case is $(2 \times 300)/2.4 = 250$ sec. This is distinctly less than the time of 312.5 sec when he swims up and down the river. In the early 1880's an experiment based upon the principles just described was tried by Michelson and Morley at Chicago. In the above example, the speed of the river could have been treated as an unknown. Then from the different times, 312.5 and 250 sec, the speed of the river relative to the earth or of the earth relative to the river could have been obtained. But in their analogous experiment with light Michelson and Morley obtained the astounding result that they could detect no certain difference in the two times. Their result was published in 1887, the year of Hertz's discovery.

13.2. Michelson-Morley Experiment.—In order to measure the velocity of the earth through the ether Michelson and Morley used the Michelson interferometer which we shall now describe. A beam of light

AB as in Fig. 13.1 is partly reflected by the back surface of a block of glass and partly transmitted. The reflected beam strikes the mirror *E* and is reflected back along the path *EDCK*. The transmitted beam strikes the mirror *H* and is reflected back along the path *HGFCK*. If the beam *AB* contains light of only one wave-length, there is interference when the path difference between the beams which strike the two mirrors and return to *C* is equal to an odd number of half wave-lengths and there is re-enforcement at *K* when the path difference is a whole number of wave-lengths. Let us suppose that *CH* and *CE* are equal in length. Then, if *CH* is in the direction in which the earth is moving and *CE* is perpendicular to this direction, there should be a difference between the time for the light to travel over the path *CEC* and the time for the path *CHC*. Let c be the velocity of light relative to the ether and v the velocity of the earth. Then, while the light is traveling from *C* to *H*, its velocity relative to the apparatus is $c - v$. If d is the distance *CH*, the time for light to travel from *C* to *H* is $d/(c - v)$, while the time for light to travel from *H* to *C* is $d/(c + v)$. Hence, the total time for light to travel from *C* to *H* to *C* is

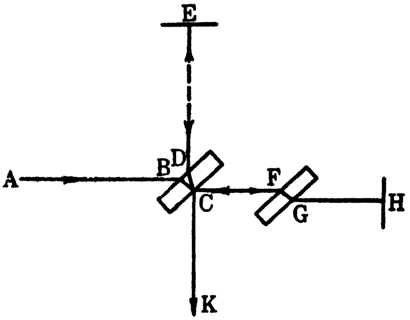


Fig. 13.1.—Michelson interferometer.

$$t_1 = \frac{d}{c - v} + \frac{d}{c + v} = \frac{2cd}{c^2 - v^2} = \frac{2d\left(1 + \frac{v^2}{c^2}\right)}{c} \qquad (15\text{-}1)$$

approximately.

Next, consider the time for light to travel from *C* to *E* to *C*. Referring to Fig. 13.2, the point *C* moves to *C′*, due to the motion of the earth, while the light travels from *C* to *E* and back. The path in the ether is *C* to *E* to *C′* and the velocity in the ether is c, and this is the velocity along *CE* and along *EC′*. The velocity relative to the apparatus is along *DE*. This relative velocity is the vector difference of c along *CE* and v along *CD*. Therefore, in the triangle *CDE*, the line *CE* is proportional to c, *CD* to v, and the relative velocity of the light to the apparatus is proportional to *DE*. Hence, the velocity of the light relative to the apparatus is $\sqrt{c^2 - v^2}$. The distance to be traveled on the apparatus, however, is $DE + ED$, which is $2d$ as measured on the apparatus. Hence, the time to go from C to E to C in Fig. 13.1 is

$$t_2 = \frac{2d}{\sqrt{c^2 - v^2}} = \frac{2d}{c}\left(1 + \frac{v^2}{2c^2}\right) \tag{13-2}$$

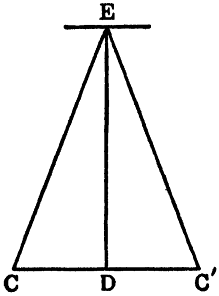


FIG. 13.2.

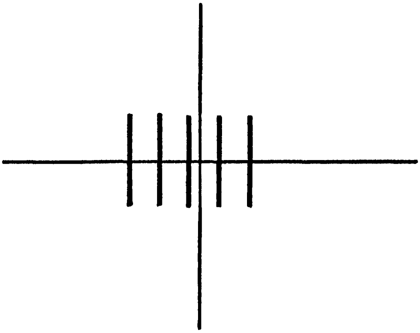


FIG. 13.3.—Interference fringes.

Hence, the difference of the times for the two paths is

$$t_1 - t_2 = \frac{2d}{c} \times \frac{v^2}{2c^2} \tag{13-3}$$

In this time light travels a distance $c(t_1 - t_2)$ and the number of wave-lengths in this distance is $c(t_1 - t_2)/\lambda$ where λ is the wave-length of the light. Hence the effective optical path difference of the two paths CHC and CEC, Fig. 13.1, is

$$\text{Path difference} = \frac{v^2 d}{c^2 \lambda} \text{ wave-lengths} \tag{13-4}$$

The Michelson interferometer is used with a telescope at K, Fig. 13.1. In the field of the telescope light and dark fringes are seen as in Fig. 13.3. Two adjacent dark fringes are produced by paths whose differences CHC–CEC differ by one wave-length. If now in the Michelson-Morley experiment the apparatus is turned through a right angle (the apparatus floats on a sea of mercury so as to reduce friction and strains in the apparatus), the path difference CHC–CEC changes from $+v^2d/c^2\lambda$ to $-v^2d/c^2\lambda$ wave-lengths, or a total change of $2v^2d/c^2\lambda$. Hence, in turning the apparatus through a right angle the number of fringes which pass the cross wires of the telescope is $2v^2d/c^2\lambda$.

In one of Michelson and Morley's first experiments the distance d was 11 meters. With a wave-length of 5000 angstroms, a shift of $2 \times 18.5^2 \times 1100/(186{,}000^2 \times 5000 \times 10^{-8}) = 0.4$ of a fringe width was expected. The actual displacement observed was less than $\frac{1}{20}$ of this amount. More recent experiments as late as 1929 have shown that the effect, if it exists at all, is much less than $\frac{1}{20}$ of the expected amount.

The Michelson-Morley experiment gave a negative result and is an example of an important experiment which gives such a result. The theory of the expected result is therefore wrong and it is necessary to devise a new theory.

13.3. Fizeau's Experiment.—Fizeau measured the velocity of light in water which was moving with a velocity v and compared the velocity with that when the water was not moving. He expected to find that either the velocity of light in water was unaltered by the velocity of the water, or that it was increased by an amount equal to the velocity of the water. He found neither. He found that, if the velocity of light in still water is V and the velocity of the moving water is v, the velocity of light in the moving water is $V + kv$, where k is a fraction. The water neither drags the ether completely along with it nor slips through the ether. The fraction k is sometimes called Fresnel's coefficient of ether drag since Fresnel developed the theory involved. The value of k is found to be given by $1 - 1/\mu^2$ where μ is the index of refraction in the water. Other liquids were tried and k was still found to be given by this formula. Since μ is very nearly unity for air, it seems that there should be no ether drag in the Michelson-Morley experiment. This makes it all the more remarkable that Michelson and Morley obtained a negative result.

13.4. Trouton-Noble Experiment.—The apparatus in this experiment consists of an electrical condenser suspended with its plates vertical by a very fine wire. The charges on the plates are passed into them by means of this wire and by a wire hung from beneath and dipping into mercury. The electric field between the plates, since it is moving through the ether due to the earth's motion, produces a magnetic field, as we have seen in Chapter V. The magnetic field varies as $\sin \theta$ where θ is the angle between the electric lines of force of the condenser and the direction of the earth's motion. Since the energy of the magnetic field varies as H^2, the energy of the field varies as $\sin^2 \theta$. The condenser therefore tends to turn so as to make the magnetic energy a minimum and this occurs when $\sin\theta = 0$. In the experiment a charge is given to the plates and a turning effect is looked for, but no such effect is observed, and we have another negative result, showing that the theory of the expected but unfound result is wrong.

13.3. Einstein's Theory.—Various explanations of the result of the Michelson-Morley experiment were attempted in the period 1887 through 1904. Fitzgerald of Dublin, Ireland, suggested that the result could be explained if the length CH becomes $d\sqrt{1 - v^2/c^2}$. This suggestion was examined from the point of view of Maxwell's theory by Larmor of England and Lorentz of Holland. They showed that the theory demanded that the electric lines of force given out by a charged sphere tended to concentrate in a plane at right angles to the motion of the sphere. Since there is a tension along the lines of force this concentration might tend to pull the sphere into an oblate spheroid which would be contracted in the direction of motion of the sphere. These ideas were applied to the electron and Lorentz showed from the Maxwell theory that there should be a contraction, which came to be known as the Lorentz-Fitzgerald contraction. However, there must be non-electrical forces holding the electron together since Earnshaw's theorem in electrostatics states that a system of electrical charges cannot be in equilibrium under the action of the electrical forces between the charges alone. But the nature of the nonelectrical forces was unknown and the exact amount of the contraction was not calculable.

Principally as a consequence of the negative results of the Michelson-Morley and Trouton-Noble experiments it began to dawn on the physicists of the time that there must be something very wrong with the "swimmer and river" theory of Sec. 13.1 and with the analogous theory given in Sec. 13.2. The theoretical predictions are all based on the idea of the earth moving through space. But what is space? A man nay walk along a train and he moves relatively to the train. The train noves relatively to the ground. But the earth is rotating on its own ixis and is also revolving around the sun. The man's motion in space s therefore very complicated. But is there such a thing as absolute notion? Just how would one conduct an experiment on a train to find ut whether or not the train is moving if one is not allowed to look out f the windows? There seems to be no way of finding out whether or ot the train is moving unless we are allowed to observe some outside ody such as the earth. Then we can say that the train has a velocity elative to the earth. But how do we know that it is not the earth which the formula for any physical law in Frame 2 will be the same as the formula in Frame 1. If this were not so, similar experiments in the two frames would give different results. For instance, an electron at rest in Frame 1 will produce no magnetic field in Frame 1, and therefore an electron at rest in Frame 2 will produce no magnetic field in Frame 2, although it will produce a magnetic field in Frame 1. A magnetic field produced by lines of electrostatic force is therefore a matter of relative velocity between the electrostatic field and the observer. In the case of the Trouton-Noble experiment the electrostatic field between the plates of the condenser is at rest relative to the observer (both being at rest to the same frame of reference—the laboratory) and hence no magnetic field should be produced, which is in agreement with the experimental result.

13.7. Lorentz Transformation Equations.—Let us suppose that Frame 2 in Fig. 13.5 has the axes $O'X'$, $O'Y'$ and $O'Z'$ parallel to the axes OX, OY and OZ of Frame 1. Let us suppose that the time of an event as measured in Frame 2 is represented by t' and that the time of the same event as measured in Frame 1 is represented by t, and let us further suppose that at $t = 0$ and $t' = 0$ the origin O' coincides with the origin O. Also let us suppose that Frame 2 is moving parallel to OX with a velocity v relative to Frame 1. Consider a point P fixed relative to Frame 2 and having the coordinates x', y' and z' in this frame. At $t = 0$ let the coordinates of P in Frame 1 be x, y, z. Quantities measured in Frame 2 are distinguished from the corresponding quantities measured in Frame 1 by a prime ($'$). Now consider the coordinates x, y, z of P at some subsequent time $t = t$ in Frame 1. According to the old Newtonian ideas of space we have

$$\left.\begin{aligned} x &= x' + vt' \\ y &= y' \\ z &= z' \end{aligned}\right\} \qquad (13\text{–}5)$$

Also according to Newtonian principles the time t' in Frame 2 is related to the time t in Frame 1 by

$$t = t' \qquad (13\text{–}6)$$

Lorentz, however, found that Maxwell's system of electromagnetic equations which we have discussed in Chapter V became different when changed from Frame 1 to Frame 2 if the Newtonian relations (13–5) and (13–6) are assumed to be true. Hence, according to Lorentz, the law of physics expressed by Maxwell's equations is different in Frame 2 from the law in Frame 1 and it is therefore possible to find out by experiments in either frame without reference to any other frame of reference whether the frame is moving in space. This is opposed to the postulate that such motion cannot be shown, and Lorentz concluded that the Newtonian relations (13–5) and (13–6) are incorrect. The relations (13–5) and (13–6) are known as transformation equations. Lorentz then sought a system of new transformation equations such that Maxwell's equations remain the same when changed from Frame 1 to Frame 2. The Lorentz system of transformation equations is

$$\left.\begin{aligned} x &= k(x' + vt') \\ y &= y' \\ z &= z' \end{aligned}\right\} \tag{13–7}$$

and

$$t = k\left(t' + \frac{vx'}{c^2}\right) \tag{13–8}$$

where

$$k = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{13–9}$$

A most remarkable consequence of these new transformation equations is that t is not equal to t'. In other words, clocks in the two frames seem to run at different rates. If two events happen at the same place in Frame 2 but at times t'_1 and t'_2, so that the time interval between them is $t'_2 - t'_1$ as measured by a clock in Frame 2, the interval between the same two events as measured by a clock in Frame 1 is $t_2 - t_1$, which from (13–8) is

$$t_2 - t_1 = k(t'_2 - t'_1) \tag{13–10}$$

But k is greater than unity, so that the interval as measured in Frame 1 is longer.

Consider the stick shown in Fig. 13.6. This stick is at rest in Frame 2. The positions of the two ends of the stick relative to Frame 2 are x'_A and x'_B. The length $x'_B - x'_A$ can be measured by means of a meter stick or a surveyor's chain by an observer in Frame 2. This length does not involve the time of an event at either end of the stick. But the length of the stick to an observer in Frame 1 cannot be measured by such simple operations. A system of clocks must be placed along the axis of x in Frame 1. These clocks are synchronized. The time t_B when the end B of the stick passes the point x_B in Frame 1 is noted. At the same time $t_A = t_B$, the position x_A of the end A of the stick is noted. The length of $x_B - x_A$ is then called the length of the stick in Frame 1. We now introduce the times t'_A and t'_B in Frame 2, which correspond to the times t_A and t_B in Frame 1. Transforming by means of (13–7), we obtain

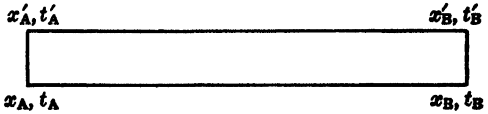


FIG. 13.6.—Stick at rest in Frame 2.

$$x_A = k(x'_A + vt'_A)$$

$$x_B = k(x'_B + vt'_B)$$

whence

$$x_B - x_A = k(x'_B - x'_A) + kv(t'_B - t'_A) \tag{13–11}$$

Next, transforming by means of (13–8), we obtain

$$t_A = k\left(t'_A + \frac{vx'_A}{c^2}\right)$$

$$t_B = k\left(t'_B + \frac{vx'_B}{c^2}\right)$$

whence

$$t_B - t_A = k\left[(t'_B - t'_A) + \frac{v(x'_B - x'_A)}{c^2}\right] \tag{13–12}$$

we have assumed that the positions x_A and x_B in Frame 1 were taken at the same time so that $t_A = t_B$. Hence, from (13–12)

$$t'_B - t'_A = -\frac{v(x'_B - x'_A)}{c^2} \tag{13–13}$$

Substituting in (13–11) and remembering (13–9), we obtain

$$x_B - x_A = \frac{x'_B - x'_A}{k} \tag{13–14}$$

so that there is an apparent contraction since $k > 1$. It should be remembered, however, that, whereas the measurement of $x_B - x_A$ involves the use of clocks, the measurement of $x'_B - x'_A$ does not. This brings in the operational viewpoint which we shall discuss in a later chapter.

From (13–13) and (13–14)

$$t'_B - t'_A = -\frac{kv(x_B - x_A)}{c^2} \qquad (13\text{–}15)$$

so that events which are simultaneous ($t_A = t_B$) in Frame 1 are not simultaneous in Frame 2. If $x_B > x_A$, event B happens before event A in Frame 2. It is a matter of the relative situation of the observer to the events.

13.8. Addition of Velocities.—The Lorentz transformation equations lead to the relativity formula for the addition of velocities. Again, we consider a stick at rest in Frame 2, but we do not assume that events A and B are simultaneous in Frame 1. Let a body be traveling in Frame 2 with a velocity u' such that the body passes the left-hand end of the stick at time t'_A and the right-hand end at time t'_B. The distance $(x'_B - x'_A)$ divided by the time $(t'_B - t'_A)$ gives the velocity in Frame 2, so that

$$u' = \frac{(x'_B - x'_A)}{(t'_B - t'_A)} \qquad (13\text{–}16)$$

As observed from Frame 1, however, the velocity is

$$u = \frac{(x_B - x_A)}{(t_B - t_A)} \qquad (13\text{–}17)$$

which in virtue of (13–11) and (13–12) becomes

$$u = \frac{k(x'_B - x'_A) + kv(t'_B - t'_A)}{k(t'_B - t'_A) + \dfrac{kv(x'_B - x'_A)}{c^2}} \qquad (13\text{–}18)$$

Dividing numerator and denominator by $(t'_B - t'_A)$, we obtain in virtue of (13–16)

$$u = \frac{u' + v}{1 + \dfrac{u'v}{c^2}} \qquad (13\text{-}$$

the relativity formula for the addition of velocities.

It is worth while to test (13–19) for particular cases. For the case of $u' = c$ and $v = v$ which is less than c, (13–19) gives $u = c$ and Einstein's condition that the resultant velocity cannot be greater than c is fulfilled. Hence, if a velocity v is added to the velocity c, the resultant is still c. Likewise, if a velocity v is subtracted from the velocity c, the resultant is again c. Thus in the Michelson-Morley experiment the velocity of light from C to H, Fig. 13.1, is not $c - v$, as we previously supposed, but is c. It is therefore seen that Eq. (13–19) for the addition of velocities satisfies Einstein's requirement that the velocity of light is always the same.

13.9. Variation of Mass with Velocity.—In order to derive the formula for the variation of mass with velocity, we shall give one definition and make one tentative assumption. We define the momentum of a body as mv, where m is the mass and v the velocity of the body. However, in relativity dynamics m varies with the velocity, while in Newtonian dynamics m is a constant and independent of the velocity. We tentatively assume that when a number of bodies interact and no energy is lost the total mass of the bodies in a given frame of reference remains constant. Let us suppose that in Frame 2, two bodies of equal masses m' and traveling with velocities u' and $-u'$ parallel to the axis x' collide and after a collision coalesce into one body. In virtue of our tentative assumption, the mass of the coalesced bodies after collision is $2m'$. The principle of the conservation of momentum also holds so that the coalesced bodies are at rest in Frame 2 after collision. Now let us consider the appearance of the experiment which we have just described to an observer in Frame 1. The velocities u' and $-u'$ transform into u_1 and u_2 where

$$u_1 = \frac{u' + v}{1 + \frac{u'v}{c^2}} \tag{13–20}$$

and

$$u_2 = \frac{-u' + v}{1 - \frac{u'v}{c^2}} \tag{13–21}$$

The mass of the body traveling with velocity u_1 is m_1 and that of the body traveling with velocity u_2 is m_2. The sum of the momenta before collision is equal to the total momentum after collision, so that

$$m_1u_1 + m_2u_2 = (m_1 + m_2)v \tag{13–22}$$

After the collision the coalesced bodies travel with velocity v in Frame 1 since they are at rest in Frame 2. Substituting the expressions for u_1 and u_2 given by (13–20) and (13–21) in (13–22) we obtain

$$\frac{m_1}{m_2} = \frac{\left(1 + \dfrac{u'v}{c^2}\right)}{\left(1 - \dfrac{u'v}{c^2}\right)} \tag{13–23}$$

(13–20) and (13–21) may be written respectively in the forms

$$u_1 + \frac{u_1u'v}{c^2} = v + u' \tag{13–24}$$

$$u_2 - \frac{u_2u'v}{c^2} = v - u' \tag{13–25}$$

Squaring, subtracting, and rearranging, we obtain

$$\left(\frac{u'v}{c^2}\right)^2 - 2\left(\frac{2c^2 - u_1^2 - u_2^2}{u_1^2 - u_2^2}\right)\left(\frac{u'v}{c^2}\right) + 1 = 0 \tag{13–26}$$

This is a quadratic in $u'v/c^2$ and its solution is

$$\frac{u'v}{c^2} = \frac{2c^2 - u_1^2 - u_2^2 - 2\sqrt{(c^2 - u_1^2)(c^2 - u_2^2)}}{u_1^2 - u_2^2} \tag{13–27}$$

Substituting in (13–23), we obtain

$$\frac{m_1}{m_2} = \frac{\sqrt{1 - \dfrac{u_2^2}{c^2}}}{\sqrt{1 - \dfrac{u_1^2}{c^2}}} \tag{13–28}$$

If the body of mass m_2 is moving with zero velocity in Frame 1 before the collision so that $u_2 = 0$, then

$$m_1 = \frac{m_2}{\sqrt{1 - \dfrac{u_1^2}{c^2}}} \tag{13–29}$$

Now both bodies have the same mass when moving with the same velocity so that (13–29) means that m_1 is the mass of a body whose rest mass is m_2 when the body is moving with velocity u_1. In the usual notation $m_1 = m$, $m_2 = m_0$ and $u_1 = u$, so that

$$m = \frac{m_0}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{13–30}$$

This is the relativity formula for the change of mass with velocity. We have proved (13–30) for a head-on collision where the bodies coalesce. Tolman has shown that the formula holds for collisions of any type.

13.10. Momentum and Kinetic Energy.— The relativity formula for momentum is mu. But, since m is related to the rest mass m_0 according to (13–30), we have

$$\text{Momentum} = m_0\beta c/\sqrt{1 - \beta^2} \tag{13–31}$$

where $\beta = u/c$.

In the relativity theory force is not given by the familiar relation $F = ma$, but is defined by Newton's second law, which states that the rate of change of momentum is proportional (or equal, if we devise our units to make it so) to the unbalanced force acting on the body. Hence

$$F = \frac{d(mu)}{dt} \tag{13–32}$$

Since m is not a constant in the relativity theory, (13–32) becomes

$$F = m\frac{du}{dt} + u\frac{dm}{dt} \tag{13–33}$$

If the force F moves a body through a distance dx the work done is Fdx. If the body is free to move, the kinetic energy of the body is increased by an amount $dW = Fdx$. If the body starts from rest, the total kinetic energy developed is

$$W = \int Fdx = \int m\frac{du}{dt} \cdot dx + \int u\frac{dm}{dt} \cdot dx \tag{13–34}$$

Now $(du/dt) \cdot dx$ may be written $du \cdot dx/dt$, and, since $dx/dt = u$, we have $(du/dt) \cdot dx = udu$. Similarly, also $(dm/dt) \cdot dx = udm$. Hence, (13–34) may be written in the form

$$W = \int mudu + \int u^2dm \tag{13–35}$$

But m is given by (13–30), so that

$$W = m_0 \int \frac{u\,du}{\sqrt{1 - \frac{u^2}{c^2}}} + \frac{m_0}{c^2} \int \frac{u^3\,du}{\left(1 - \frac{u^2}{c^2}\right)^{3/2}} \qquad (13\text{–}36)$$

Since the body starts from rest, the limits of the integrals are $u = 0$ and $u = u$. The following two integrals are necessary

$$\int \frac{u\,du}{\sqrt{1 - \frac{u^2}{c^2}}} = -\,c^2\sqrt{1 - \frac{u^2}{c^2}} \qquad (13\text{–}37)$$

and

$$\int \frac{u^3\,du}{\left(1 - \frac{u^2}{c^2}\right)^{3/2}} = \frac{c^4\left(2 - \frac{u^2}{c^2}\right)}{\left(1 - \frac{u^2}{c^2}\right)^{1/2}} \qquad (13\text{–}38)$$

Putting in limits, we find that (13–36) becomes

$$\text{Kinetic energy} = m_0c^2\left(\frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} - 1\right) \qquad (13\text{–}39)$$

This reduces to the Newtonian formula, kinetic energy = $(\frac{1}{2})mu^2$, for values of u small with respect to c the velocity of light.

13.11. Mass and Energy.—Using (13–30) we may write (13–39) in the form

$$W = c^2(m - m_0) \qquad (13\text{–}40)$$

or

$$m - m_0 = \frac{W}{c^2} \qquad (13\text{–}41)$$

Thus, owing to its kinetic energy W, the mass of the body is increased by an amount W/c^2. We therefore think of the energy W as having a mass W/c^2. This has been proved here for the case of kinetic energy. Einstein, however, has proved the relation

$$\text{mass} = \text{energy}/c^2 \qquad (13\text{–}42)$$

for any kind of energy. As examples, a fly-wheel which is rotating has a greater mass than the same wheel at rest; also a wound-up watch spring has a greater mass than when unwound. Of course in both these instances the increase of mass is not possible of detection because the energy is divided by c^2 or by 9×10^{20} and this makes the increase of mass exceedingly small.

In Sec. 13.9 we made a tentative assumption which we may describe as the principle of the *conservation of mass.* When an elastic collision occurs between two bodies the kinetic energy is conserved and no heat is produced. However, when the two bodies coalesce the collision is inelastic, the kinetic energy is not conserved, and heat is produced. The rest mass of each body of mass m' in Frame 2 is, from (13–30), $m'\sqrt{1 - u'^2/c^2}$ and so the total kinetic energy of the two bodies before collision was, from (13–39), $2m'c^2(1 - \sqrt{1 - u'^2/c^2})$. This kinetic energy appears as heat in the coalesced bodies. This heat energy has a mass $2m'(1 - \sqrt{1 - u'^2/c^2})$. Adding this to the mass $2m'\sqrt{1 - u'^2/c^2}$, we obtain $2m'$, the total mass of the two bodies before collision, and this agrees with our tentative assumption. Also we note that a hot body has a greater mass than a cold body.

CHAPTER XIII

PROBLEMS

1. A β-particle is moving to the left with a speed of 2.75×10^{10} cm/sec while a second β-particle is moving to the right with a speed of 2.50×10^{10} cm/sec. Find the relative velocity of one β-particle to the other.

2. Find the kinetic energy in ergs of a β-particle when it is moving with each of the following speeds: 2.0×10^{10} cm/sec, 2.5×10^{10} cm/sec, and 2.95×10^{10} cm/sec.

3. A meter stick is at rest in Frame 2. What is the length of the stick seen by an observer in Frame 1, given that Frame 2 is moving in the direction of the length of the stick relatively to Frame 1 with a speed of 2.5×10^{10} cm/sec?

4. A clock in Frame 2 of Problem 3 indicates that 1 hour has passed between two events in Frame 2. What is the length of time between the same two events to an observer in Frame 1, given that the two events occur at the same place in Frame 2?

5. What is the mass of an electron which is moving relatively to the observer with a speed of 1.8×10^{10} cm/sec?

6. A fly-wheel in the shape of a disc of radius 90 cm is rotating at a speed of 1,000 revolutions per minute. Given that the mass of the wheel is 5,500 kilograms, find (*a*) the moment of inertia of the wheel, (*b*) the kinetic energy of the wheel, and (*c*) the increase of mass of the wheel due to its kinetic energy.

CHAPTER XIV

X-RAYS

14.1. Beginning of the New Era in Physics.—We have seen in Sec. 7.6 that Plücker discovered cathode rays in 1859. However, it was not until 1897 that J. J. Thomson and his associates at Cambridge, England, recognized the true nature of cathode rays. We now recognize these rays as streams of electrons. Physicists, in their pride, often marvel at the rapidity of advances in physics. However, there are many cases—the recognition of the electron, for instance—of the extreme slowness of advance in a particular field of physics.

In early 1895 three concepts were firmly entrenched in the minds of physicists: the atom, the ether, and electricity. The atom was indestructible and impenetrable. The ether was necessary as a medium for the propagation of light waves in a vacuum and for the propagation of the electromagnetic waves discovered by Hertz in 1887. Although no one seemed to know just what electricity was, yet all physicists believed they knew how electricity behaved. Electricity in its electrical engineering aspects was understood. These aspects are explained in Chapters IV and V of this book. But there was something wrong in spite of the general complacency of the physicists. First, the cathode rays would not fit into the picture. Then there was the Michelson-Morley experiment which was first performed in 1887 and is described in Chapter XIII. This was followed by the discovery of the photoelectric effect in 1888 by Hallwachs. The true nature of this effect was recognized by Einstein in 1904. Lastly, in 1894 Lenard of Germany passed cathode rays through a solid, though thin, aluminum window out into the air. This gave the Crookes school of physicists in England a jolt for this school believed that cathode rays were charged molecules (or, perhaps, atoms) and could not penetrate the molecules (or atoms) of the aluminum window. The penetration could be understood if the window were porous but its lack of porosity was attested by the fact that it could support a "vacuum" of 0.001 mm of mercury against the pressure of the outside air. The German school of physicists believed that cathode rays were ethereal vibrations and so might be able to pass through aluminum just as light passes through glass. And so Lenard, a German, found cathode rays out in the air—a phenomenon which no English physicist, because of his cherished notions and established habits of thought, could possibly have looked for.

In 1894 the principal physicists working in the field of cathode rays were, perhaps, Crookes in England and Lenard in Germany. In spite of Lenard's experiment, the three concepts of the atom, the ether, and electricity remained intact. It seems that a sufficient jolt had not yet been given to the minds of the physicists of the time to shake these minds loose from the three entrammeling concepts. It was necessary for a fresh mind to come upon the scene. In late October 1895, W. C. Roentgen of a small and obscure university at Würzburg, Germany, began experimenting with a Crookes tube of the shape shown in Fig. 14.1. The anode *A* was tucked away in a side tube and the cathode *C* was a flat metal disc. The tube was evacuated to a pressure of about 0.001 mm of mercury. Upon sending a discharge through the tube the inside of the glass wall at *B* glowed with a greenish-yellow fluorescence. On Friday, November 8, 1895, Roentgen, after completely covering the Crookes tube with black cardboard and applying the high voltage (about 50,000 volts) from an induction coil to *K* and *A*, noticed that a sheet of barium platino-cyanide paper lying on a near-by table fluoresced whenever the voltage was applied to the Crookes tube. Roentgen observed this phenomenon in the earlier part of that memorable Friday. Realizing instantly that he had something new, something startling, he worked far into the night. He quickly found that obstacles placed between the tube and the fluorescent screen cast shadows. This fact convinced Roentgen that he was dealing with some kind of rays. Because of their unknown nature, Roentgen named them *x-rays*. He found that the rays penetrate all substances to a greater or less degree. In general, dense substances are more opaque to the rays than light substances. A

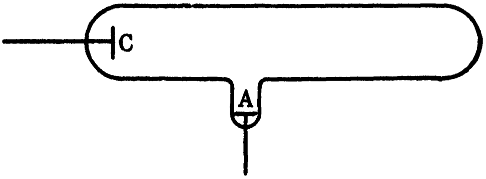


FIG. 14.1.—An early x-ray tube.

Crookes tube which emits x-rays began to be known as an x-ray tube. When the hand is interposed between an x-ray and a fluorescent screen, the rays penetrate the fleshy parts of the hand more easily than the bones, so that the bones stand out as shadows on the screen.

Roentgen was so surprised by his startling discovery that he delayed publication of his finding for nearly two months. He wanted to be quite sure. His main interests previous to his discovery of x-rays were in the physical properties of crystals and in the effect of pressure on the physical properties of various substances. Entering the field of cathode rays in October 1895, Roentgen stumbled on to the phenomenon of x-rays after about two weeks of work in this field. In the author's opinion, the discovery was highly accidental. But, accidental or not, Roentgen's discovery produced a tremendous stir among both the public and the physicists. X-ray photographs of the bones of the hand were exhibited publicly in Berlin on January 4, 1896. The idea that the bones of a *living* hand could be photographed amazed the public. The newspapers of the period carried headlines of this new marvel. Many ridiculous claims were made. The furor caused by the discovery of x-rays has only recently been equaled by the furor created by the announcement of the controlled release of atomic energy. All over the world physicists who possessed Crookes tubes and induction coils soon verified Roentgen's findings.

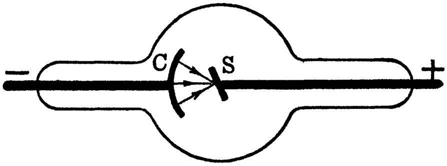


FIG. 14.2.—Gas x-ray tube.

14.2. Gas Tube.—The first x-ray tubes were Crookes tubes without a target. It was early found that the x-rays proceeded from that part of the glass wall of the Crookes tube which was hit by the cathode rays and which fluoresced with a greenish-yellow color. Many physicists, including J. J. Thomson, at first believed that the emission of x-rays was connected with the fluorescence of the glass. However, by June 1896, Crookes tubes with a concave cathode such as *C* in Fig. 14.2 and a target *S* began to come into use. As we have noted in Chapter VII, the cathode rays leave the cathode approximately at right angles to its surface. In Fig. 14.2, the cathode rays are focused at a small spot on the metal target. This spot is the center of emission for the x-rays. A metal does not fluoresce and so x-rays are not a consequence of fluorescence. The spot on the target is known as the *focal spot.* X-rays from a Crookes tube similar to that of Fig. 14.1 do not give distinct shadows on a fluorescent screen or a photographic plate because of the umbra-penumbra effect. The rays from an x-ray tube with a small focal spot give shadows with distinct outlines. The x-ray tube of Fig. 14.2 is exhausted to a pressure between 0.01 and 0.001 mm of mercury and then sealed off. The tube operates because although the pressure is low, there is still some gas in the tube. Such a tube is now known as a gas tube.

As a gas tube is operated, the gas within the tube is absorbed in the walls of the tube and the pressure of the gas diminishes. A greater voltage is therefore necessary to force a discharge through the tube and the tube is said to become harder. In general, the harder the x-ray tube the more penetrating are the rays emitted by the tube. It has become the practice to say that a hard tube emits hard rays, so that very penetrating rays are called hard rays. A tube which emits rays that are not so penetrating is called a soft tube and the rays are called soft rays. The voltage across an x-ray tube is usually of the order of 50,000 to 200,000 volts.

14.3. Coolidge Tube.—From 1896 to 1913, x-ray tubes were of the type shown in Fig. 14.2. However, all gas tubes harden on use. Various devices were made to enable the operator to soften the tube but these devices were cumbersome and not very successful. In 1913 Coolidge, of the General Electric Laboratories at Schenectady, introduced a new type of x-ray tube which has displaced the gas tube except for very special purposes.

In this type of tube the vacuum is made as high as possible, the pressure being of the order of 1/10,000 to 1/100,000 mm. There is so little gas in such a tube that no ordinary x-ray voltage can force a current through the tube. Use is made of the Richardson thermionic effect to obtain a supply of electrons. The concave cathode of the gas tube is replaced by a filament *F* as shown in Fig. 14.3. Opposite the filament is placed the target *T*. Two voltages are applied to the tube—a low voltage of some 10 or 12 volts for lighting the filament and a high voltage of some 50,000 or 100,000 volts between the filament and the target. Electrons evaporating from the filament find themselves in a very intense electric field, which throws them against the target with a velocity v given by

$$m_0c^2\left\{\frac{1}{\sqrt{1-\beta^2}}-1\right\}=Ve \tag{14-1}$$

whence, on putting in values, $\beta = .41$, or $v = \beta c = .41 \times 3 \times 10^{10} = 1.23 \times 10^{10}$ cm/sec. The relativity formula for kinetic energy is used because v is of the same order as c, the velocity of light.

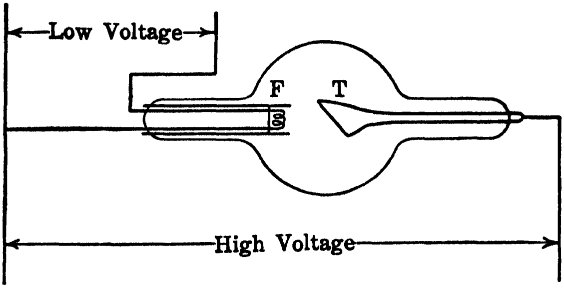


FIG. 14.3.—Coolidge x-ray tube.

The filament in a Coolidge tube is placed in a metal cup. The effect of this cup is to focus the electrons on a certain spot of the target, which is known as the *focal spot.* This is somewhat circular in shape and has a diameter ranging in various tubes from 2 to 5 mm. This focal spot is the origin of most of the x-rays which leave the tube. However, there are stray electrons which strike the target at other places than at the focal spot with the result that every point on the whole surface of the target is to some extent a source of x-rays. A pinhole picture of a Coolidge x-ray tube can be made by placing a plate of lead in which a pinhole has been punched between the tube and a photographic film enclosed in a light-tight envelope. An exposure of a minute or so will give a picture showing the whole target. A much shorter exposure gives a picture showing only the focal spot, which on close examination will be seen to have a spiral structure corresponding to the spiral filament, from which the electrons have evaporated. One disadvantage of the Coolidge tube as compared with the gas tube is the lack of focusing of the electrons at a point on the target. The x-rays from a Coolidge tube do not cast such well-defined shadows as those from a gas tube.

14.4. Transformer Connections.—The high voltage for an x-ray tube is usually supplied by a transformer. There are a large number of turns in the secondary and a small number of turns in the primary of the transformer. The primary is supplied with an alternating e.m.f. which can be varied from about 50 to 220 volts. Since the root mean square e.m.f. developed in the secondary is, say, 50,000 volts, great care must be taken to insulate the secondary from the primary. Accordingly, the transformer is immersed in oil. The middle point of the secondary is usually grounded as shown in Fig. 14.4, so that, if the maximum instantaneous e.m.f. developed in the secondary is 70,000 volts, the maximum difference of potential between a terminal of the secondary and ground is (½) × 70,000 or 35,000 volts.

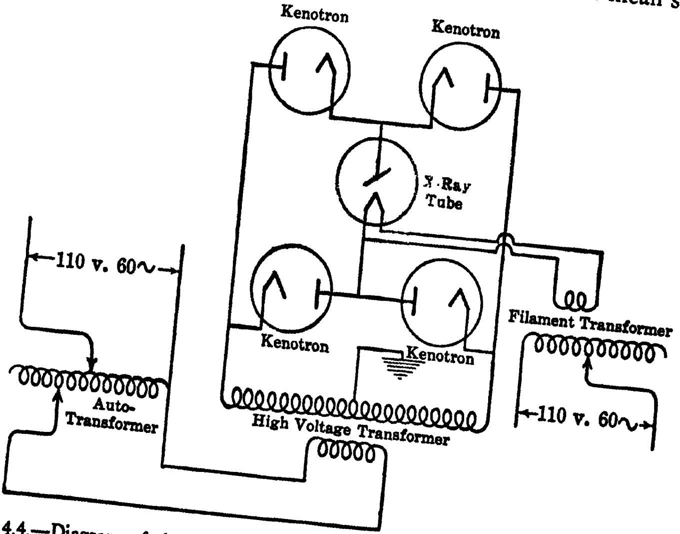


Fig. 14.4.—Diagram of the connections to an x-ray tube, full wave rectification being produced by means of four kenotrons.

The voltage developed in the secondary is alternating. If the filament of an x-ray tube is lighted and the terminals of the secondary are connected to the filament and target of an x-ray tube respectively, the tube will only operate during that half of a cycle when the filament is negative with respect to the target. During the next half of a cycle the tube will not produce x-rays unless the target becomes so hot that it begins to emit electrons. A Coolidge tube thus may act as a rectifier allowing current to pass in one direction but not in the other. However, it is found in practice that the life of a tube is lengthened by supplying the tube with high voltage which has already been rectified. One way of producing this rectification is by means of a kenotron as explained in Sec. 10.7. A single kenotron merely protects the x-ray tube against the high reverse voltage when the tube is not taking any current. Full-wave rectification can be obtained by means of four kenotrons as described in Sec. 10.7 and as shown in Fig. 14.4 above.

The voltage developed in the secondary of the high-voltage transformer can be varied by changing the voltage supplied to the primary of the high-voltage transformer. The primary is connected to an autotransformer as shown in Fig. 14.4. The characteristic feature of an *autotransformer* is that the primary and secondary coils are not separate from each other. Instead, there is only one coil, the primary current being supplied to a certain number of turns and the secondary current being taken from a different number of turns as shown in Fig. 14.4. The ratio of the secondary to the primary e.m.f. is determined by the ratio of the number of turns used by the secondary current to the number used by the primary current. If, therefore, the primary current is supplied through sliding contacts as shown, the ratio of the number of turns in the secondary to the number in the primary can be varied. Since the primary is supplied with 110-volt 60-cycle a.c., the secondary e.m.f. can be varied by moving the contact points. This secondary e.m.f. becomes the primary e.m.f. of the high-voltage transformer, and so the secondary e.m.f. of the high-voltage transformer can be varied by the operator.

We have supposed the filament to be lighted. The current which heats the filament is usually supplied by a step-down transformer, which is called the filament transformer to distinguish it from the high-voltage transformer. The primary of the filament transformer is usually supplied with 110-volt 60-cycle a.c., about 10 or 12-volt 60-cycle a.c. being developed in the secondary. However, it is seen by reference to Fig. 14.4 that the filament is also connected to the negative end of the high voltage. Hence, although the voltage developed in the secondary of the filament transformer may only be 10 or 12 volts, yet the difference of potential between the primary and secondary of the filament transformer may be as high as 35,000 volts if the potential across the x-ray tube is 50,000 volts r.m.s. For this reason, great care must be taken in insulating the secondary from the primary. Since the current from the secondary of the filament transformer is used for heating the filament and since the heating effect of a current does not depend on its direction, there is no need to rectify the filament current. The voltage across the filament can be varied by changing the ratio of the numbers of turns in the primary and secondary of the filament transformer. The temperature of the filament and therefore also the electron current through the tube can thus be controlled by the operator.

The great advantage of the Coolidge tube is that the high voltage across the tube and the electron current through the tube are under the control of the operator. He can, at will, make the tube emit harder or softer rays and can control the intensity of the x-rays coming from the tube.

14.5. Ionization of Gases. - In his first paper on x-rays, Roentgen announced that x-rays caused a fluorescent screen to light up and a photographic plate to blacken when subsequently developed. He mentioned nothing about the power of x-rays to impart electrical conductivity to air and other gases. In early 1896 J. J. Thomson and his associates at Cambridge, England were among the first to observe this phenomenon. Under ordinary circumstances air and other gases are insulators. Thomson soon showed that the conduction of electricity through gases is carried on by a mechanism similar to that in liquid electrolytes. The gas molecules are ionized by the x-rays—that is, some molecules become charged positively and others negatively. These charged molecules collect clusters of neutral molecules about them. These clusters constitute what are known as positive and negative ions.

X-rays pass through a thin aluminum window *W* into a hollow brass cylinder *AB* as shown in Fig. 14.5. A brass rod *CD* is supported by a rod *EF* which is held in place by insulating material as shown. The rod *EF* is connected to an electroscope *H*. The outside of the chamber *AB* is connected to one terminal of a battery, the other terminal of which is connected to ground. The chamber *AB* is filled with a gas usually at atmospheric pressure. If no x-rays are passing through the gas in the chamber no current will flow through the gas. If, however, x-rays are passing through the gas, the gas becomes ionized. With the connections as shown in Fig. 14.5 the metal walls of the chamber are at a positive potential relative to the inner electrode *CD*. Positive ions are therefore drawn to the inner electrode *CD* and negative ions to the walls of *AB* and a current thus passes through the gas. A positive charge thus begins to accumulate on the electrode *CD*, the electroscope *H* and the connecting wires between *CD* and *H*. The leaf *L* of the electroscope is repelled from the rod *R*. The leaf is observed by means of a microscope containing a scale in its focal plane. As the electroscope becomes charged, the leaf is seen to move, and either the time for a given deflection or the deflection in a given time is determined. Since the deflection of an electroscope is usually proportional neither to the charge on it nor to the difference of potential between the leaf and the metal case of the electroscope, it is usual to determine the time of a given deflection. The ionization current passing through *AB* is then inversely proportional to the time of this given deflection.

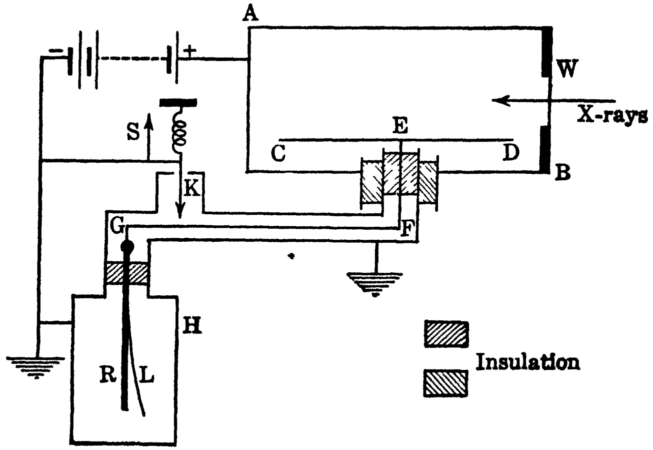


FIG. 14.5.—Ionization chamber and connections.

The chamber *AB* is called an ionization chamber. The ionization current is usually of the order of 10^{-13} amp. It is measured by means of an electroscope or electrometer instead of by a galvanometer because the current is too small to be measured by the latter type of instrument. The rate of deflection of the electroscope for a given ionization current varies inversely as the capacity of the system consisting of the electroscope, the connecting wires and the electrode *CD*. A not unusual value of this capacity is 100 cm. (Note: the centimeter is the unit of capacity in esu and is equal to 1.11 micro-microfarads.) An ordinary sensitivity of the electrometer is 5000 divisions per volt, so that if the ionization current is 10^{-13} amp, the rate of deflection of the electrometer is 4.5 divisions per second, which is very easily observable.

Because an electrostatic method is used for measuring the ionization current, it is necessary to have the wires connecting the inner electrode *CD* of the ionization chamber to the electroscope electrostatically shielded. This is shown in Fig. 14.5. The connecting wires are surrounded by metal tubes which are grounded. The outside case of the electroscope is also grounded. This grounded tubing pierces the insulating plug in the ionization chamber as shown in Fig. 14.5. The walls of the chamber are usually at a potential of 50 to 100 volts. Initially the inner electrode *CD* is grounded through the key *K*. This is held pressed against *GF* by means of a spring as shown. The ionization current then flows to ground. When it is desired to take a reading the grounding key is raised by means of the string *S* and the ionization current instead of flowing to ground now flows into the electroscope and begins charging it, so that its potential increases at a rate of 9×10^{-4} volt/sec in the example given above. When the given deflection of the electroscope has been obtained the grounding key *K* is let down into contact with *GF*. The electroscope leaf then comes back to zero and by raising the grounding key another reading can be taken.

The grounded tubing pierces the insulating plug in order that any leakage across the insulation may go to ground and not to the rod *EF*.

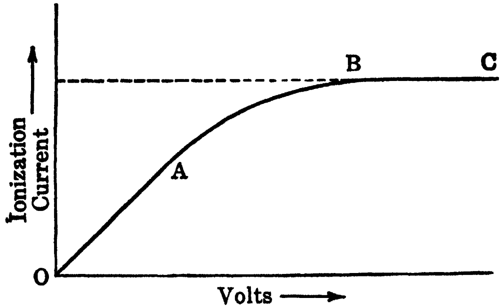


FIG. 14.6.—Variation of the current through an ionization chamber with the voltage across the chamber.

14.6. Ionization Current vs. Voltage Curve.—If the outside of the ionization chamber *AB*, Fig. 14.5 is connected to a battery of variable voltage and the ionization current for each voltage is obtained, it is found that the ionization current varies with the voltage in a manner shown by the graph in Fig. 14.6. The graph consists of three portions. The first portion *OA* is straight, the second portion *AB* is curved and the third portion is parallel to the voltage axis. Along *OA* the current is proportional to the voltage across the chamber so that Ohm's law is obeyed. Along *AB* the current does not increase with voltage as rapidly as it should according to Ohm's law. With further increase of voltage the current departs more and more from the value required by Ohm's law until we come to the portion *BC* where the current remains constant with increase of voltage so long as the voltage does not approach the sparking potential. In the stage represented by *BC* the current is said to be saturated. The least value of the voltage necessary to give the saturation current is called the saturation voltage. The curve *OABC* is the characteristic for the ionization chamber considered as a circuit element in the sense of Sec. 12.1.

Since the ions which have been produced by the x-rays are charged positively and negatively, they tend to recombine. The speed with which the ions are removed from the gas increases with increase of voltage across the chamber, so that at small voltages the ions are removed so slowly that considerable recombination of the ions takes place and so the number of positive ions collected per second by the inner electrode is small and consequently a small rate of deflection is registered by the electroscope. As the voltage is increased, a smaller number of ions is lost by recombination, and so the current increases and we have the portion *OA*. However, there is only a certain number of ions being produced per second, and, as the number of positive ions collected per second by the inner electrode approaches the number of positive ions produced per second, the current increases less rapidly with the voltage, thus giving the portion *AB* of the graph. Finally, every positive ion which is produced is collected by the inner electrode and every negative ion is collected by the outer electrode (the walls) of the chamber. When this stage arrives, an increase of voltage can cause no increase in the rate of collection of ions and therefore no increase in the ionization current.

The rate of production of ions depends only on the x-rays entering the ionization chamber. For x-rays of a given hardness the rate of production of ions is proportional to the intensity of the x-rays. The intensities of two beams of x-rays of the same hardness is therefore proportional to the rates of production of ions produced by the respective beams of x-rays. Since the rate of ion production is proportional to the saturation current, the intensities of two beams of x-rays of the same hardness can be compared by observing the saturation currents produced in an ionization chamber first by one beam and then by the other.

14.7. Absorption of X-Rays.—If x-rays from an x-ray tube *A*, which is enclosed in a lead box *B* as in Fig. 14.7, are passed through the apertures *C* and *D* in lead plates, a definite beam of x-rays is obtained. Let this beam pass into an ionization chamber *E*. Also let the voltage of the outside of the chamber be above saturation, so that the current registered by the electroscope (not shown in the figure) is the saturation current. The ionization current then gives a measure of the intensity of the x-rays.

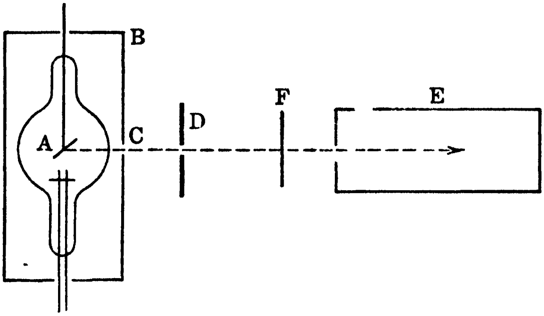


Fig. 14.7.—Arrangement for the measurement of the absorption of X-ray.

Let this current be represented by I_0. Now place a sheet of some material in the beam as at *F*. It is then found that the ionization current is reduced to *I*. If the thickness of the sheet is very small the reduction in the intensity, $I_0 - I$, is found to be proportional to the thickness. Since both the thickness and the reduction of intensity are small, we shall represent them by *dt* and *dI* respectively. Further, it is found that the reduction *dI* is proportional to the original intensity *I*. Hence, since *dI* varies as *Idt*, we have

$$dI = -\mu I dt \tag{14–2}$$

where μ is a constant of proportionality and the negative sign is inserted because *dI* is a reduction and not an increase of intensity. Rearranging (14–2), we obtain

$$\frac{dI}{I} = -\mu dt \tag{14–3}$$

Integrating both sides, we have

$$\log I = -\mu t + \text{const} \tag{14–4}$$

If we suppose that the intensity is I_0 when $t = 0$, then the constant in (14–4) equals $\log I_0$. Taking $\log I_0$ over to the left side and combining the two logarithms we obtain

$$\log\left(\frac{I}{I_0}\right) = -\mu t \tag{14–5}$$

or

$$I = I_0 \exp(-\mu t) \tag{14–6}$$

Eq. (14–6) gives us the relation between I_0, the intensity without the sheet F, Fig. 14.7, and I with the sheet F when the thickness of the sheet is not small. The quantity μ is known as the absorption coefficient. The absorption coefficient of a given material, such as aluminum, can be found by determining the thickness which will reduce the intensity of the beam penetrating through the sheet to one-half the original intensity. In this case $I/I_0 = 0.5$ and the value of x which makes $\exp(-x) = 0.5$ is found by reference to tables to be 0.693, so that

$$\mu = \frac{0.693}{t_{1/2}} \tag{14–7}$$

where $t_{1/2}$ represents the thickness necessary to reduce the intensity to one-half.

The hardness of x-rays is determined by their absorption coefficient in some substance, usually aluminum. The rays are harder when the absorption coefficient is smaller. It is found that, if the intensity of the x-rays from a tube is reduced to one-half by passing through aluminum and then these rays are passed through more aluminum and the intensity again reduced to one-half so that altogether the intensity has been reduced to one-quarter of the original intensity, the thickness of aluminum necessary to produce the second reduction of one-half is greater than that necessary to produce the first reduction of one-half. The absorption coefficient therefore decreases as the rays are passed through more and more aluminum. Another way of stating the same thing is to say that the hardness of the penetrating rays increases with the thickness of aluminum which the rays have penetrated. The explanation of this increase of hardness is that the rays from an x-ray tube are not homogeneous but heterogeneous. The rays from a tube consist of both hard and soft rays. On passing through an absorbing sheet, the softer components are filtered out, leaving the harder components to penetrate through. The average hardness of the penetrating beam thus increases with the thickness penetrated.

We shall see later that it is possible to obtain homogeneous x-rays. When homogeneous rays are sent through absorbing sheets of a given material, the absorption coefficient μ is a constant and is independent of the thickness t. In this case the relation (14–6) holds exactly. For heterogeneous rays, relation (14–6) is only approximate since μ is not a constant but decreases with increase of t.

✓ For rays of a given hardness it is found that the absorption coefficient in a given chemical substance varies as the density of the substance. In the case of a substance such as water, which can exist either as solid, liquid, or gas, the density can vary over a wide range. In such a substance μ varies as ρ, the density of the substance, so that μ/ρ is a constant. For rays of a given hardness, the quantity μ/ρ has a value which is characteristic of the substance and is independent of the state of the substance. Consequently, μ ρ has been given a special name and it is called the ✓ mass absorption coefficient. In general the mass absorption coefficient of an elementary substance for x-rays of a given hardness increases rapidly with the atomic weight of the substance. The mass absorption coefficient of lead (atomic weight 207) is therefore very much greater than that of aluminum (atomic weight 27). This explains why lead is used to protect the operator of an x-ray tube from the harmful effects of x-rays. We shall return to the further consideration of absorption of x-rays in a later section.

14.8. Scattering of X-Rays.—If a beam of x-rays is allowed to fall upon a sheet of paraffin as shown at B, Fig. 14.8, it is found that the part of the paraffin which is illuminated by the rays becomes itself a source of x-rays and these x-rays proceed outwards in all directions. These rays are called scattered rays, and the phenomenon is known as scattering. The scattered rays are emitted not only by the surface of the scattering sheet, but by points throughout the body of the sheet which are illuminated by x-rays coming from the tube. The rays coming directly from the tube are called primary rays to distinguish them from the scattered rays. The scattered rays are of approximately the same hardness as the primary rays.

If the thickness of the scattering sheet B, Fig. 14.8 is small, the intensity of the scattered rays entering the ionization chamber is proportional to the area A of the window of the chamber, to the thickness dt of the scattering sheet in the direction of the primary rays, and to the intensity I_P of the primary x-rays from the tube. The intensity of the scattered rays entering the ionization chamber also varies inversely as the square of the distance of the chamber window from the portion of the scattering sheet which is illuminated by the primary x-rays. Hence, if dI_ϕ represents the intensity of the scattered rays entering the chamber when it is set so as to receive scattered rays making an angle ϕ with the primary beam,

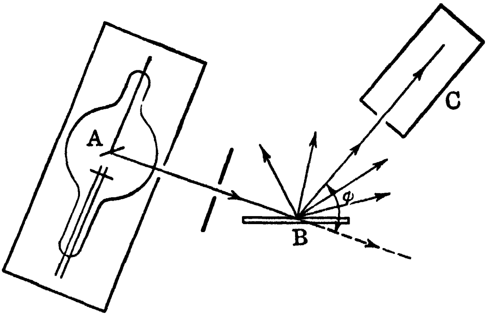


FIG. 14.8.—Experimental arrangement for measuring the scattering of x-rays.

$$dI_\phi = \frac{S_\phi I_P A dt}{R^2} \tag{14–8}$$

where S_ϕ is a constant of proportionality and is known as the spatial scattering coefficient. The subscript ϕ is added to S in order to indicate that the spatial scattering coefficient is a function of ϕ, the angle of scattering. It has been found that for scattering substances of small atomic weight—that is, atomic weight below about 30—the spatial scattering coefficient varies approximately as $1 + \cos^2 \phi$.

If the ionization chamber is set at different angles of scattering, it is possible to determine the intensity of the total scattering in all directions or, in other words, the intensity of the scattered rays crossing a sphere surrounding the scattering substance. This total scattering involves what is known as the spherical scattering coefficient. If dI_S is the total intensity of the rays scattered over a sphere,

$$dI_S = \sigma I_P dt \tag{14–9}$$

where σ is a constant of proportionality and is known as the spherical scattering coefficient.

As in the case of the absorption coefficient, the spatial and spherical scattering coefficients of a substance are divided by its density in order to obtain the spatial and spherical mass scattering coefficients respectively. It is found that the spherical mass scattering coefficient σ/ρ for x-rays of moderate hardness has approximately the constant value 0.2 for elements whose atomic weights are below about 30. There is one exception to this rule and that is hydrogen, for which σ/ρ has approximately the value 0.4.

14.9. Polarization of X-Rays.—A beam of x-rays, traveling parallel to *OX* as in Fig. 14.9, strikes a scattering sheet at *Z*. This sheet scatters x-rays in all directions. By means of a system of apertures as shown in Fig. 14.9 a beam in the direction *ZO* which is perpendicular to *OX* is isolated. This beam then strikes a second scattering sheet at *O* and this second sheet also scatters x-rays. If ionization chambers are placed to receive x-rays in the directions *OX* and *OY*, respectively, it is found that the intensity of the rays scattered in the direction *OX* is a maximum while the intensity in the direction *OY*, which is perpendicular to both *OX* and *OZ*, is very nearly zero. The beam *ZO* is therefore said to be polarized. Primary rays from an x-ray tube are only partially polarized, but x-rays which have been scattered at right angles are almost completely polarized. In order to detect the polarization these once-scattered rays must be scattered again. The scatterer at *Z*, Fig. 14.9 may be thought of as the polarizer and the scatterer at *O* as the analyzer. The polarization of x-rays was discovered by Barkla in 1904 at Cambridge, England.

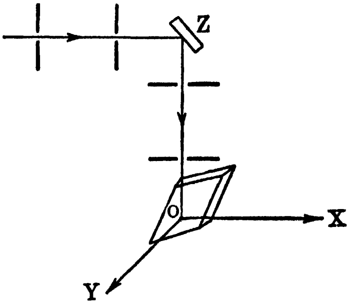


FIG. 14.9.—Experimental arrangement for showing the polarization of x-rays.

14.10. Characteristic X-Rays.—In 1908 Barkla and Sadler announced the discovery of secondary x-rays which differ from the scattered rays which we have described in Sec. 14.8 in that they are distinctly softer than the primary x-rays which produce them. If primary x-rays from an x-ray tube are allowed to fall upon a copper sheet *A*, Fig. 14.10, the sheet is found to be the source of x-rays which are much more intense than the scattered rays previously described. If these secondary rays are allowed to enter an ionization chamber *B* as in Fig. 14.10, their absorption coefficient in aluminum can be measured. This coefficient is considerably greater than that of the primary rays in aluminum, so that the secondary rays are softer than the primary rays. Furthermore, it is found that the absorption coefficient in aluminum is very nearly independent of the thickness of the aluminum absorbing sheet, so that the secondary rays, according to Sec. 14.7, are very nearly homogeneous. It was at first thought that these secondary rays were exactly homogeneous. They have a definite absorption coefficient which is characteristic of the material at *A*. Thus, if the material at *A* is copper, the mass absorption coefficient of the secondary rays in aluminum is 47.7, while, if the material at *A* is silver, the mass absorption coefficient of the secondary rays is 2.5. Since these secondary rays are characteristic of the element emitting them, the name characteristic rays has been given to the rays.

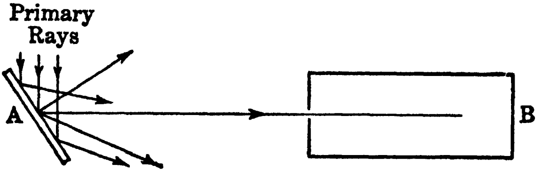


FIG. 14.10.

The characteristic rays of the various elements increase in hardness as the atomic weight of the element increases. If copper x-rays, as the characteristic rays from copper are called, are allowed to fall on a sheet of iron, the characteristic rays of iron are excited. On the other hand, if iron x-rays are allowed to fall on copper, the characteristic rays of copper are not excited. Hence, since copper (63.6) has a higher atomic weight than iron (55.8), it is seen that the harder rays can excite the softer but the softer cannot excite the harder rays.

Barkla found that elements of high atomic weight such as lead (207) when irradiated by x-rays emit characteristic rays which are soft. Further research showed that elements emit both hard and soft rays each of which is characteristic of the element. Barkla designated these different rays by the letters *K* and *L*. Silver *K* x-rays are hard and silver *L* x-rays are soft. The difference between the mass absorption coefficients in the case of *K* and *L* x-rays is very great. In the case of tungsten, μ/ρ in aluminum for the *K* rays is 0.27, while μ/ρ for the *L* rays is more than one hundred times as great, being about 36. The softer *L* rays are easily filtered out by thin aluminum sheets. In the case of copper rays, the rays which are observed are the *K* rays because the *L* rays are absorbed in the air itself and do not enter the ionization chamber.

Barkla stated that only elements of atomic weight 40 and over could be caused to emit characteristic rays. We now know that elements of atomic weights lower than 40 can be caused to emit their characteristic rays. However, these rays are so soft that they are absorbed in the air and they are consequently difficult to observe. In addition to the characteristic rays emitted by all elements, the scattered rays which we have described in Sec. 14.8 are also emitted. The difference between scattered and characteristic x-rays is that the hardness of the former is determined by the hardness of the primary rays, while the hardness of the latter is characteristic of the substance emitting them. Although a substance such as copper gives out both scattered and characteristic rays when irradiated with x-rays, the characteristic rays are so much more intense than the scattered rays that the scattered rays are masked unless special means are taken to observe them.

14.11. Secondary Cathode Rays.—When x-rays fall on a copper sheet, it not only emits scattered and characteristic x-rays but also negatively charged particles, which are called secondary cathode rays. That these particles are electrons can be shown by measuring their electrostatic and magnetic deflections by methods similar to those described in previous chapters. The velocity of the electrons is found to be of all values between zero and a maximum. This variation is due to the fact that the x-rays penetrate the copper sheet and cause the ejection of electrons from different depths and the ejected electrons therefore lose different amounts of energy in penetrating through to the surface of the metal. The maximum velocity is of the order of 50,000 to 100,000 volts. It is a very curious fact that the maximum velocity of the secondary cathode rays is equal to that of the cathode rays which hit the target in the x-ray tube.

14.12. Theories of X-Rays.—Roentgen, in his paper of December 1895, hazarded the opinion that x-rays were longitudinal vibrations in the ether. Towards the latter part of 1896, Stokes of England suggested that x-rays were produced by the sudden stopping of the stream of negatively charged molecules when these molecules hit the target of an x-ray tube. But the electron had not been recognized by the end of 1896. The English school of physicists still believed that cathode rays were charged molecules. After the recognition of the electron in 1897, Stokes' idea became fruitful. According to the Stokes theory the electrons on hitting the target of an x-ray tube undergo very great negative acceleration. We have stated in Sec. 6.4 that an accelerated electron emits electromagnetic radiation. According to Stokes' theory, this radiation proceeds outwards in the form of a pulse in the ether. J. J. Thomson made use of this pulse theory to explain the process of the scattering of x-rays. The pulse proceeds outward with the speed of light. Within the pulse the electric field strength is very strong. As the pulse passes over an atom, the electrons constituting the atom are suddenly acted upon by a strong electric field and are given a sudden jerk. This sudden jerk causes the electrons to emit radiation as stated in Sec. 6.4. Hence the electrons over which an x-ray pulse is passing become themselves emitters of x-rays. Thomson calculated the value of the spherical mass scattering coefficient and obtained

$$\frac{\sigma}{\rho} = \frac{8\pi NZe'^4}{3Wm^2c^4} \tag{14–10}$$

where N is Avogadro's number, Z is the number of electrons in an atom of the scatterer, W is the atomic weight of the scatterer, e' and m are the charge and mass of the electron respectively, and c is the velocity of light. Putting in known values, we find that (14–10) reduces to

$$\frac{\sigma}{\rho} = 0.402\left(\frac{Z}{W}\right) \tag{14–11}$$

Since Barkla obtained $\sigma/\rho = .2$ for light elements other than hydrogen, $Z = W/2$ approximately. The number of electrons in an atom is thus approximately equal to half the atomic weight of the atom except in the case of the hydrogen atom which contains one electron.

However, we have noted in Sec. 14.11 that x-rays have the power of ejecting electrons from substances. This power is similar to the photoelectric effect of light. Electrons are only ejected from certain atoms although an x-ray pulse passes over all atoms. Further, the ionization of gases by x-rays is explained by this power of x-rays to eject electrons from atoms. Although an ether pulse passes over every atom in the path of an x-ray beam, yet only an exceedingly small fraction of the atoms become ionized. Hence, an electromagnetic pulse theory of x-rays encounters similar difficulties to those encountered by the wave theory of light in attempting to explain the photoelectric effect. Accordingly,

W. H. Bragg of England was led to propose a corpuscular theory of x-rays. He supposed the target to emit x-ray corpuscles which in turn knocked electrons out of the atoms of a gas and so ionized the gas.

The situation at the beginning of 1912 was that the ether pulse theory explained the scattering of x-rays and Bragg's corpuscular theory explained the ionization of gases by x-rays. There was considerable controversy as to the nature of x-rays. Are they a form of wave motion in the ether or are they some kind of corpuscles? In the latter part of 1912 an experiment by Laue, Friedrich, and Knipping in Germany showed conclusively at that time that x-rays are ether waves. This experiment is described in the next section.

14.13. Diffraction of X-Rays.—In the case of light, it is possible to explain all classes of phenomena except one by means of a corpuscular theory. Usually, polarization of light is explained by the wave-theory, but even this phenomenon can be explained by endowing the corpuscles of a corpuscular theory with the appropriate properties. The one class of phenomena which persistently refuses to be explained by a corpuscular theory is that class of phenomena which is connected with interference. When two beams of light, originating from the same source, are brought together, the result is darkness or brightness according as the difference of the two paths from the source is an odd or even number of half wave-lengths. A particular application of the phenomenon of interference is made in the diffraction grating. Parallel light falling normally on a grating as shown in Fig. 14.11 is diffracted in the direction θ. The various parallel beams pass through the lens L and are brought to a focus at P. The optical difference of path between the beams AHP and BKP is AG. If λ is the wave-length of the light and n is a whole number, there is brightness at P when

$$n\lambda = d \sin \theta \tag{14–12}$$

where $d = AB$, the distance between consecutive slits in the diffraction grating. The portion of the primary wave at A is in phase with the portions of the wave at B and C. Even if the light does not fall normally on the diffraction grating and the portions of the waves at A, B and C are not in phase with each other, yet there is a definite relation between the phase of the portion of the wave at A and that of the portion at B. Similarly there is a relation between the phases at A and C and between the phases at B and C. This definite phase relationship in the wave theory is called coherence.

Attempts have been made to explain interference by supposing the corpuscles of a corpuscular theory to vibrate or pulsate in some way, but there is no reason why the pulsating corpuscle as it passes through the slit at *A*, Fig. 14.11 should be in phase with a second pulsating corpuscle, which passes simultaneously through the slit at *B*. There is no relation between the phases of the pulsating corpuscles. It is this lack of coherence between the corpuscles which makes it impossible to explain interference phenomena by means of a corpuscular theory.

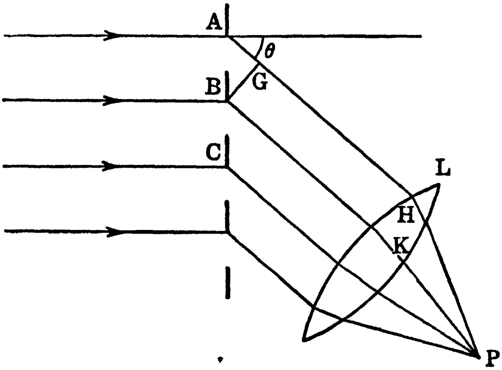


FIG. 14.11.—Optical diffraction grating.

The Einstein photoelectric equation

$$(\tfrac{1}{2})mv^2 = h\nu - W \tag{14-13}$$

which we derived in Sec. 11.7, shows that, if the frequency of the light is high, the velocity of the ejected photoelectron is great. If the work function W in (14-13) may be neglected, we have shown in Sec. 11.9 that (14-13) gives rise to the relation

$$\lambda \text{ (in angstroms)} = \frac{12345}{V \text{ (in volts)}} \tag{14-14}$$

We have seen in Sec. 14.11 that the maximum velocity of the secondary cathode rays produced by x-rays is of the order of 50,000 volts. Hence, if x-rays are of the same nature as light, their wave-length must be of the order of 0.25 A. Ordinary visible light has a wave-length of the order of 5000 A. An optical diffraction grating is ruled with about 10,000 lines to the centimeter, so that it seemed to be necessary for an x-ray diffration grating to be ruled with about 20,000 times as many lines to the centimeter, making d in (14-12) equal to 1/200,000,000 or 0.5 ×

10^{-8} cm. No grating can be ruled with lines this close together. However, in 1912 it occurred to Laue that nature herself provides such a grating. Crystals are formed in their peculiar geometrical shapes because their atoms are arranged in a perfectly definite order. In the case of a cubic crystal the atoms are equally spaced in three mutually perpendicular directions. We shall see later that these spaces in rocksalt (NaCl) are equal to 2.82 A. If then each atom diffracts x-rays as the slits diffract light in a diffraction grating, a crystal like rocksalt should diffract x-rays. However, there is an important difference between an optical grating and a crystal grating. The slits in an optical grating are in one plane whereas the atoms of a crystal grating are arranged in space. The optical grating is thus a two-dimensional grating, while a crystal is a three-dimensional grating.

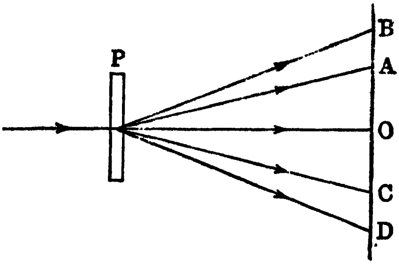


Fig. 14.12.

In the latter part of 1912, Laue, Friedrich, and Knipping announced the discovery of the diffraction of x-rays. They passed a narrow beam of x-rays through a thin crystal of zinc blende (ZnS) as shown in Fig. 14.12. The rays passed through the crystal *P* and on to a photographic plate *BD*. Most of the rays passed straight through the crystal *P* and formed a central black spot at *O* on the plate. Some rays were diffracted to make black spots at *A*, *B*, *C*, *D*. However, since the crystal grating is three dimensional, there are black spots formed on the plate at various points outside of the plane of Fig. 14.12. The spots on the plate are arranged according to a geometrical pattern, as shown in Fig. 14.13 for the case of potassium chloride. These spots are known as Laue spots. The particular pattern formed depends upon the particular crystalline system to which the crystal belongs.

Since the crystal constitutes a three-dimensional grating, the mathematics necessary to the development of a formula corresponding to

(14–12) is cumbersome and is beyond the scope of this book. Fortunately, however, W. L. Bragg, son of W. H. Bragg, showed that the Laue spots could be explained by supposing that reflection occurred from the various planes of atoms in a crystal. This approach to the problem simplifies the mathematics considerably and is given in the next section.

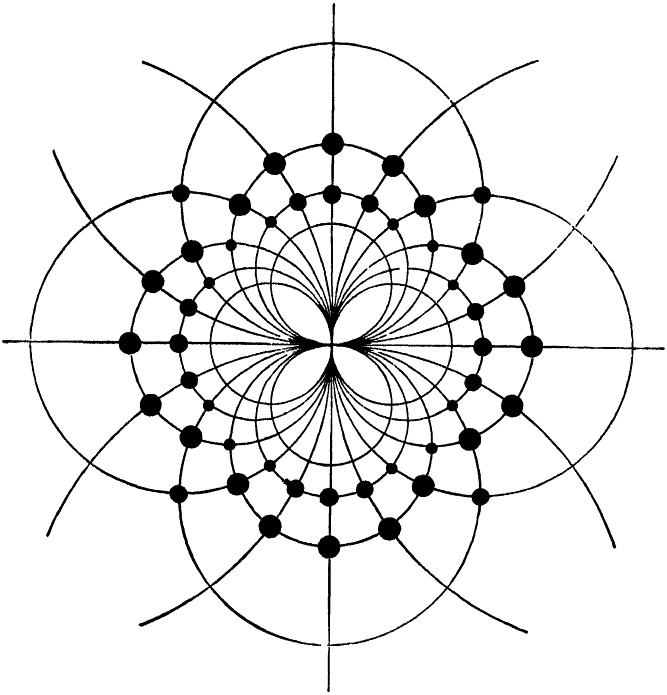


FIG. 14.13.—Laue spots for potassium chloride (KCl). The intensity of a spot is porportional to its size in the diagram.

14.14. Bragg's Law.—The atoms in a plane section of a crystal are represented in Fig. 14.14. Planes may be drawn as represented by the lines *AB*, *AC*, *DE*. Consider a beam of x-rays traveling in the vertically downward direction as shown in Fig. 14.14. At each plane of atoms there is, according to Bragg, partial reflection of x-rays, the reflected beams being shown in Fig. 14.14. It is these reflected beams which give rise to the Laue spots. The question then arises as to the conditions necessary for reflection of x-rays from a plane of atoms. Let *AB*, Fig. 14.15 represent one plane of atoms, *CD* the next parallel plane of atoms, and *EF* the next and so on. In the case of x-rays it is not only necessary to consider the rays reflected from the plane *AB* but also the rays reflected from the planes *CD*, *EF* and so on. Reflection of ordinary light takes place from the surface, but x-rays penetrate down into the substance, and so reflection of x-rays occurs at a series (or set) of planes. Because of this peculiar circumstance, it happens that, if x-rays are of the same nature as light, the rays reflected from successive planes must be in phase with each other. Hence, if the ray *LM* is reflected by the plane *AB* in the direction *MN* and the ray *OP* is reflected by the plane *CD* along *PQ*, the path difference of the two rays must be $n\lambda$, where λ is the wave-length of the x-rays and n is a whole number. Through *M* draw perpendiculars *MR* and *MS* to *OP* and *PQ*, respectively. *MR* then represents the trace of the wave front incident upon the plane *AB*, and *MS* the trace of the wave front reflected from the plane *AB*. The path difference between the wave front reflected from *AB* and that reflected from *CD* is $RP + PS$. Since the angle of reflection equals the angle of incidence, $PS = RP$. Hence, for the reflected

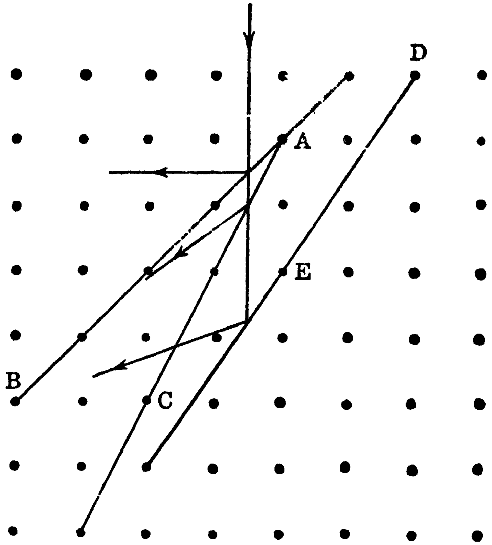


Fig. 14.14.—Reflection of x-rays from planes of atoms in a crystal.

waves from the two planes *AB* and *CD* to be in phase, we must have

$$n\lambda = 2 \times RP \qquad (14\text{–}15)$$

But $2 \times \overline{RP} = 2 \times \overline{PM} \times \sin \angle PMR$ and $\angle PMR = \angle LMA$, so that

$$n\lambda = 2d \sin \theta \qquad (14\text{–}16)$$

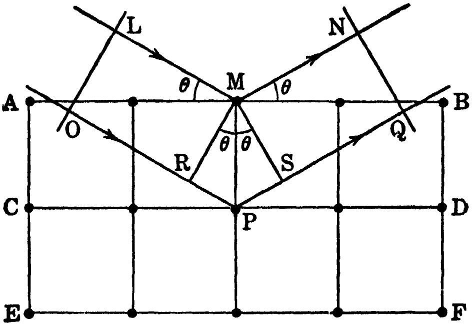


FIG. 14.15.—Reflection of x-rays from a crystal. Bragg's Law.

where $\theta = \angle LMA$ and d is the distance between consecutive planes of atoms. θ is the angle between the incident beam and the plane *AB* and is called the *glancing* or *grazing* angle of incidence. d is called the *grating space*. The relation expressed by (14–16) is called Bragg's law.

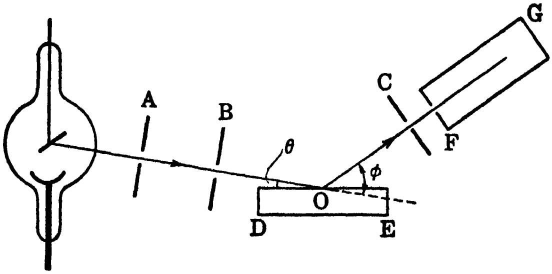


FIG. 14.16.—Bragg x-ray spectrometer.

In 1913, W. H. and W. L. Bragg set up an apparatus as represented in Fig. 14.16. A narrow beam of x-rays was obtained by means of the slits *A* and *B*. This beam fell upon a cleavage face of a crystal *DE*. The crystal *DE* could be rotated about an axis through *O*, and an ionization chamber *FG* was mounted upon an arm which also could be rotated about the axis through O. A slit C was mounted on the arm carrying the ionization chamber FG. With the ionization chamber set at an angle ϕ, the glancing angle θ could be varied by rotating the crystal. It was found that the ionization current showed a sharp maximum when $\theta = \phi/2$—that is, when the angle of incidence equals the angle of reflection.

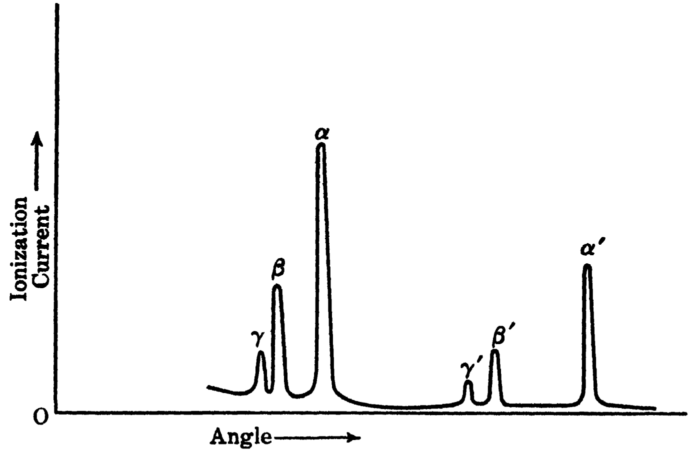


Fig. 14.17.—Curve showing the existence of x-ray spectrum lines.

14.15. X-Ray Spectra.—If an x-ray tube with a molybdenum target is used in an apparatus as shown in Fig. 14.16 and various settings of the angle θ are made while ϕ is made equal to 2θ for each setting of θ, it is found that the intensity of the reflected rays varies with θ according to the graph shown in Fig. 14.17. There are certain values of θ for which the reflected rays are very intense as shown by the sharp maxima, γ, β, and α in Fig. 14.17. If the crystal used is rocksalt, these peaks occur at values of θ equal to 6°20′, 6°27′, and 7°15′ for the γ, β, and α maxima respectively.

If Bragg's equation (14–16) is examined, it will be seen that x-ray reflection differs from optical reflection in that only the rays of appropriate wave-length are reflected at a given angle θ. If θ is varied, the wave-length of the reflected x-rays varies. Since the ionization chamber is always placed in a position to receive the reflected rays, the strong reflections shown at α, β, and γ in Fig. 14.17 indicate that x-rays of the wave-lengths corresponding to angles of 6°20′, 6°27′, and 7°15′ in Bragg's equation must be present in large quantities in the rays coming from the tube. It is found that the positions of the peaks in Fig. 14.17 depend on the metal used for the target of the tube. If the curve of Fig. 14.17 is followed to still greater angles, it is found that there are less prominent peaks α', β', and γ' at angles whose sines are double the sines of 6°20′, 6°27′, and 7°15′, respectively. It is seen from (14–16) that sin θ varies as n for a given wave-length. The peaks α, β, γ in Fig. 14.17 thus correspond to $n = 1$, while the peaks α', β', γ' correspond to $n = 2$. The first set of peaks is called the first order spectrum and the second set the second order spectrum. With sensitive apparatus, third and higher order spectra can be observed.

The Bragg method of obtaining x-ray spectra as above described is somewhat tedious. Dozens and perhaps hundreds of electrometer readings must be taken in order to plot a graph such as that in Fig. 14.17. Maurice De Broglie of France devised a photographic method which is automatic. In this method, the crystal is rocked to and fro through a range of angles covering 2 or 3 degrees of arc. If a molybdenum tube and a rocksalt crystal are used, the crystal is rotated from, say, $\theta = 5°$ to $\theta = 8°$ and then back again to $\theta = 5°$, the process being repeated many times. A photographic plate which is held stationary is used instead of the ionization chamber. As the crystal passes through an angle appropriate to the reflection of one of the spectrum lines, that line is strongly reflected on to the photographic plate and a dark line is produced. As the crystal passes through an angle appropriate to the reflection of another spectrum line, that line is strongly reflected and another dark line is produced on the photographic plate. The plate when developed has the appearance shown in Fig. 14.18. Knowing the distance of the photographic plate from the axis of rotation of the crystal and the position of the lines on the photographic plate, the values of θ corresponding to the different lines can be found.

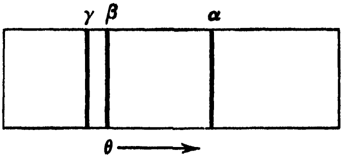


FIG. 14.18.—Typical K x-ray spectrum.

14.16. Miller Indices.—Before we can calculate the wave-length of x-rays we must determine d in Bragg's equation (14–16). However, the grating space d depends on the set of planes used in the crystal. A crystal of rocksalt for instance can be cleaved in various ways, and so different sets of planes for reflection can be obtained. It is necessary to devise some means of designating which set of planes is being used for reflection.

Consider a crystal of rocksalt as shown in Fig. 14.19. This crystal

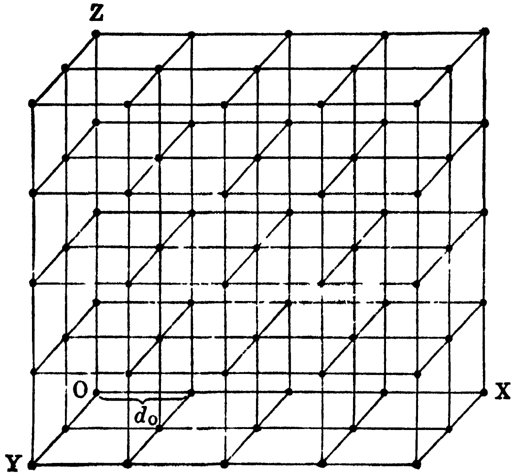


Fig. 14.19.

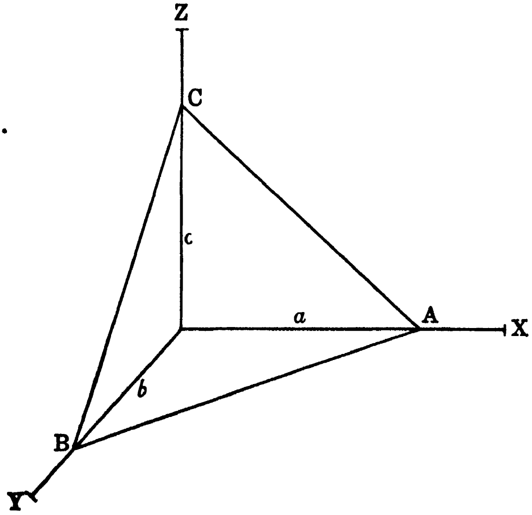


Fig. 14.20.

belongs to the cubic system. It has three mutually perpendicular principal axes *OX*, *OY*, and *OZ*. These axes are shown also in Fig. 14.20.

Imagine a plane ABC to cut these axes in A, B, and C respectively. Let the lengths of the intercepts OA, OB and OC be a, b, and c, respectively; then the Miller indices of the plane are $1/a$, $1/b$, $1/c$. It is usual to multiply the indices by some factor so as to obtain whole numbers. The Miller indices do not define a particular plane but a set of parallel planes. The indices are enclosed in parentheses, so that we speak of (1, 0, 0), (1, 1, 0), (1, 1, 1), (2, 1, 1), etc., planes. It is only the ratio of the indices which is of importance. For instance, the (4, 2, 2) planes are the same as the (2, 1, 1) planes. In Fig. 14.19 the distance between consecutive planes of atoms parallel to a principal plane such as XOY is called the principal grating space. This is represented by d_0 in the figure.

The (2, 2, 3) planes make intercepts along OX, OY, OZ, which are proportional to $\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{3}$, respectively, that is, proportional to 3, 3, 2, respectively. One of the (2, 2, 3) planes is therefore the plane which passes through the atoms at distances $3d_0$, $3d_0$, $2d_0$ along the axes OX, OY, OZ, respectively, from the origin O. The intercepts of the planes (1, 0, 0) are proportional to 1, ∞, ∞. Hence the (1, 0, 0) planes do not cut OY and OZ and so such planes are parallel to the principal plane YOZ. The grating space of the (1, 0, 0) planes is thus d_0.

14.17. Principal Grating Space.—The molecular weight of rocksalt (NaCl) is $23 + 35.5 = 58.5$. The number of NaCl molecules in 58.5 grams of rocksalt is 6.03×10^{23} (Avogadro's number). The density of rocksalt is 2.163 gm/cm^3 so that the number of molecules in 1 cm^3 is

$$\left(\frac{2.163}{58.5}\right) \times 6.03 \times 10^{23} = 22.3 \times 10^{21}$$

However, it is the separate atoms and not the molecules which are arranged in the order shown in Fig. 14.19. As one goes along a line parallel to a principal axis one encounters sodium and chlorine atoms alternately. Hence, the number of atoms in 1 cm^3 of rocksalt is 44.6×10^{21}. The number of atoms along an edge of a one centimeter cube is $\sqrt[3]{44.6 \times 10^{21}} = 3.55 \times 10^7$. Hence, the principal grating space of rocksalt is

$$d_0 = \frac{1}{3.55} \times 10^{-7} = 2.82 \times 10^{-8} \text{ cm}$$

Calcite ($CaCO_3$) is another crystal which is used in the reflection of x-rays. However, this crystal does not belong to the cubic system and the angles between the principal axes are not right angles. The calculation of the principal grating space is therefore more complicated. However, when these calculations are made, the grating space for calcite is found to be

$$d_0 = 3.03 \times 10^{-8} \text{ cm}$$

14.18. Atomic Weight and Atomic Number.—By reference to any textbook on chemistry it will be found that the elements can be arranged in what is known as the Periodic Table. The elements appearing in the same vertical column possess similar chemical properties. At first, the elements were arranged in this table in the order of increasing atomic weights. When this is done, it is found that every so often the chemical properties of an element are similar to those of a previous element. When this similarity occurs, the latter element is placed in the same vertical column as the previous element. Thus, sodium (At. Wt. 23) has similar chemical properties to lithium (At. Wt. 7). A portion of the periodic Table is shown in Table 14.1. In this table the first number to the right of the symbol for the element is the atomic weight. It will be seen that the elements are arranged in the order of increasing atomic weight until we come to argon (A) and potassium (K). However, argon is a gas similar to neon (Ne), while potassium is a metal similar to sodium (Na), and so the arrangement for these two elements is in the order of chemical properties rather than in the order of the atomic weights. There are other exceptions to the general rule that the elements are arranged in the order of increasing atomic weight, such an exception being shown by the elements iron (Fe), cobalt (Co), and nickel (Ni). The order of increasing atomic weights is Fe, Ni, Co, while the order according to chemical properties is Fe, Co, Ni. Until the end of 1940 the complete Periodic Table contained 92 places for elements.

TABLE 14.1

PORTION OF THE PERIODIC TABLE

H 1.008, 1							He 4, 2
Li 7, 3	Be 9, 4	B 11, 5	C 12, 6	N 14, 7	O 16, 8	F 19, 9	Ne 20, 10
Na 23, 11	Mg 24, 12	Al 27, 13	Si 28, 14	P 31, 15	S 32, 16	Cl 35.5, 17	A 40, 18
K 39, 19							

Since there are exceptions to the arrangement according to the order of atomic weights, another number known as the atomic number has come into use. This number is simply the number of the element as it appears in the table. Hydrogen is the 1st element, helium the 2d, nitrogen the 7th, argon the 18th, potassium the 19th and so on. The second number to the right of the symbol for an element in Table 14.1 is the atomic number. It will be noticed that with the exception of hydrogen the atomic number of an element is very nearly one-half the atomic weight. As remarked in Sec. 14.12, the experimental values of the mass scattering coefficient of x-rays in various light elements indicates that with the exception of hydrogen the number of electrons in an atom is approximately equal to half the atomic weight of the atom. According to modern theory the number of electrons in a neutral atom is exactly equal to the atomic number of the atom, and this is so even in the case of hydrogen.

14.19. Wave-Length of X-Rays.—The angles given for the molybdenum x-rays spectrum in Sec. 14.15 are for the case where the rays are reflected from the (1, 0, 0) planes of rocksalt, so that d in the Bragg equation is 2.82 A. Substituting in (14–16), the wave-lengths are .619, .631, and .710 A for the γ, β, and α lines, respectively. The x-ray spectra for several elements are shown in Table 14.2.

TABLE 14.2

K SPECTRA

Element	Atomic weight	Atomic number	Angstroms		
			α	β	γ
Fe.	55.8	26	1.934	1.752	1.741
Co.	58.9	27	1.787	1.617	1.605
Ni.	58.7	28	1.656	1.497	1.485
Cu.	63.6	29	1.539	1.382	1.378
Zr.	91.2	40	0.786	0.700	0.688
Mo.	96.0	42	0.710	0.631	0.619
Ag.	107.9	47	0.560	0.496	0.486
W.	184.0	74	0.211	0.184	0.179
Pb.	207.2	82	0.168	0.146	0.141

In all these spectra the α line is the strongest and the γ line the weakest. Further, the lines of all of the spectra are arranged in a manner similar to Fig. 14.21.

The separation in angstroms between the α and β lines is always much greater than that between β and γ lines. The x-ray spectra of all the elements cited are therefore similar. This is quite different to the situation in light, where the arrangement of the lines differs from element to element.

If the spectrum of the x-rays coming from a tube with a tungsten target is carefully examined, it is found that besides the wave-lengths shown in Table 14.2 there is a second spectrum whose lines have wave-lengths which are much longer than the wave-lengths of the tungsten (W) spectrum shown in the table. More careful examination of the x-rays coming from targets of other elements shows the existence of this second spectrum.

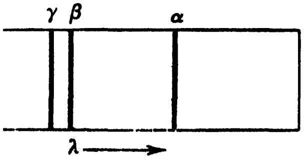


FIG. 14.21.—Typical *K* x-ray spectrum.

In Sec. 14.10, we have spoken of the discovery by Barkla of characteristic x-rays. These rays continue to have the same absorption coefficient in aluminum no matter how much aluminum they have previously passed through. In our present view, the x-rays from a tube consist of rays of a continuous spectrum together with a line spectrum characteristic of the target. X-rays of a given wave-length are found to have a definite absorption coefficient in aluminum and so are homogeneous. The absorption coefficient in aluminum diminishes with decrease of wave-length. It is now known that when the x-rays from a tube fall upon a sheet of copper as in Barkla's experiments, Sec. 14.10, the line spectrum of the copper is excited. Now this spectrum is not absolutely homogeneous because it contains lines of three different wave-lengths, but, even so, it is much more homogeneous than the continuous spectrum from the tube. With the kind of apparatus available at the time Barkla performed his experiments, it was not possible to show by absorption measurements that the characteristic x-rays were not exactly homogeneous. It was remarked in Sec. 14.10 that Barkla found that each element when irradiated by x-rays emits both hard and soft characteristic x-rays. Barkla designated these two kinds of rays by the letters *K* and *L*. The *L* characteristic rays consist of the lines in the second spectrum. Hence the first and second spectra of an element have become known as the *K* and *L* spectra, respectively. Further research has shown the existence of spectra whose lines are of still longer wave-length. These spectra are known as the *M* and *N* spectra. The wave-lengths shown in Table 14.2 are for the lines of the *K* spectra of the elements. The wave-lengths of the most prominent lines in the L spectra of some elements are shown in Table 14.3.

TABLE 14.3

L SPECTRA

Element	Atomic weight	Atomic number	Angstroms			
			α	β_1	β_2	γ
Mo........	96.0	42	5.40	5.17	4.91	4.71
Ag	107.9	47	4.15	3.93	3.69	3.51
W	184.0	74	1.48	1.28	1.24	1.095

It is a very important fact that the x-ray spectrum of an element is very nearly the same whether or not the element is combined chemically with one or more elements. In the case of a compound, each element in the compound emits its own characteristic x-ray spectrum.

The lines of the tungsten L spectrum are arranged as shown in Fig. 14.22.

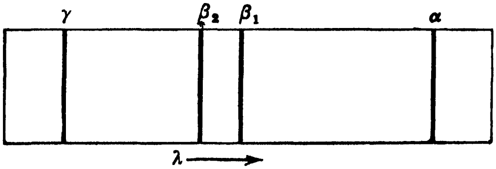


FIG. 14.22.—Prominent lines of the L x-ray spectrum of tungsten.

14.20. Moseley's Law.—Since x-rays are of the same nature as light, the frequency of the rays can be found from the relation

$$\nu = \frac{c}{\lambda} \tag{14-17}$$

where $c = 3 \times 10^{10}$ cm/sec, the velocity of light. Moseley plotted the square root of the frequency of one of the spectrum lines, say, the α line in the K spectra, against the atomic number.and obtained a line as shown by the line α in Fig. 14.23. The spectrum lines β and γ also give rise to the lines β and γ in Fig. 14.23. Since the lines α, β, γ of Fig. 14.23 are very nearly straight, the relation between $\sqrt{\nu}$ and Z, the atomic number, for a given spectrum line such as the α line is

$$\sqrt{\nu} = a(Z - b) \tag{14-18}$$

where *a* and *b* are constants but are different for each of the α, β, and γ lines. Eq. (14–18) expresses what is known as Moseley's law. This equation also holds for the lines of the *L* spectra, the constants *a* and *b*

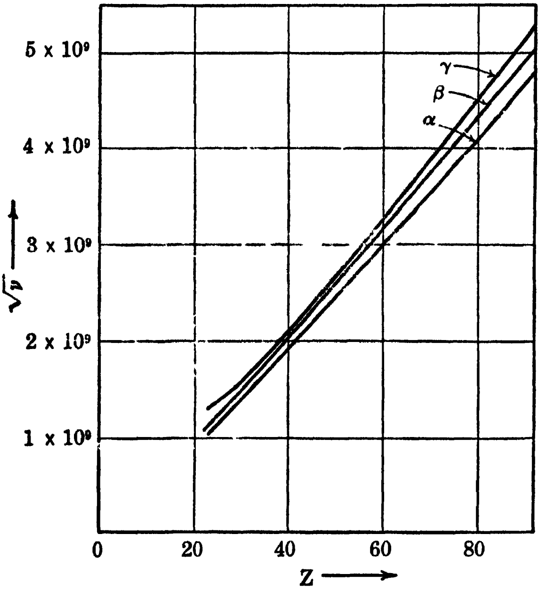


FIG. 14.23.—Moseley diagrams for the α, β, and γ, lines of the *K* spectra of the elements.

having appropriate values. In the case of the α, β, and γ lines of the *K* spectra the constant *b* is so small that we may say that the frequency of a spectrum line varies as the square of the atomic number. This is an approximate statement of Moseley's law.

14.21. Continuous Spectrum.—Referring to Fig. 14.17 it is seen that the curve does not drop to zero between the spectrum lines. There are other wave-lengths present beside the spectrum lines. These wave-lengths belong to the continuous spectrum. The continuous spectrum from an x-ray tube is shown in Fig. 14.24, where the intensity as measured by an ionization chamber is plotted against the wave-length λ. In 1915 Duane and Hunt of Harvard excited an x-ray tube with a steady voltage from a storage battery giving several tens of thousands of volts. They carefully examined the continuous spectrum and found that for a given voltage of excitation there was a definite minimum wave-length as represented by *OA* in Fig. 14.24. Wave-lengths shorter than that represented by *OA* could not be found. By careful measurements they showed that $\lambda_{\min}$ is related to the excitation voltage by

$$\lambda_{\min} = \frac{12{,}345}{V} \tag{14-19}$$

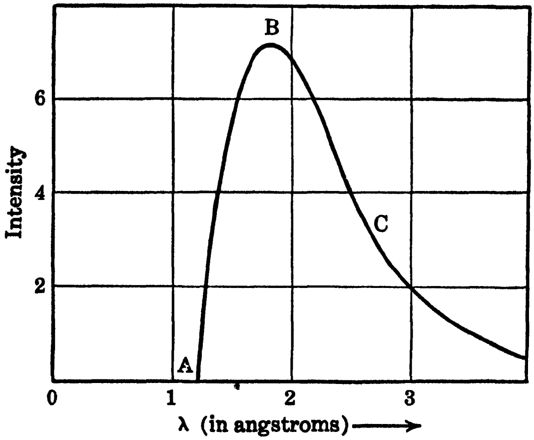


Fig. 14.24.—Distribution of intensity in the continuous x-ray spectrum with respect to wave-length.

where V is given in volts and $\lambda_{\min}$ in angstroms. It will be remembered that a similar relation holds when in Einstein's photoelectric equation

$$m_0c^2\left(\frac{1}{\sqrt{1-\beta^2}} - 1\right) = h\nu - W \tag{14-20}$$

the work function W is neglected and the energy of the expelled electron is measured in electron-volts. Eq. (14–19) is known as the Duane-Hunt relation and the phenomenon which it describes is known as the ***inverse photoelectric effect.*** In the photoelectric effect, light of frequency ν causes the ejection of an electron from an atom with an energy given by (14–20). In the inverse effect, an electron hitting an atom of a target in an x-ray tube excites x-rays of a frequency given by (14–20). The relativity expression for the kinetic energy of the electron is given on the left side of (14–20) because in the case of x-rays the velocity of the electron is not small with respect to the velocity of light.

The distribution of the intensity amongst the different wave-lengths of the continuous spectrum is shown in Fig. 14.24. The maximum intensity occurs at a wave-length which is about 1.5 to 2 times the minimum wave-length.

14.22. Excitation of Spectra.—The *K* spectrum of an element is not excited until the voltage across the x-ray tube is such that the minimum wave-length of the continuous spectrum is just smaller than the wave-length of the γ line. In the case of the *K* spectrum of molybdenum, the γ line has a wave-length of .619 A. The *K* spectrum is excited when $\lambda_{\min}$ of the continuous spectrum is .618 A, so that, applying the Duane-Hunt relation, the voltage across the tube is $12{,}345/.618 = 20{,}000$ volts. For practical purposes the wave-length of the γ line may be used in the formula. It should be noted that all the lines of the *K* spectrum of molybdenum are excited when the potential across the tube is 20,000 volts or more, but that, if the potential is lower than this critical voltage, none of the lines, not even the α or β line, of the *K* spectrum of molybdenum is excited.

The excitation potentials for the *K* spectra of several elements are shown in Table 14.4.

TABLE 14.4

EXCITATION POTENTIALS
K Spectra

Element	Atomic number	Volts
Fe	26	7,110
Co	27	7,720
Ni	28	8,300
Cu	29	8,950
Zr	40	18,000
Mo	42	20,000
Ag	47	25,450
W	74	69,300
Pb	82	89,100

The *L* spectrum of an element is excited by potentials which are much lower than the potential required to excite the *K* spectrum of the same element. The *L* spectra are much more complicated than the *K* spectra, there being many more lines than those listed in Table 14.3. In the case of molybdenum it is found that at a potential below 2520 volts none of the lines of the *L* spectrum are excited. At potentials between 2520 and 2630 volts only some but not all of the lines are excited. If the potential is between 2630 and 2870 volts, more lines appear. It is not until the potential is above 2870 volts that the whole *L* spectrum is excited. There are thus three excitation potentials for the *L* spectrum of each of the elements whose atomic number is above 18. There are five excitation potentials for the *M* spectra of the heavy elements. Because three different sets of lines in the *L* spectrum are excited by different excitation potentials, we speak of the L_1, L_2, and L_3 component spectra of the *L* spectrum. Likewise we speak of the M_1, M_2, M_3, M_4 and M_5 spectra.

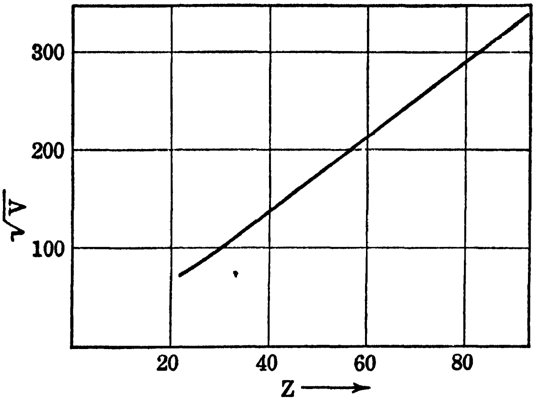


FIG. 14.25.—Relation between the excitation voltage for the *K* x-ray spectrum and the atomic number of an element.

If the square root of the excitation potential for a given spectrum, such as the *K* spectrum, is plotted against the atomic number of the element, a straight line is obtained which very nearly passes through the origin as in Fig. 14.25. Hence, since the intercept on the *Z* axis is small, the excitation potential of a given spectrum is very nearly proportional to the square of the atomic number of the element emitting the spectrum.

14.23. Absorption and Wave-Length.—In Sec. 14.7 we have described the method of measuring the absorption coefficient of x-rays in substances. If the crystal and ionization chamber in Fig. 14.16 are set so that x-rays of a wave-length λ enter the chamber, and if absorbing sheets are placed between the crystal and the chamber, the absorption coefficient of x-rays of any wave-length in the sheet can be measured.

It is necessary here to distinguish between the "total" and the "true" absorption coefficients. When x-rays pass through a sheet of absorbing substance, part of the rays are scattered in all directions as described in Sec. 14.8, and part of the rays are truly absorbed and disappear as x-rays. In absorption experiments, the loss due to both true absorption and scattering is measured, and the absorption coefficient which is caluclated is the total absorption coefficient. This is equal to the sum of the true absorption coefficient and the scattering coefficient. It is usual to represent the true absorption coefficient by τ, the relation of τ to μ and σ (see Secs. 14.7 and 14.8) being

$$\tau = \mu - \sigma \tag{14–21}$$

If τ is divided by the density ρ, the true mass absorption coefficient is obtained. Values of τ/ρ are obtained from the experimental values of μ/ρ and σ/ρ. For many elements σ/ρ is approximately 0.2 per gm/cm^2.

For a given element τ/ρ varies nearly as λ^3 and for a given wave-length τ/ρ varies nearly as Z^3, so that we may write

$$\frac{\tau}{\rho} = C\lambda^3 Z^3 \tag{14–22}$$

where C is a constant of proportionality. Equation (14–22) holds over a wide range of wave-lengths and of atomic numbers, but does not hold for all wave-lengths and all atomic numbers. If, for instance, τ/ρ is determined for x-rays of various wave-lengths absorbed in copper, a curve as shown in Fig. 14.26 is obtained. If we begin with a wave-length of, say, 4 A and then decrease the wave-length, we find that τ/ρ decreases with the cube of the wave-length according to (14–22) as shown by AB in Fig. 14.26. The decrease in τ/ρ continues until a wave-length of 1.378 A is reached, when there is a sudden increase in τ/ρ as shown by BC. As the wave-length is still further decreased, τ/ρ again decreases with the cube of the wave-length as shown by CD according to (14–22). Along AB the constant C has one value while along CD this constant has another value. The wave-length at which the sudden jump in τ/ρ occurs is known as the critical absorption wave-length. This wave-length for copper is 1.378 A.

It is found that K characteristic rays of copper (see Sec. 14.10) are not excited unless the wave-length of the x-rays falling upon the copper is less than 1.378 A. According to the Duane-Hunt relation, this wave-length corresponds to a potential of 12,345/1.378 or 8950 volts. This is the potential necessary to excite the K spectrum of a copper target of an x-ray tube (see Table 14.4). The extra energy which is absorbed for wave-lengths less than 1.378 A therefore goes to excite the *K* spectrum of copper.

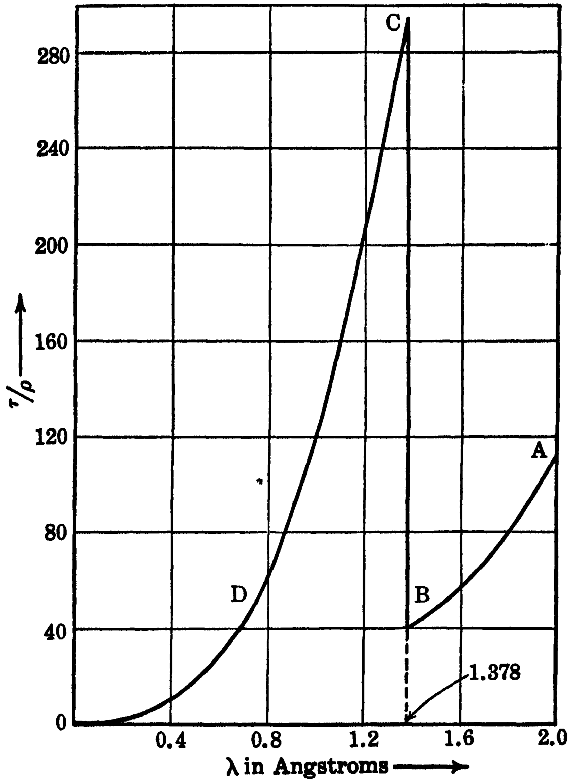


FIG. 14.26.—Variation of the true absorption coefficient of x-rays in copper ($Z = 29$) with the wave-length.

If x-rays of very long wave-lengths are used, it is found that there are three critical absorption wave-lengths corresponding to the three excitation potentials for the *L* spectrum. For instance, in the case of molybdenum the three excitation potentials for the *L* spectrum are 2520, 2630, and 2870 volts as we have seen in Sec. 14.22. Corresponding to these are the critical absorption wave-lengths for molybdenum of 4.90, 4.70 and 4.30 A.

Since there are critical absorption wave-lengths corresponding to the various spectra of an element, we speak of the *K* and the L_1, L_2, and L_3 critical absorption wave-lengths.

If τ/ρ is measured for a given wave-length in various elements, a curve similar to that shown in Fig. 14.27 is obtained. This curve is for x-rays of wave-length 1.490 A. Along *DC*, τ/ρ varies as Z^3. At $Z =$ 28 (nickel) there is a sudden drop in τ/ρ. Along *BA*, τ/ρ again varies as Z^3. The constant *C* in (14–22) has one value along *AB* and another value along *CD*, Fig. 14.27. The wave-length 1.490 A is short enough to excite the *K* spectrum of nickel but is not short enough to excite the *K* spectrum of copper, the next higher element.

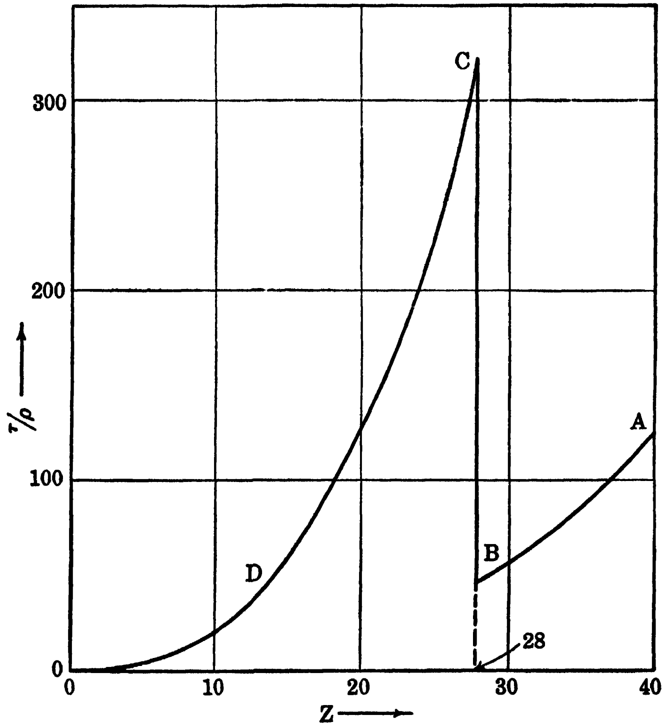


Fig. 14.27.—Variation of the true absorption coefficient of x-rays of wave-length 1.490A with the atomic number of the absorber.

The mass absorption coefficient of an element for x-rays of a given wave-length is, except for very small variation, the same whether the element is in combination with other elements or not. Hence the mass absorption coefficient of a compound can be calculated. For instance, the mass absorption coefficient or μ/ρ of aluminum for x-rays of wave-length 0.7 A is 5.00 per gm/cm^2, while the μ/ρ of oxygen for the same wave-length is 1.10. The formula for aluminum oxide is Al_2O_3, so that, since the atomic weights of aluminum and oxygen are 27 and 16 respectively, the molecular weight of Al_2O_3 is $2 \times 27 + 3 \times 16 = 102$. One gram of Al_2O_3 therefore contains $2 \times 27/102 = 0.53$ and $3 \times 16/102 = 0.47$ grams of Al and O respectively. The μ/ρ of Al_2O_3 for $\lambda = 0.7$ A is therefore $0.53 \times 5.00 + 0.47 \times 1.10$ or 3.167 per gm/cm^2.

If Tables 14.2 and 14.4 are examined it is seen that the wave-lengths of the α, β, γ lines of the K spectra and the values of the excitation potentials of the K spectra of iron, cobalt, and nickel follow the order Fe, Co, Ni and not the order Fe, Ni, Co. Hence, it is the atomic number, rather than the atomic weight, of an element which determines its properties as regards x-rays. The critical absorption wave-lengths for these three elements also decrease according to the order Fe, Co, Ni. Further, it is to be noted that the atomic number Z rather than the atomic weight occurs in Moseley's law as given in (14–18) and in the absorption law as given in (14–22).

The mass absorption coefficients of several elements for x-rays of various wave-lengths are shown in Table 14.6.

TABLE 14.6

MASS ABSORPTION COEFFICIENTS (μ/ρ)

Element and atomic number	Wave-length in angstroms				
	0.25	0.50	0.70	1.00	1.40
C 6	0.18	0.31	0.54	1.30	4.0
O 8	0.21	0.50	1.10	2.87	
Al 13	0.37	1.91	5.00	13.8	38.0
Fe 26	1.93	14.3	36.3	95.0	270.0
Cu 29	2.75	18.8	49.2	133.0	40.0
Mo 42	7.42	48.6	18.8	51.0	
W 74	5.60	38.0	....	....	141.0
Pb 82	8.40	59.3	133.0	77.0	185.0

The table shows that there is a critical absorption wave-length for copper (Cu) between 1.00 and 1.40 A, since the value of μ/ρ for the latter wave-length is less than that for the former wave-length. Also, as we proceed down the column headed 0.25 A, there is a drop in μ/ρ as we go from molybdenum (Mo) to tungsten (W). This means that 0.25 A is less than the K critical absorption wave-length for molybdenum but greater than the K critical absorption wave-length for tungsten.

14.24. Nature of X-Rays.—We have remarked in Sec. 14.12 that at the beginning of 1912 there was a controversy as to the nature of x-rays. Are x-rays a form of wave-motion or are they corpuscles? With Laue's discovery of the diffraction of x-rays by crystals and Bragg's discovery of the reflection of x-rays by crystals, the wave nature of x-rays seemed to be established. In Sec. 14.8 we have described the scattering of x-rays, and in Sec. 14.12 we have shown that the classical electromagnetic theory as applied by Thomson to x-rays seems to give an adequate explanation of the scattering phenomenon. In the sections following Sec. 14.12, we have seen that x-rays have a wave-length and on that account can be considered as electromagnetic vibrations. On this view, when an electromagnetic wave passes over an atom, the electrons in the atom are set vibrating. The frequency of the vibrating electrons will of course be identical with the frequency of the electromagnetic waves passing over them. We have seen in Sec. 6.4 that a vibrating electron emits electromagnetic radiation (or waves) which has the same frequency as that of the vibrating electron. Thomson originally calculated the formula for the mass scattering coefficient of substances for x-rays and obtained (14–10). Thomson's calculation was based upon the electromagnetic pulse theory of x-rays, but the same formula (14–10) is obtained if the x-rays are assumed to be electromagnetic vibrations with a definite frequency or wave-length. Since Barkla has shown that (14–10) is in such good agreement with the experimental value of σ/ρ for many substances, it seems that an electromagnetic wave-theory of x-rays gives an excellent explanation of the scattering of x-rays. Moreover, an electromagnetic wave theory gives an explanation of the polarization phenomenon described in Sec. 14.9. The x-rays coming from an x-ray tube are for the most part unpolarized. A beam of such x-rays traveling along AB, Fig. 14.28, consists of electromagnetic waves in which the electric field is vibrating in all directions in the plane perpendicular to AB as shown at CD. However, we can resolve these vibrations into two component sets of vibrations, EF and GH. An electron at O is set vibrating and the vibration has components along EF and GH. We have seen in Sec. 6.4 that a vibrating electron does not emit electromagnetic radiation in the direction of vibration. Hence the component vibration HG does not emit any radiation in the direction OP. A second electron at *P*, where *OP* is perpendicular to *AB*, receives only those waves which have their electrical vibrations along *LM*, which is parallel to *EF* and perpendicular to *OP*. Thus the radiation which is scattered by the electron at *O* in the direction *OP* is completely polarized. This radiation falling on the second electron at *P* causes it to vibrate in the direction *LM*. This vibrating electron then emits radiation in all directions excepting in the direction of its vibration. An ionization chamber placed with its window at *Q* therefore receives no scattered x-rays from the second electron at *P*, but receives scattered x-rays from the electron at *P* if the window is at *R*.

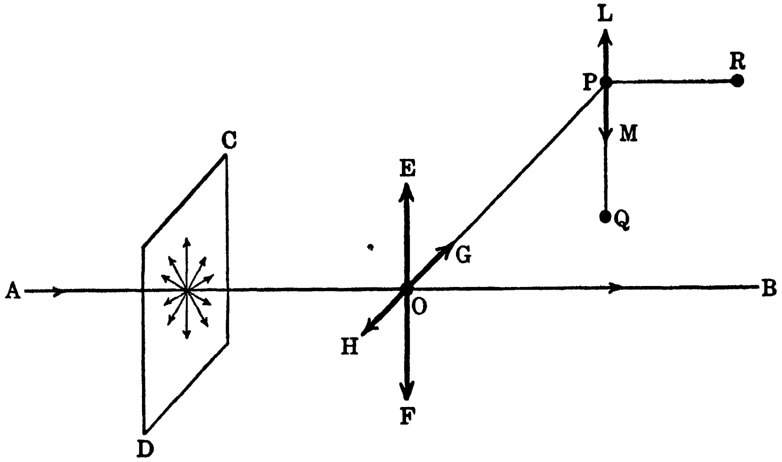


FIG. 14.28

However, the other facts about x-rays, such as the ejection of secondary cathode rays as described in Sec. 14.12 and the ionization of only a few molecules when x-rays pass through air, persisted, and after a while it came to be understood that the Laue and Bragg experiments only show that x-rays are of the same nature as light. The ejection of secondary cathode rays is an example of the photoelectric effect. The photoelectric effect, as we have seen in Sec. 11.7, seems to require a form of corpuscular theory of light. The quantum theory is the modern form of the corpuscular theory. A cathode electron in the x-ray tube strikes an atom in the target and the inverse photoelectric effect occurs, so that an x-ray quantum is produced according to the relation

$$h\nu = m_0c^2\left(\frac{1}{\sqrt{1-\beta^2}} - 1\right)$$

This quantum then proceeds from the x-ray tube until it strikes an atom, when the photoelectric effect occurs and an electron is ejected. If this is an atom of gas, the atom becomes ionized. The ejected electron travels with such high speed that it knocks electrons out of other atoms, thus leaving a train of ionized atoms, the effects of which can be made visible under proper conditions. If the ions thus produced find themselves in the electric field of an ionization chamber, they are drawn to the electrodes—the positive ions to the negative electrode and the negative ions to the positive electrode.

The question as to whether x-rays (or light) are a form of wave motion or consist of a swarm of corpuscles still persists. Previous to 1923 it was held by one school of physicists that the quantum process occurred only when light was emitted or absorbed but that between the atom which emitted the light and the atom which absorbed it the light traveled through the ether as an electromagnetic wave-motion according to the Maxwell theory. This school maintained that there are no such entities as light quanta, but that light is emitted and absorbed in quanta. The emission and absorption of light in quanta was explained as being due to the structure of the atom itself. This view is expressed in some degree in Secs. 9.11 and 9.12. However, in 1922–23 A. H. Compton, then of Washington University, St. Louis, performed an experiment which seemed to show that x-rays are not only emitted and absorbed as quanta but are also quanta as they travel through space. These discrete quanta of light or x-rays are now known as photons.

14.25. Compton Effect.—In the previous section we have seen that the phenomena connected with the scattering of x-rays are very well explained by the classical electromagnetic wave-theory. It is another of the ironies of physics that the best evidence in favor of x-rays being corpuscular in form is given by a phenomenon connected with the scattering of x-rays. In 1923 it occurred to A. H. Compton to irradiate a block of scattering substance with x-rays of a definite wave-length and then to examine the scattered rays by means of an x-ray spectroscope. The classical theory always demands that the scattered rays shall be of the same frequency as the primary waves. When an electron is set oscillating by a vibrating field, the frequency of the vibrations of the electron must equal the frequency of the vibrations of the electric field,

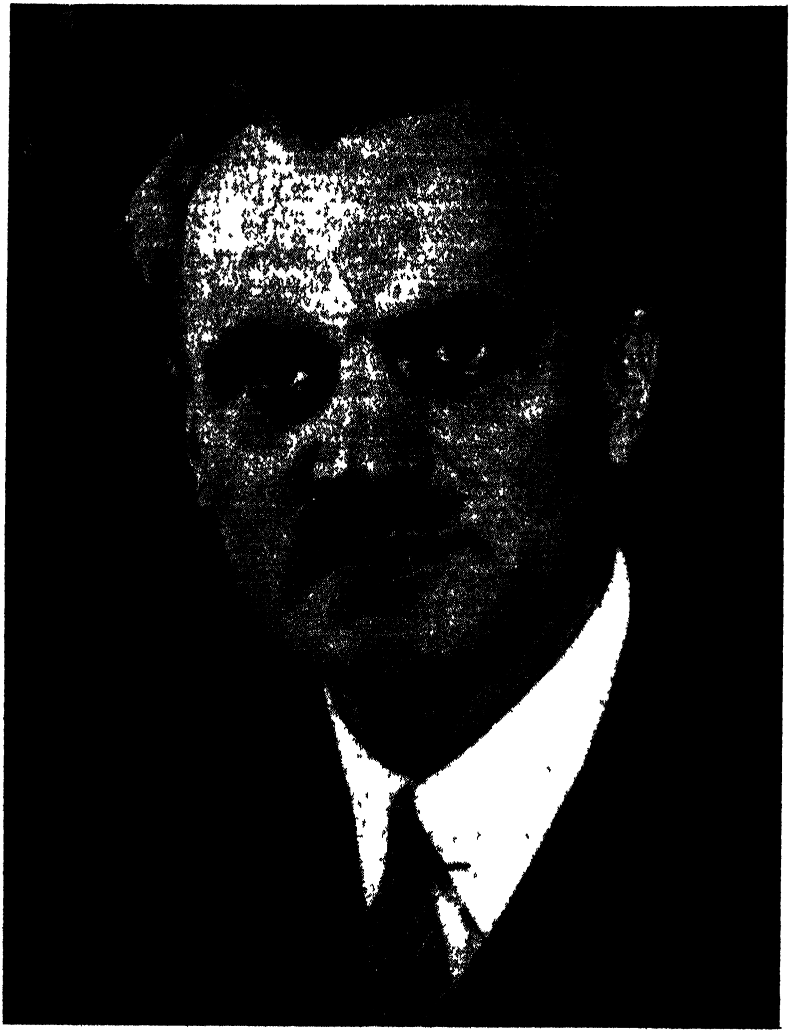


FIG. 14.29.—Arthur Holly Compton.

and, furthermore, the frequency of the radiation emitted by the electron must be equal to that of the electron and to that of the original vibrating electric field. Nevertheless, Compton, on performing his experiment, found that the scattered rays consisted of two frequencies, one equal to and the other less than that of the primary rays. The experimental arrangement is shown diagrammatically in Fig. 14.30. X-rays were allowed to fall upon a carbon block *C*. The rays scattered in the direction *CP* were reflected by a crystal *P* into the ionization chamber *A*. An x-ray tube with a molybdenum target was used and the *K* spectrum of molybdenum excited. If the scattered rays are of the same frequency as the primary rays, the spectrum of the scattered rays should not be the same as that of the primary rays. However, Compton obtained not only the *Kα* line of molybdenum in the scattered rays but another line on the long wave-length side of the *Kα* line. Later, it was also found that the *Kβ* line in the scattered spectrum was accompanied by a line on the long wave-length side. Ross, of Stanford University, using a photographic method, obtained results as shown in Fig. 14.31. The new lines are shown as α' and β'. Further, it was found that the wave-length separation of α and α' is the same as that of β and β'. Moreover, if an x-ray tube with a silver or tungsten target is used, the wave-length separation of the α and α' lines equals that of the α and α' lines when an x-ray tube with a molybdenum target is used, so that the separation does not depend on the nature of the target. However, the wave-length separation does depend on the angle of scattering which is represented by ϕ in Fig. 14.30. It is found that the wave-length separation which we shall designate by $\lambda' - \lambda$ is given by

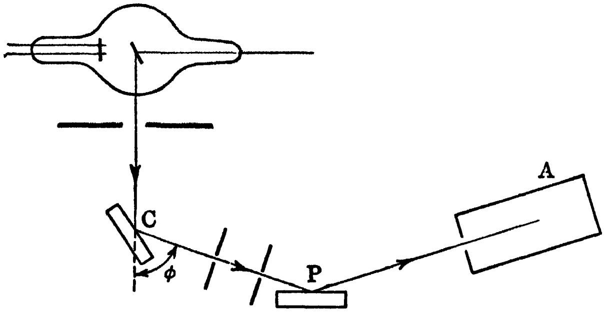


FIG. 14.30—Experimental arrangement for showing the Compton effect.

$$\lambda' - \lambda = 0.024(1 - \cos\phi) \qquad (14\text{–}23)$$

when $\lambda' - \lambda$ is measured in angstroms. Again, it is found that under certain conditions the line of the same wave-length as the line in the primary spectrum can be made to disappear. Compton thus found that there is a change of wave-length on scattering. However, as indicated in Fig. 14.31, Compton found two lines—the unmodified line which has the same wave-length as that of the primary rays and the modified line whose wave-length is changed from that of the primary rays according to (14–23).

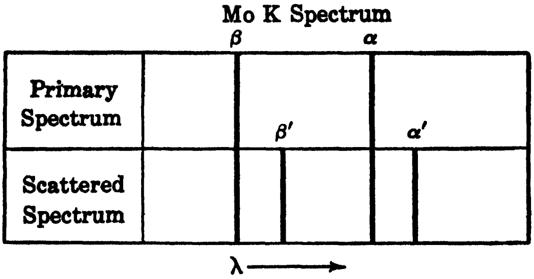


FIG. 14.31.—Comparision of the spectrum of scattered with that of primary x-rays.

14.26. Theory of the Compton Effect.—According to Chapter XIII the kinetic energy plus the mass energy of a body is $Mc^2/\sqrt{1 - \beta^2}$ where M is the rest-mass and βc is the velocity of the body. The momentum of the body is $M\beta c/\sqrt{1 - \beta^2}$, so that the ratio

$$\frac{\text{Momentum}}{\text{Total energy}} = \frac{\beta}{c} \tag{14–24}$$

Let us suppose that a beam of light or x-rays consists of a swarm of photons (a photon is the modern term for a corpuscle of light) and that each of these photons travels with a quantum of energy $h\nu$ and velocity c. Then $\beta = 1$ in (14–24) and we have

$$\text{Momentum of a photon} = h\nu/c \tag{14–25}$$

Imagine that an x-ray photon collides with an electron as indicated in Fig. 14.32. The photon bounces off the electron at B and then travels in the direction BC, which makes an angle ϕ with the original direction of travel AB. It is assumed that the electron was at rest before the collision. After the collision, the electron recoils in the direction BD. Momentum is a vector, so that if BE in Fig. 14.32 represents the momentum of the photon before collision, the lines BC and BD in the parallelogram $BCED$ represent respectively the momentum of the photon and that of the electron after the collision. According to a theorem in trigonometry

$$CE^2 = BC^2 + BE^2 - 2BC \cdot BE \cos \phi \tag{14-26}$$

where $\phi = \angle CBE$ and is the angle of scattering. The momentum of the recoiling electron is $m\beta c/\sqrt{1 - \beta^2}$, where $\beta = v/c$, v is the velocity

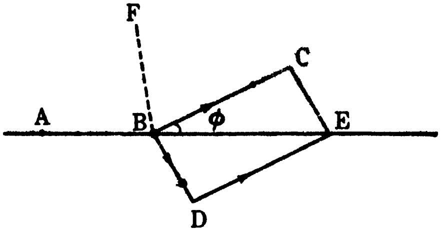


Fig. 14.32.—Theory of the Compton Effect.

of the recoiling electron, and c is the velocity of light. Hence, from the principle of the conservation of momentum, we have

$$\frac{m^2\beta^2c^2}{1 - \beta^2} = \left(\frac{h\nu'}{c}\right)^2 + \left(\frac{h\nu}{c}\right)^2 - 2\frac{h\nu}{c} \cdot \frac{h\nu'}{c} \cdot \cos \phi \tag{14-27}$$

where ν' is the frequency of the scattered photon.

We suppose the collision to be elastic, so that there is no loss of kinetic energy. Conservation of kinetic energy requires that

$$h\nu = h\nu' + mc^2\left(\frac{1}{\sqrt{1 - \beta^2}} - 1\right) \tag{14-28}$$

where $h\nu$ and $h\nu'$ are the energies of the primary and scattered photons, respectively. If we treat ν' and β as the unknowns in (14–27) and (14–28), we obtain

$$\nu' = \frac{\nu}{\left\{1 + \frac{h\nu}{mc^2}(1 - \cos \phi)\right\}} \tag{14-29}$$

Replacing the frequencies ν and ν' by the wave-lengths λ and λ' by means of the relations $\nu = c/\lambda$ and $\nu' = c/\lambda'$, we obtain

$$\lambda' - \lambda = \left(\frac{h}{mc}\right)(1 - \cos\phi) \tag{14-30}$$

Putting in the values $h = 6.62 \times 10^{-27}$ erg sec, $m = 9.0 \times 10^{-28}$ gm, and $c = 3 \times 10^{10}$ cm/sec, we obtain $h/mc = 0.024 \times 10^{-8}$ cm or 0.024 A, agreeing with the constant in (14–23).

In the foregoing theory, it was assumed that the electron was at rest before the collision. In 1925 Jauncey extended the analysis to the case where the electron has a velocity inside the atom before being struck by the x-ray photon. Two cases were distinguished. In the first, the electron is knocked out of the atom notwithstanding the direction of its velocity relative to that of the incident photon. However, the change of wave-length of the scattered photon depends on the pre-collision velocity of the electron according to

$$\lambda - \lambda' = (h/mc)\cdot(1 - \cos\phi) - 2\lambda(v/c)\sin{}^{\phi}/_{2} \tag{14-31}$$

where v is the component of the velocity of the electron just before collision in the direction BF, which is the bisector of the angle ABC in Fig. 14.32. According to the modern theory of the atom, the electrons in the atom move with various velocities in various directions. Hence where a beam of photons is scattered by the electrons of a large number of atoms each collision gives rise to a different change of wave-length. But only the effect of many collisions can be observed and the Compton modified line is found to have a wave-length greater than that of the primary line. This width is in accordance with (14–31), when v can be positive as well as negative.

In the second case Jauncey found that for certain velocities and directions of these velocities the amount of energy communicated by the photon to the electron in the atom was not sufficient to knock the electron out of the atom. In this case the electron is held rigidly to the atom and the mass m which enters into the Compton equation (14–30) is the mass of the atom rather than the mass of the electron. In such a case the change of wave-length is far too small to be observed, and we have the unmodified line.

14.27. Diffraction from a Ruled Grating.—In Sec. 14.13 it was stated that it seemed to be necessary to rule a grating with something like 100,000,000 lines to the centimeter in order to produce diffraction of x-rays by a ruled grating. This was the view in 1912 when Laue made his discovery. However, in 1922 A. H. Compton discovered that x-rays can be totally reflected from a substance like glass if the glancing angle of incidence is small enough. For instance, x-rays of wave length 1.28 A are totally reflected from crown glass of density 2.52 at glancing angles less than a critical angle of $\theta = 11$ minutes of arc. It therefore seemed possible to obtain diffraction of x-rays from a ruled grating at glancing angles within the region of total reflection. It turns out that due to the small glancing angles involved the necessary number of lines per cm which must be ruled is the very opposite of being prohibitively great. The simple theory of the ruled grating is as follows: X-rays of wave-length λ fall at a glancing angle θ upon a ruled grating as shown in Fig. 14.33. The glancing angle of diffraction is ϕ. A and B represent consecutive rulings on the grating. For strong diffraction in the direction ϕ the path difference, $CB - AD$, must be a whole number of wave-lengths. Hence, if d represents AB,

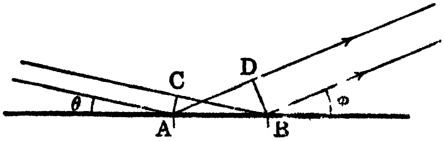


FIG. 14.33.—Diffraction of x-rays from a ruled grating.

$$n\lambda = d(\cos\theta - \cos\phi) \tag{14-32}$$

But, for small angles,

$$\cos\theta = 1 - \theta^2/2$$

so that (14–32) becomes

$$n\lambda = d(\phi^2 - \theta^2)/2 \tag{14-33}$$

Putting $n = 1$, $\lambda = 0.71$ A, $\phi = 8'$, $\theta = 5'$, we find $d = 0.0043$ cm. The grating thus needs to be ruled with only $1/.0043 = 212$ lines to the centimeter.

Consider the diffraction of a spectrum line (say, the Mo $K\alpha_1$ line) from a ruled grating whose grating space d has been made a certain value by ruling. The wave-length can then be found from (14–33). This same spectrum line is then reflected from a calcite crystal. Since the wave-length is known, the grating space for calcite can be found by applying Bragg's law. From Sec. 14.17 we see that d_0 the grating space of rocksalt is related to the molecular weight W and density ρ of the crystal by the formula

$$d_0^3 = W/2\rho N \tag{14-34}$$

where N is Avogadro's number. Solving for N, we have

$$N = W/2\rho d_0^3 \tag{14-35}$$

For calcite the formula is

$$N = W/(2 \times 1.096 \times \rho d_0^3) \tag{14-36}$$

where the number 1.096 is due to the complicated structure of calcite. Since d_0 is known, N can be calculated. Accurate measurements give $N = 6.03 \times 10^{23}$. The faraday is the charge necessary to deposit one gram-atom of hydrogen in an electrolytic cell. Its value is 2.893×10^{14} esu/gm-atom. If each hydrogen ion carries a charge e' esu, $Ne' = 2.893 \times 10^{14}$, whence

$$e' = 2.893 \times 10^{14}/6.03 \times 10^{23} = 4.8 \times 10^{-10} \text{ esu}$$

This is in satisfactory agreement with the most recent value for e' obtained by Millikan's oil-drop method.

CHAPTER XIV

PROBLEMS

1. Find the speed in cm/sec with which the electrons in a Coolidge x-ray tube hit the target when the potential across the tube is 70,000 volts.

2. The mass absorption coefficient of x-rays of a certain hardness in aluminum is 3.18 per gm/cm². Find the thickness of aluminum which will reduce the intensity of the x-rays penetrating the aluminum to 0.2 of the original intensity, being given that the density of aluminum is 2.7 gm/cm³.

3. A thickness of 1.05 mm of copper reduces the intenstiy of a beam of x-rays to 0.075 of its original intensity. Find the mass absorption coefficient of the copper, being given that the density of copper is 8.93 gm/cm³.

4. The x-rays from a silver target are examined with a Bragg spectrometer, upon the axis of which a crystal of rocksalt is mounted. Strong reflections in the first order are obtained from a (1, 0, 0) crystal face at the following glancing angles: 4°57′, 5°4′, 5°43′. Determine the wave-lengths of the lines of the K x-ray spectrum of silver.

5. Sylvine (potassium chloride — KCl) crystallizes in a form similar to that of rocksalt (sodium chloride — NaCl). The density of sylvine is 1.99 gm/cm³ and the atomic weight of potassium is 39.1. Determine the principal grating space of sylvine.

6. Using the principal grating space of sylvine as found in Problem 5, determine the glancing angles at which the line of the K x-ray spectrum of cobalt is reflected in the third order from a (1, 0, 0) face of sylvine.

7. Determine the glancing angles at which α, β_1, β_2, and γ lines of the L x-ray spectrum of tungsten are reflected in the second order from a natural face of a crystal of calcite ($CaCO_3$), given that the principal grating space of calcite is 3.03 angstroms.

8. Using Moseley's law, determine the approximate wave-lengths of the α, β, and γ lines of the K spectrum of bromine (atomic number = 35) from the wave-length of the α, β, and γ lines of the K spectrum of molybdenum.

9. A potential difference of 85,000 volts is applied to an x-ray tube with a silver target. Find the shortest wave-length emitted by the tube.

10. Determine the excitation voltage for the K x-ray spectrum of iodine (atomic number = 53).

11. From the absorption laws, determine the mass absorption coefficient of x-rays of wave-length 0.9 A in (*a*) zirconium (atomic number = 40), and (*b*) zinc (atomic number = 30).

12. The wave-lengths of the $K\alpha$ and $K\beta$ lines of molybdenum are .710 A and .631 respectively. The K critical excitation potential of zirconium is 18,000 volts. Determine the mass absorption coefficient of zirconium for (*a*) the $K\alpha$ line and (*b*) the $K\beta$ line of molybdenum. Determine the amount of zirconium which will reduce the intensity of the $K\beta$ line of molydbenum to one-tenth. Next, find the fraction to which this amount will reduce the intensity of the $K\alpha$ line of molybdenum. Hence describe how a thin sheet of zirconium may be used to filter out the $K\beta$ (and also the $K\gamma$) line of molybdenum from the x-rays coming from a molydbenum target, leaving the $K\alpha$ line in the beam.

13. Find the change of wave-length when the $K\beta$ line of silver is scattered from carbon at angles of (*a*) 60°, (*b*) 90°, and (*c*) 150°. Determine the wave-lengths of the modified and unmodified lines in the Compton effect at each of these angles.

14. Determine the mass absorption coefficient of potassium sulphate (K_2SO_4) for x-rays of wave-length 0.5 A. Use Table 14.5 and plot curves.

CHAPTER XV

SOME APPLICATIONS OF X RAYS

15.1. Crystal Analysis.—We have shown in the previous chapter that the wave-length of x-rays can be obtained from Bragg's law

$$n\lambda = 2d \sin \theta \tag{15-1}$$

by measuring θ, d being known. Conversely, if λ is assumed known and θ is measured, the grating space d can be calculated. In Sec. 14.17 we have described the method of finding the principal grating space of rocksalt. This grating space is the distance between consecutive (1, 0, 0) planes of atoms. But there are other planes, such as the (1, 1, 0) and (1, 1, 1) planes. The grating space of each of these sets of planes is different from that of the (1, 0, 0) planes. By cutting a crystal in such a way that the new crystal face is parallel to the (1, 1, 0) planes and reflecting x-rays of definite wave-length, as, for instance, that of the molybdenum $K\alpha$ line, from this face, the value of d for the (1, 1, 0) planes can be found. Similarly by reflection from a face parallel to the (1, 1, 1) planes, the value of d for the (1, 1, 1) planes can be found. Faces can be cut parallel to other sets of planes in the crystal and the corresponding d's can be found. We now have a set of d's, each d corresponding to a given set of crystal planes. If the ratio of each d to the principal grating space for the corresponding planes is recorded in a table, the values of these ratios depend upon the way in which the atoms are arranged in the crystal. This arrangement of the atoms in a crystal is called the crystal structure.

Crystals have many different forms. Two of these forms are the cubic and the rhombohedral form, examples being rocksalt (NaCl) and calcite ($CaCO_3$) respectively. The unraveling of crystal structure by means of x-rays is very complicated, so that we shall content ourselves with describing the unraveling of cubic crystals since this is the simplest case. However, the principles underlying the unraveling of the crystal structure in the more complicated cases are the same as those used in the simple case of cubic crystals.

There are three types of cubic crystals—(1) simple cube, (2) face-centered cube, and (3) body-centered cube. The simple cubic form is shown in Fig. 15.1a, the face-centered cubic form is shown in Fig. 15.1b, and the body-centered cubic form in Fig. 15.1c. Each of the lattices shown in Fig. 15.1 extends indefinitely in all directions. We shall take the atom at the inside lower left-hand corner as being at the origin of coordinates. It is then seen that the coordinates of an atom in a cubic lattice are $(h'd_{100}, k'd_{100}, l'd_{100})$, where h', k', and l' are integers, positive or negative, including zero. For a simple cubic lattice there are no other restrictions on h', k', and l'; for a face-centered cubic lattice $(h' + k' + l')$ must be an even number; and for a body-centered lattice h', k', and l' must be either all even or all odd numbers.

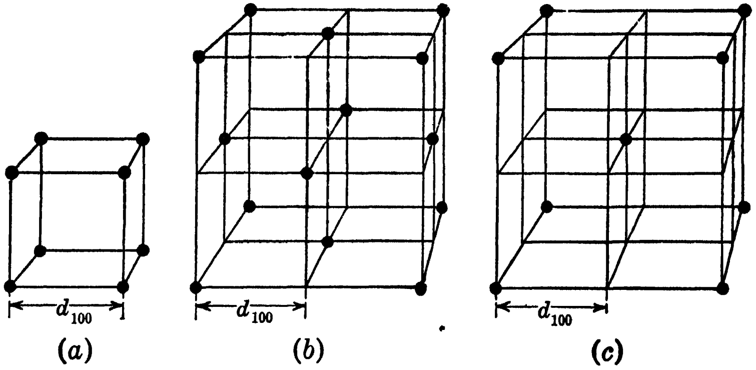


FIG. 15.1.—Types of cubic lattices—(a) simple cube,(b) face centered cube, and (c) body centered cube.

In Sec. 14.16 we have described what is meant by (h, k, l) planes. If the (h, k, l) plane is in a crystal, it must contain atoms if x-rays are to be reflected according to Bragg's law. For a simple cubic crystal the distance between two consecutive parallel planes is

$$d_{hkl} = d_{100}/\sqrt{h^2 + k^2 + l^2} \tag{15-2}$$

where h, k, and l are any integers, positive or negative. We have noted in Sec. 14.16 that the (4, 2, 2) planes, for instance, are the same as the (2, 1, 1) planes since there is a common factor 2 in 4, 2, and 2. We shall now speak of an (h, k, l) reflection where h, k, and l may contain a common factor. For instance, we speak of a (4, 2, 2) reflection from a simple cubic crystal. Since there is a common factor 2, this is taken to refer to the second order reflection from the (2, 1, 1) planes. Hence, for a simple cubic crystal, Bragg's law becomes

$$\sin\theta = (\lambda/2d_{100})\sqrt{h^2 + k^2 + l^2} \tag{15-3}$$

where the notation (h, k, l) refers to the reflection and h, k, and l may be any integers, positive or negative, including zero.

Since there are restrictions on h', k', l' in the coordinates $(h'd_{100}, k'd_{100}, l'd_{100})$ of an atom in a face-centered or body-centered cubic lattice, there are corresponding, but not the same, restrictions on h, k, and l in (15-3). For these two lattices an (h, k, l) reflection occurs at θ, as given by

$$\sin\theta = (\lambda/4d_{100})\sqrt{h^2 + k^2 + l^2} \tag{15-4}$$

For a face-centered cubic lattice, h, k, l must be all even or all odd, whereas, for a body-centered cubic lattice, $(h + k + l)$ must be an even number. Reflections of the (1, 0, 0) type are not given by these two lattices. However, a (1, 1, 1) reflection occurs from a face-centered lattice at a value of $(2d_{100}\sin\theta)/\lambda$ smaller than the value of the same quantity for a simple cubic lattice.

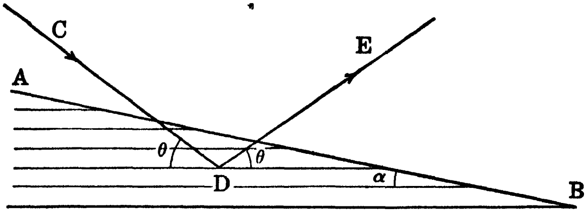


FIG. 15.2.—Reflection of x-rays occurs at the crystal planes and not at the surface of a crystal.

15.2. Powdered Crystal Method.—It is not necessary, although in the Bragg method it is convenient, for the face of a crystal to be parallel to any particular set of planes for reflection from that set of planes to occur. For instance, a crystal may be cut with its face making an angle α with the crystal planes as shown in Fig. 15.2. The reflection takes place from the crystal planes and not from the crystal face. The reflection occurs as shown in Fig. 15.2.

If now a crystal is broken up into a fine powder, each of the grains of the powder is a small crystal with the various sets of planes, (1, 0, 0), (1, 1, 0), (1, 1, 1) and so on. If the powder is irradiated by x-rays of a given wave-length, say, the $K\alpha$ line of molybdenum or of copper, a fraction of the powder grains will have their crystal planes oriented in such a direction that they reflect the $K\alpha$ rays in the first order ($n = 1$, in Bragg's equation) from the (1, 0, 0) planes, another fraction of the grains will have their crystal planes oriented in such a direction that they reflect the $K\alpha$ rays in the first order from the (1, 1, 0) planes, and so on. There are reflections not only from the different sets of planes but also there are reflections in the second ($n = 2$) and higher orders from the various sets of planes.

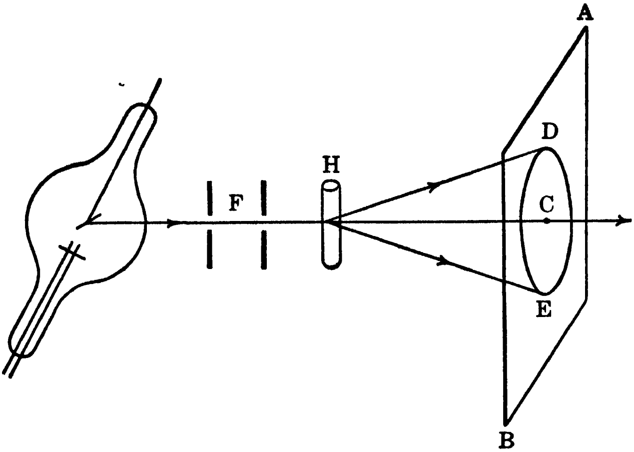


Fig. 15.3.—Experimental arrangement for obtaining a powdered crystal x-ray photograph.

In the experimental arrangement a narrow pencil of x-rays falls upon a holder H containing the powdered crystal as shown in Fig. 15.3. A photographic film which is enclosed in a light-tight envelope is placed at AB. Some of the primary rays pass through the powdered crystal and produce a central spot at C. The rays reflected in a given order and from a given set of planes produce a circle of blackness on the photographic film as shown by DE in Fig. 15.3 instead of a line of blackness. The circle is formed because the crystal grains in H which are reflecting x-rays to produce the circle DE can have their crystal planes oriented in any azimuth about FC as axis. Instead of a film AB as shown in Fig. 15.3 it is customary to use a thin strip and this strip is bent in the form of a cylinder of small height about H as axis. The photograph obtained is then as shown in Fig. 15.4. Portions of circles are obtained. Each of these circular lines represents a reflection of the same wave-length in a certain order from a certain set of planes. The crystal structure can be obtained from the arrangement of the lines and their relative intensities.

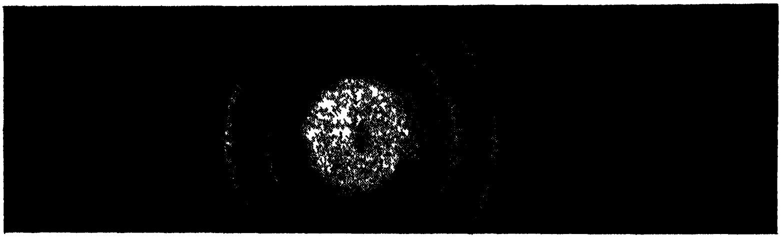


FIG. 15.4.—Powdered crystal x-ray photograph of potassium chloride taken in the author's laboratory. Mo Kα rays were used, the Kβ rays being mostly removed by a filter of zirconium nitrate. X-ray tube operated at 45,000 volts r.m.s. and 25 milliamperes; exposure, 13 hours; target to powdered crystal, 18 cm; powdered crystal to film, 5.5 cm.

In the case of simple cubic crystals the circles as shown in Fig. 15.4 correspond to values of sin θ which are proportional to $\sqrt{h^2 + k^2 + l^2}$ where h, k, and l are whole numbers. By reference to Table 15.1, it is seen that it is impossible to obtain $\sqrt{7}$ and $\sqrt{15}$ from $\sqrt{h^2 + k^2 + l^2}$ and the circles corresponding to sin θ proportional to $\sqrt{7}$ and $\sqrt{15}$ are missing in a powdered crystal photograph of a simple cubic crystal.

TABLE 15.1

h, k, l	$h^2 + k^2 + l^2$	h, k, l	$h^2 + k^2 + l^2$	h, k, l	$h^2 + k^2 + l^2$
1, 0, 0	1		7	3, 2, 0	13
1, 1, 0	2	2, 2, 0	8	3, 2, 1	14
1, 1, 1	3	3, 0, 0	9	. . .	15
2, 0, 0	4	3, 1, 0	10	4, 0, 0	16
2, 1, 0	5	3, 1, 1	11	4, 1, 0	17
2, 1, 1	6	2, 2, 2	12	4, 1, 1	18

There are other rules for other types of crystals.

15.3. Seeman Spectroscope.—The Bragg and de Broglie methods of obtaining x-ray spectra usually take a fairly large amount of time. Using the de Broglie photographic method as described in Sec. 14.15 it requires about an hour to obtain the K spectrum of molybdenum when the tube is operated at about 50,000 volts and 5 milliamperes. We have mentioned that the focal spot on the target of a Coolidge tube has an appreciable size. Seeman makes use of this size of the spot in his method of obtaining x-ray spectra. The scheme is shown in Fig. 15.5. *AB* represents the target of an x-ray tube and *CD* the focal spot. *EK* is a crystal and *F* is a wedge of lead. Just under the edge of the wedge at *G* x-rays are incident on the face of the crystal at glancing angles between ∠ *CGH* and ∠ *DGH*. If this range of angles includes the angle of reflection for a line of the spectrum of the target, such as the ∠ *LGH*, this line will be strongly reflected in the direction *GM*. If a photographic film is placed at *NP*, a black line appears at *M* when the film is developed. If ∠ *CGD* also includes the directions at which other spectrum lines are reflected by the crystal, black lines corresponding to these other spectrum lines also appear on the film. Using this method, the α and β lines of the *K* spectrum of molybdenum can be obtained with an exposure of under two minutes, the x-ray tube being operated at 45,000 volts and 25 milliamperes. Such a spectrum is shown in Fig. 15.6. Ross used a Seeman spectroscope when he obtained the Compton effect as mentioned in Sec. 14.25. This spectroscope is very useful for obtain-

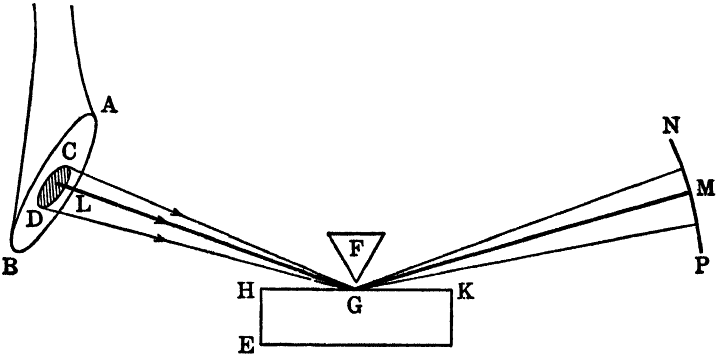


FIG. 15.5.—Showing the principle of the Seeman x-ray spectroscope.

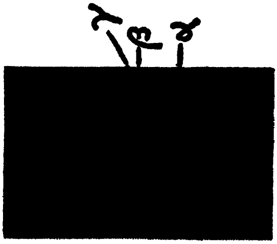


FIG. 15.6.—Molybdenum *K* x-ray spectrum obtained with a Seeman spectroscope in the author's laboratory. X-ray tube operated at 45,000 volts r.m.s. and 25 milliamperes. Target to wedge, 18 cm; wedge to film, 40 cm; exposure to bring out α line, 0.5 min; exposure to bring out γ line, 3 min.

ing the x-ray spectrum of an unknown substance which is placed on the target. The spectrum of the unknown substance is compared with that of a known substance which is recorded on the same film.

15.4. Photographic Effect of X-Rays.—When light falls upon a sensitive photographic film and the film is subsequently developed, the developed film is blackened. The blackness B of the developed film depends upon the product of I, the intensity of the light, and t, the time of exposure of the original film to the light. We may therefore write

$$B = f(It) \tag{15–5}$$

where $f(It)$ represents a function of It. For small values of It, B increases with It. However, beyond a certain value of It, B diminishes as It increases. This phenomenon is known as solarization or photographic reversal. The same kind of phenomenon is found with x-rays. Large values of It sometimes give less blackening than smaller values. Unless this is remembered, strange conclusions may be drawn from x-ray photographs. The center of an x-ray spectrum line under certain conditions may be very intense. Hence this spectrum line may appear on a negative as a narrow black band with a less black line down its middle. When two spectrum lines are compared on the same film it does not necessarily follow that the line with the greater blackness is the line with the greater value of It. In 1940 Baltzer and Nafe, of Washington University, found four reversals (2 maxima and 2 minima) of blackness as It was increased.

In 1934 Jauncey and Richardson found that B in (15–5) is also a function of the temperature of the film when x-rays are used. The value of It to obtain a certain blackening on a film cooled to a temperature of 103° Abs is three times that to obtain the same blackening on a film at room temperature (295° Abs). In certain experiments on the intensity of x-rays it is important for the experimenter to be aware of the effect of the temperature of the photographic film used.

15.5. Vibrations of Crystal Atoms.—One of the very interesting applications of x-rays is to give almost visual proof that the atoms of a solid do vibrate in the way described in Debye's theory of specific heats in Sec. 9.15. In the derivation of Bragg's law in Sec. 14.14, we supposed the centers of the atoms of a crystal to be exactly in planes. However, according to Debye's theory of the specific heats of solids there are elastic waves traveling back and forth in a solid and these produce standing waves or vibrational modes. Consequently the atoms of a crystal are vibrating in and out of the crystal planes and so do not form the exact planes shown in Fig. 14.15. As the temperature of the crystal increases the mean square displacement of the atoms from the planes increases and the planes of atoms become less distinct. This decrease in the distinctness of the crystal planes causes the intensity of the reflected x-rays to diminish. Finally, when the crystal melts, all semblance of planes of atoms disappears and there is then no reflection of x-rays at all, although there is general scattering of x-rays in all directions.

It can be shown from theory that, when x-rays are reflected from a crystal like sylvine (KCl) in a given order of reflection at two temperatures T_1 and T_2, the ratio of the reflected intensities I_1 and I_2 is

$$I_1/I_2 = \exp\left\{-16\pi^2(\overline{z_1^2} - \overline{z_2^2})\cdot(\sin^2\theta)/\lambda^2\right\} \qquad (15\text{–}6)$$

where $\overline{z_1^2}$ and $\overline{z_2^2}$ are the mean square displacements of atoms from planes at the temperature T_1 and T_2. From measurements on potassium chloride, James and Brindley have found that $(\overline{z_1^2} - \overline{z_2^2}) = .0140\ \text{A}^2$ for $T_1 = 290°$ Abs and $T_2 = 86°$ Abs.

According to (3–27), the mean square displacement of $\overline{u^2}$ of the particles or atoms of a solid which is in a vibrational mode of frequency ν is

$$\overline{u^2} = W/4\pi^2\nu^2M \qquad (15\text{–}7)$$

where W is the energy of the solid and M its mass. However, this displacement $\overline{u^2}$ is in any direction. For x-ray reflection it is only the mean square displacement $\overline{z^2}$ perpendicular to a set of planes of atoms which is important. In a cubic crystal we have

$$\overline{u^2} = \overline{x^2} + \overline{y^2} + \overline{z^2} \qquad (15\text{–}8)$$

where $\overline{x^2}$, $\overline{y^2}$, and $\overline{z^2}$ represent the mean square component displacements in the x, y, and z directions respectively. Also, since in a cubic crystal the displacement of an atom is just as likely in one direction as another, we have

$$\overline{x^2} = \overline{y^2} = \overline{z^2} = \overline{u^2}/3 \qquad (15\text{–}9)$$

Hence from (15–7)

$$\overline{z^2} = W/12\pi^2\nu^2M \qquad (15\text{–}10)$$

But, according to (9–46) of the chapter on specific heats, the energy of the vibrational modes in a gram-atom of a solid in the frequency range ν to $\nu + d\nu$ is

$$dW = 4\pi v\left(\frac{1}{V_L{}^3} + \frac{2}{V_T{}^3}\right)\frac{h\nu^3 d\nu}{e^{h\nu/kT} - 1} \tag{15–11}$$

If we replace W in (15–10) by dW as given by (15–11), we obtain the mean square displacement $\overline{z^2}$ which is due to the vibrational modes in the frequency range ν to $\nu + d\nu$. Integrating, we then obtain the mean square displacement $\overline{z^2}$ for all vibrational modes. This is

$$\overline{z^2} = \frac{v}{3\pi M}\left(\frac{1}{V_L{}^3} + \frac{2}{V_T{}^3}\right)\int_0^{\nu_m} \frac{h\nu}{e^{h\nu/kT} - 1}\, d\nu \tag{15–12}$$

where, in the case of a crystal such as potassium chloride, v is the volume of a gram-molecule and M is the molecular weight. Just as in the case of the theory of specific heats we again integrate from $\nu = 0$ to ν_m. Also we introduce the characteristic temperature T_c, defined by (9–49) and (9–50). Putting $x = h\nu/kT$, we obtain

$$\overline{z^2} = \frac{3nh^2T_c^2}{4\pi^2MkT_c{}^3}\int_0^{T_c/T} \frac{x}{e^x - 1}\, dx \tag{15–13}$$

where M is the molecular weight of the crystal and n is the number of atoms in a gram-molecule. For sylvine $n = 2N$, where N is Avogadro's number. The integral on the right side of (15–13) is a complicated function of T_c/T. We shall write $(T/T_c)^2$ times this complicated function as $f(T_c/T)$ and so (15–13) becomes

$$\overline{z^2} = \frac{6Nh^2}{4\pi^2MkT_c}f(T_c/T) \tag{15–14}$$

for a crystal like sylvine. The right hand side of (15–14) should contain a term due to the energy of vibration which, according to recent theory, exists at absolute zero; but, if we consider the difference of the $\overline{z^2}$ at two temperatures, this term cancels out and we have

$$\overline{z_1{}^2} - \overline{z_2{}^2} = \frac{6Nh^2}{4\pi^2MkT_c}\left\{f\left(\frac{T_c}{T_1}\right) - f\left(\frac{T_c}{T_2}\right)\right\} \tag{15–15}$$

For potassium chloride the characteristic temperature $T_c = 230°$ Abs. This gives theoretical molecular heat values which agree excellently with experimental values at different temperatures. A table of values of

(T_c/T) is shown in Table 15.2. Applying (15–15) to potassium chloride at 290° Abs and 86° Abs we obtain $(\overline{z_1^2} - \overline{z_2^2}) = 0.0138\ \text{A}^2$ in excellent agreement with the experimental value obtained from x-ray reflection. This agreement gives confirmation of the ideas in Debye's theory of specific heats.

TABLE 15.2

T_c/T	$f(T_c/T)$	T_c/T	$f(T_c/T)$	T_c/T	$f(T_c/T)$
0.0	∞	0 8	1.02	1.6	0.417
.2	4.75	1 0	0.778	1.8	.353
.4	2.26	1.2	617	2.0	.303
.6	1.43	1 4	·03	2.5	.216
				3.0	.159

15.6. Diffuse Reflection of X-Rays.—On carefully examining a Laue picture obtained by passing a narrow beam of x-rays containing a continuous spectrum of wave-lengths through a thin slab of crystal, it is seen that there is a slight general blackening in addition to the black Laue spots. This general blackening is caused by *diffuse scattering* of x-rays by the crystal. We shall not discuss the diffuse scattering further. However, the general blackening is accompanied by diffuse spots and streaks which are blacker than the general background of blackening and less black than the Laue spots. These diffuse spots and streaks were obtained on a photograph by Friedrich of Germany soon after the initial discovery of the diffraction of x-rays by crystals in 1912. It was believed that the diffuse spots and streaks were caused by imperfections in the particular crystal used. Moreover, many Laue spot pictures did not show the diffuse spots and streaks. In 1919 Faxen of Sweden called attention to the diffuse spots and streaks in Friedrich's picture and said that they may be caused by the temperature vibrations of the crystal. The matter then lay dormant until 1938, when Laval of France and Wadlund of Trinity College, Connecticut, rediscovered the diffuse spots and streaks. Wadlund found the same pattern of diffuse spots and streaks for specimens of rocksalt from different sources. Consequently, the diffuse spots and streaks were not caused by the imperfection or strain of a particular crystal. Laval, using an electrometer and an ionization chamber, made an exhaustive experimental study of the x-rays diffracted from the front face of a crystal of sylvine (KCl) when it is turned away from the Bragg position for reflection of monochromatic

Cu $K\alpha$ or Mo $K\alpha$ x-rays. In order to obtain the diffuse spots and streaks photographically it is sometimes necessary to make an exposure so long that the true Laue spots show photographic reversal. The experimenter who is studying Laue spots is always careful to have exposures short enough to avoid photographic reversal and consequently he would not obtain obvious diffuse spots and streaks.

According to Debye's theory of specific heats there is a certain number of vibrational modes whose wave-lengths are between λ and $\lambda + d\lambda$. These are wave-lengths of the elastic waves in the crystal. Hence, in order to avoid confusion with λ, the wave-length of x-rays, we shall use L for the wave-length of an elastic (or sound) wave in the crystal. A vibrational mode is produced by the interaction of several progressive waves of wave-length L traveling in different directions. The velocity of these elastic waves is of the order of 5×10^5 cm/sec. This is very small compared with 3×10^{10} cm/sec, the velocity of light or x-rays. For simplicity we shall limit ourselves to the longitudinal waves in a crystal. The progressive waves of wave-length L consist of rarefactions and condensations of crystal atoms. With respect to the fast-traveling x-rays, these rarefactions and condensations may be considered as standing still. They constitute a set of planes for the reflection of x-rays according to Bragg's law, where d is replaced by L. However, L is usually so large that, for $n = 1$ in Bragg's law, $\sin \theta$ is so small and, consequently, the reflected beam makes such a small glancing angle with the reflecting planes of condensation that we cannot expect any easily observable reflection from these planes.

The crystal planes with a grating space d and the condensation planes with a grating space L combine to form a set of planes with a grating space δ, given by

$$1/\delta = 1/d + 1/L \qquad (15\text{–}16)$$

But this is a vector equation, the directions of $1/d$ and $1/L$ being perpendicular to the crystal planes and the condensation planes, respectively. If $1/L$ is perpendicular to $1/d$, (15–16) becomes

$$1/\delta = \sqrt{1/d^2 + 1/L^2}$$

or

$$\delta = d(1 + d^2/L^2)^{-1/2} = d[1 - (1/2)d^2/L^2] \qquad (15\text{–}17)$$

by the binomial theorem when d^2/L^2 is small. Since, for the (1,0,0) planes of rocksalt, $d = 2.82$ A and L is of the order of 50 to 100 A, δ

very nearly equals d. But, since $1/L$ is perpendicular to $1/d$, the δ-planes make an angle ϕ with the d-planes such that

$$\tan\phi = (1/L)/(1/d) = d/L \tag{15-18}$$

Hence ϕ may have values between 1.6° and 3.2°.

The interest of physicists in diffuse reflection of x-rays was world-wide. Zachariasen of Chicago, Raman of India, Kathleen Lonsdale of London, England, all made valuable contributions to the study of this phenomenon. Fig. 15.7 is the reproduction of a photograph obtained by Jauncey and Baltzer of Washington University, St. Louis in 1941. X-rays from a tube with a copper target were used. These rays consist of two very intense spectrum lines (Cu $K\alpha$, 1.54 A; Cu $K\beta$, 1.39 A) with a weak background of wave-lengths of the continuous spectrum. Second order reflections of the α and β lines from the (1,0,0) planes of rocksalt (d = 2.82 A) occur at $\theta = 33°$ and 29.5°, respectively. In Strip 1 of Fig. 15.7 the crystal was first set at $\theta = 33°$ and an exposure of 1 min made. This gave the white line shown at α. The crystal was then set at $\theta = 36°$ and an exposure of several hours was made. Second-order reflection of x-rays belonging to the continuous spectrum is shown by the spot CL of Strip 1. At the same time a diffuse reflection of the $K\alpha$

Fig. 15.7.—Three examples of diffuse reflection of x-rays obtained in the author's laboratory. Copper K radiation with a weak background of continuous radiation was incident at glancing angles of 36°, 31°, and 28° in Strips 1, 2, and 3, respectively, on the (1,0,0) planes of rocksalt. The narrow white lines on the extreme left of each strip represent first order reflection of KA rays from the (1,0,0) planes of the crystal (θ = 16°, nearly).

rays from the δ-planes produced the diffuse spot shown as α in Strip 1. In Strip 2 the crystal was set at $\theta = 31°$, which is between 29.5° and 33°. In this case the 1 min exposure at $\theta = 33°$ was omitted. The reflection by the d-planes of the appropriate wave-length is again shown as *CL*, but now we have diffuse reflections of both the $K\alpha$ and $K\beta$ rays from the δ-planes. In Strip 3, the crystal was set at $\theta = 28°$. In addition to the rays of the continuous spectrum reflected by the d-planes and producing *CL*, we note the diffuse reflections of the $K\alpha$ and $K\beta$ rays from the δ-planes. In this last case the diffuse α spot is much less intense than the diffuse β spot. Also the α spot is broader than the β spot.

It is convenient to speak of the off-set angle. In Strip 1 the off-set angle is $36° - 33° = 3°$ for the $K\alpha$ rays. As the magnitude of the off-set angle is increased the diffuse spot becomes less intense and broader. There is no evidence of a diffuse β spot in Strip 1, where the off-set angle is $36° - 29.5° = 6.5°$. Also the $K\beta$ spectrum line is much weaker than the $K\alpha$ spectrum line.

15.7. Medical Diagnosis.—One of the first and most important applications of x-rays was to medical diagnosis. Since the bones of the body are more dense than the fleshy portion of the body and also since they contain calcium (At. No. 20), they have a greater absorption coefficient for x-rays than does the fleshy portion of the body. Hence, when x-rays are passed through the human body on to a photographic plate, the developed film shows black where the x-rays have passed through the fleshy portions and shows clear where the x-rays have been absorbed by the bones of the body. The fleshy portion of the body is not only less dense than the bones but also consists of elements of lower atomic number than calcium so that the absorption coefficient is small.

Frequently it is necessary in medical practice to diagnose the intestines and other fleshy portions of the body. In this case there is not sufficient difference of density and atomic number from the density and atomic number of the surrounding tissue to obtain sufficient contrast for a good photograph. In order to obtain this contrast, it is the practice to give the patient a "barium" or "bismuth" meal. A meal consists of a quart of milk in which barium sulphate or bismuth carbonate has been mixed. The atomic numbers of barium and bismuth are 56 and 83 respectively, so that the absorption coefficients of the barium sulphate and bismuth carbonate powders are quite high. Some time after the meal has been taken these powders settle in the intestines, and, if a picture is taken at this time, the intestines stand out in contrast to the surrounding tissue, and a diagnosis can be made. Methyl iodide is injected into other parts of the body in order that x-ray photographs of these parts may show contrast. The formula for methyl iodide is CH_3I and the atomic number of iodine is 53. The average atomic number of the surrounding tissue is about 6 to 8.

15.8. Therapeutics.—Physicians not only use x-rays for diagnosis but also for their curative properties. The effect of x-rays on living material is destructive, the effect of prolonged exposure to x-rays being to kill tissue. X-rays upon absorption in the tissue give rise to photoelectrons which begin with an energy of perhaps 100,000 ev or more, depending on the voltage across the x-ray tube. These electrons pass through the atoms of the molecules of the surrounding tissue and knock out electrons from these atoms. This tends to bring about a change in the structure of the molecules, so that after some time a biological effect is observed.

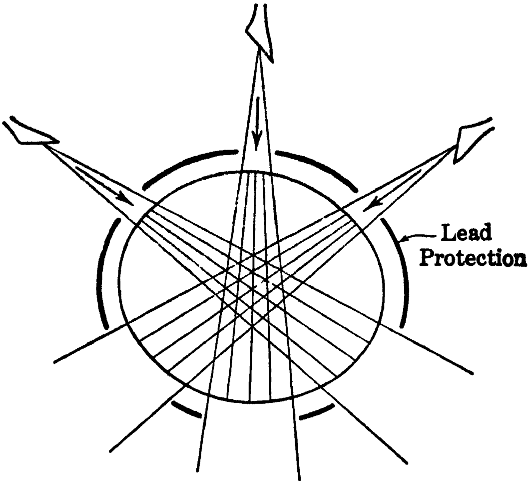


Fig. 15.8.—Cross-fire method of the treatment of abnormal tissue with x-rays.

One trouble with x-ray therapy is that x-rays not only kill abnormal but normal tissue as well. Since the absorption of x-rays is an atomic process, there seems to be no *a priori* reason why x-rays should be absorbed preferably in the abnormal tissue. In the case of the abnormal tissue being deeply embedded, the x-rays must pass through normal tissue before they arrive at the abnormal tissue. The only way of killing the abnormal tissue faster than the normal tissue seems to be to devise some scheme for irradiating the abnormal tissue with more intense x-rays than those which pass through the normal tissue. One such scheme is indicated in Fig. 15.8. This is known as the cross-fire method. In Fig. 15.8 x-rays from three different tubes cross at the location of the abnormal tissue. The killing effect of the rays should take place three times as rapidly in the abnormal tissue as in the normal tissue through which each beam passes. In some cases in practice there are as many as eight beams crossing at the location of the abnormal tissue. Possibly another scheme might be to impregnate the abnormal tissue with some element of high atomic number, but the author knows of no such method being used in practice.

15.9. Measurement of Dosage.—The biological effect of x-rays depends on the voltage across the x-ray tube, the nature of the target of the tube, the distance of the target from the tissue to be treated, and the time of exposure of the tissue to the x-rays. The intensity of x-rays arriving at a distance d from the target of an x-ray tube is approximately given by

$$I = \text{const} \times Zi\,\frac{V^2}{d^2} \tag{15-19}$$

where Z is the atomic number of the target, i the current (usually in milliamperes) through the tube, V the voltage across the tube, and the constant is a constant of proportionality. The intensity is the rate at which x-ray energy is arriving at a certain place, whereas the total energy arriving in a given time is It where t is the time. The biological effect is proportional to It. The term dosage as used by physicians corresponds to what we have called the total energy arriving during a given time. Hence the dosage

$$D = \text{const} \times ZiV^2t/d^2 \tag{15-20}$$

It is very important to measure the dosage given to a patient, so that the correct amount shall be given to him. It is therefore necessary that some kind of a unit of dosage should be devised.

The first unit to be used is known as the Erythema Skin Dose which is abbreviated to E. S. D. The E. S. D. is that amount of dose which produces an erythema or a reddening of the human skin. Unfortunately, the erythema does not appear immediately upon the skin of the patient being given one E. S. D. The reddening of the skin does not occur until several days afterward. According to one authority, the E. S. D. is that quantity of x-rays which will cause a slight hyperemia in five days, while, according to another authority, the E. S. D. is that quantity which will produce a slight reddening of the skin after eight days and a slight bronzing after fourteen days. The phrases "slight hyperemia" and "slight bronzing" are purely qualitative and not quantitative. There is no method of quantitative measurement. Then also there is considerable difference between individuals. There is thus very great difficulty in determining the E. S. D. accurately. As a result of its unsatisfactory nature, the E. S. D. has more or less been abandoned in favor of some physical method of measuring dosage.

The most recent method of measuring dosage is by means of the saturation current in an ionization chamber. Duane of Harvard has defined a unit of intensity of x-rays as that which can produce one electrostatic unit of current in each cubic centimeter of air at 0° C and a pressure of 76 cm through which it passes, provided that the current has its saturation value. This is a unit of intensity and not of dosage and it is denoted by the letter E. If x-rays whose intensity is one E unit operate for one second on tissue, the dosage is one ES unit. Hence, x-rays whose intensity is, say, 10 E units, and which operate for 5 seconds give a dosage of 10×5 ES units $= 50$ ES. The ES unit is now called the roentgen and is designated by the symbol r. The biological effect of x-rays is proportional to the number of roentgens falling upon the tissue. The E. S. D. is variously estimated as being from 700 to 1500 r.

If in (15–20) i is in milliamperes, V is in kilovolts, t in minutes and d in inches, then D is measured in roentgens when the constant is about 1.1×10^{-8}.

In deep therapy work, the scattering of x-rays must be taken into account. As x-rays pass on through tissue they are scattered in all directions. Less and less x-rays travel in the original direction of the primary x-rays, the situation being shown in Fig. 15.9. The primary rays travel downward in the body of the patient. Some of these x-rays are scattered by the electrons in the atoms and after a depth of a few centimeters has been traversed by the rays a large fraction of the rays are traveling in the horizontal and even in the vertically upward directions. In Fig. 15.8, depicting the cross-fire method of treating abnormal tissue, there is scattering of the rays in the various beams both before and after the abnormal tissue is reached. These scattered rays act for the most part on the normal tissue.

It is because of the presence of scattered x-rays that x-ray photographs of the bones in the thicker parts of the body are less distinct than in the thinner parts of the body. For instance, it is very easy to obtain a distinct picture of the bones of the hand but much more difficult to obtain a distinct picture of the vertebrae of the spine. When x-rays pass through a considerable thickness of tissue, scattering occurs, and the shadows cast by the bones become less distinct.

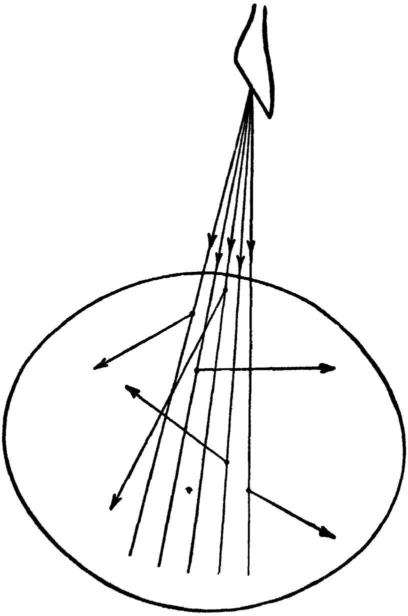


FIG. 15.9.—Illustrating the scattering of x-rays by the tissue through which the rays pass.

CHAPTER XV

PROBLEMS

1. Sylvine (KCl) is a cubic crystal. Since the atomic weights of potassium and chlorine are nearly equal, being 39.1 and 35.45 respectively, the crystal of sylvine may be treated as being made up of atoms of one kind so that sylvine may be taken as a simple cubic crystal. The value of the principal grating space of sylvine is 3.14 A. Find the glancing angles at which the α, β, and γ lines of the K x-ray spectrum of molybdenum are reflected in the third order from (*a*) a face cut parallel to the (1, 1, 0) planes, and (*b*) a face cut parallel to the (1, 1, 1) planes.

2. The metal lithium crystallizes in the body-centered cubic system. The value of its principal grating space, d_{100}, is 1.75 A. Determine the glancing angles for the (1, 1, 1) and (3, 1, 0) reflections of Mo $L\beta_1$ rays from a powdered crystal of metallic lithium. Determine the radii of the corresponding circles in Fig. 15.4 when the distance of the irradiated portion of the powdered crystal from the film is 5 cm.

3. The gas neon solidifies at −249° C and crystallizes in the face-centered cubic system. The value of its principal grating space at the temperature of liquid helium (−268.9° C) is 2.26 A. Determine the glancing angles for the (1, 1, 1), (3, 1, 0), (5, 1, 1) reflections of the Ag $K\beta$ rays from a powdered crystal of solid neon at −268.9° C.

4. A patient is exposed for 10 minutes to x-rays. The x-ray tube is operated at 100 kilovolts and 20 milliamperes, the target being tungsten and the distance of the focal spot on the target from the patient being 3 ft. A tube with a tin target is substituted and operated at 75 kilovolts and 15 milliamperes, the distance from the focal spot to the patient being 4 ft. Find the time of exposure so as to obtain the same dosage as when the first tube was used. The atomic numbers of tin and tungsten are 50 and 74 respectively.

5. A crystal of sodium fluoride (NaF) has a characteristic temperature of 441° Abs. Determine the change in the mean square departure of the atoms from the crystal planes when the temperature changes from 200° Abs to 350° Abs. Also determine the change in the intensity of Mo $K\alpha$ x-rays reflected in the fifth order from the (1, 0, 0) planes, given that $d_{100} = 2.47$ A. Assume the crystal structure of sodium fluoride to be similar to that of sylvine.

6. A molybdenum target tube is operated at 70 kv and 10 ma. Determine the dosage in roentgens received by a patient at 24 inches distance and in 5 minutes.

CHAPTER XVI

BOHR THEORY OF SPECTRA

16.1. X-Ray and Optical Spectra.—The wave-lengths of the lines in the *K* x-ray spectra of several elements are given in Table 14.2. The *K* spectra for iron, copper, molybdenum and tungsten are represented in Fig. 16.1, where the spectra of the different elements are drawn to scale.

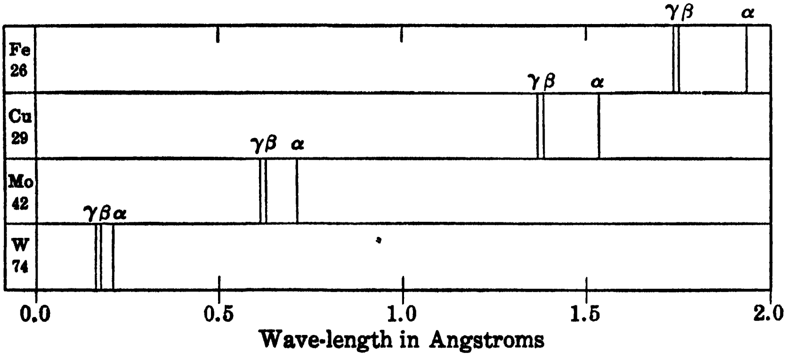


Fig. 16.1.—*K* x-ray spectra for several elements.

It is seen that the spectra are all similar, although the wave-lengths differ. The wave-lengths of the lines in the *L* x-ray specta of three elements are given in Table 14.3. The *L* spectra of these three elements are represented in Fig. 16.2. It is seen that there is a similarity among the *L* spectra, although the similarity is not as great as among the *K* spectra in Fig. 16.1.

We now come to optical spectra. There is no obvious similarity between the optical spectra of different elements. The spectrum of one element not only differs from that of a second element in that the wave-lengths of its lines are different from those of the lines of the second element, but also in that the arrangement of the lines in the spectrum of the one element differs from that of the lines of the second element. The more prominent lines of the optical spectra of several elements are represented in Fig. 16.3. The tremendous difference between the single line spectrum of sodium (Note: this line is a doublet when examined with a spectroscope of high resolving power) and the many-line spectrum of iron is obvious. The spectra which are represented in Fig. 16.3 are in the visible region. An example of a spectrum in the ultraviolet is given in Fig. 16.4.

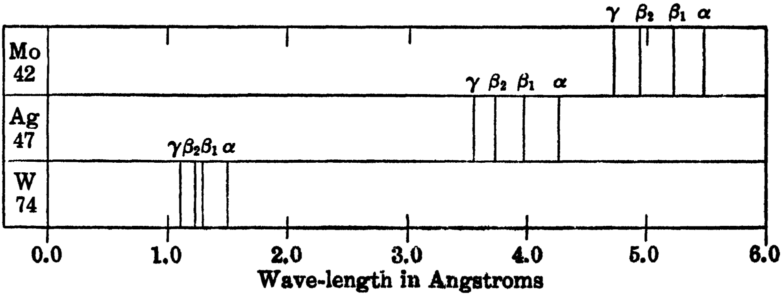


FIG. 16.2.—*L* x-ray spectra for several elements.

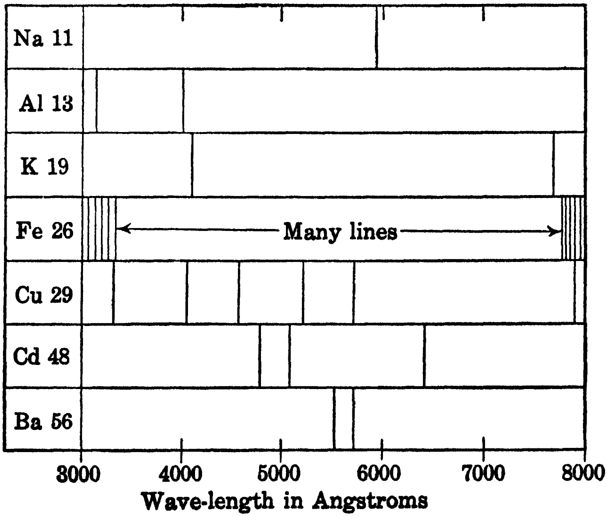


FIG. 16.3.—Prominent lines in the optical spectra of several elements.

It will be remembered that, in the case of sound, overtones are produced when an organ pipe is sounded or a stretched string is set vibrating. A stretched string can vibrate as a whole, in halves, thirds, quarters and so on, and the overtones have frequencies which are 2, 3, 4, etc., times the frequency of the fundamental tone, or the wave-lengths of the overtones in air are $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, etc., of the wave-length of the fundamental. The fundamental and overtones may be said to constitute a sound spectrum. The frequencies of the overtones of a closed organ pipe are 3, 5, 7, etc., times the frequency of the fundamental note of the pipe, and the wave-lengths of the overtones in air are $\frac{1}{3}$, $\frac{1}{5}$, $\frac{1}{7}$, etc., of the wave-length of the fundamental. Sound spectra of the stretched string and the closed organ pipe are represented in Fig. 16.5. It is seen at once that there is an order to the arrangement of the lines in each spectrum. According to the classical electromagnetic wave theory of light, an optical spectrum is produced by electrical vibrations of the atoms which are emitting the spectrum. When the optical spectrum of any given element consists of more than one line, it seems likely that some of the lines may be overtones of other lines, as in the sound spectra of Fig. 16.5. If there are relations of this type connecting certain optical spectrum lines of an element, it should be possible to express the wave-length of each of the lines so connected by some kind of a formula. In the case of the sound spectrum of a closed organ pipe the formula is

$$\lambda = \frac{\lambda_0}{(2n - 1)} \tag{16-1}$$

where λ_0 is the wave-length of the fundamental and n is an integer. When $n = 1$, $\lambda = \lambda_0$ and we have the fundamental. When $n = 2$, $\lambda = \lambda_0/3$ and we have the first overtone, and so on. If light is at all similar to sound, we might expect that the

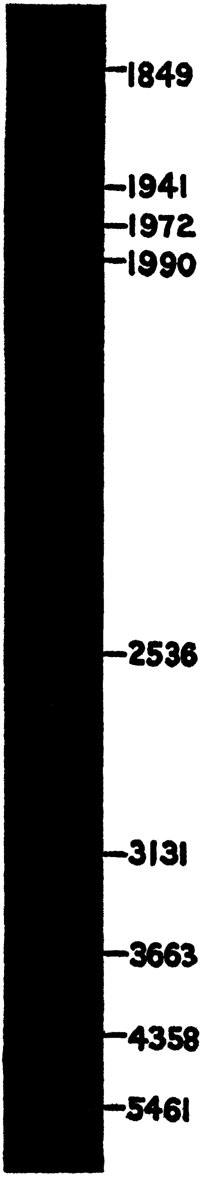


FIG. 16.4.—Photograph of the ultraviolet spectrum of mercury taken by Dr. C. F. Hagenow at Washington University. The numbers give the wave-lengths of the various lines in angstroms. Wave-lengths shorter than about 3800 A are invisible. Since ordinary glass is opaque to ultraviolet rays, a quartz mercury arc lamp and a quartz spectrograph were used. The photograph was made with a panchromatic plate treated with a mineral oil (nujorl). The oil fluoresces and brings out the lines below 2400 A. Exposure about 1 min.

formula connecting the wave-lengths of different lines in a given spectrum would contain an integer n. Such formulas were sought in vain until 1885.

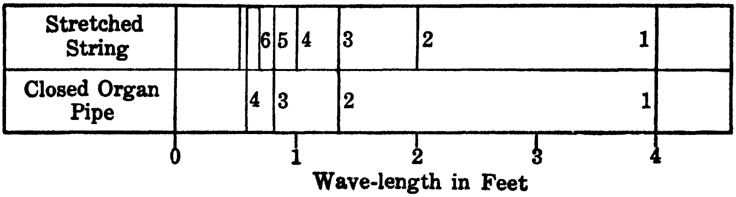


FIG. 16.5.—Sound spectra of the fundamental and overtones of a stretched string and of a closed organ pipe, wave-length in air of the fundamental being 4 feet.

16.2. Balmer Series Spectrum of Hydrogen.—Under proper conditions hydrogen can be caused to emit the spectrum which is represented in Fig. 16.6. It is seen that there appears to be a definite arrangement of the lines, the spectrum being somewhat similar to the sound spectra

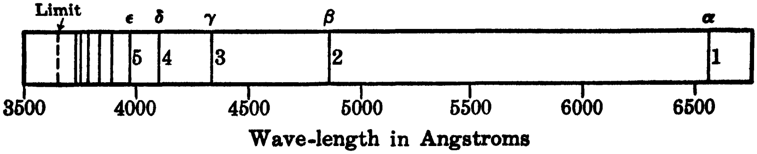


FIG. 16.6.—The Balmer series spectrum of hydrogen.

shown in Fig. 16.5. In 1885 Balmer of Switzerland found that the wave-lengths of the lines shown in Fig. 16.6 could be expressed by means of the formula

$$\lambda = 3645 \frac{n^2}{(n^2 - 4)} \qquad (16\text{–}2)$$

where λ is in angstroms and n is a whole number equal to or greater than 3. Thus, when $n = 3$, the wave-length 6560 A of the α line is obtained. When $n = 4$, the wave-length of the β line is obtained, and so on. When a set of spectrum lines can be connected together by a formula such as (16–2), the set is said to constitute a series. The lines in Fig. 16.6 constitute what is known as the Balmer series. In recent years a total of 31 lines have been obtained in the Balmer series by observing the spectra of certain nebulae and of the solar corona at a total eclipse.

16.3. Wave-Number.—It was soon found that it was more convenient to write the Balmer series in terms of the frequency ν than in terms of the wave-length λ. This is possible by reason of the relation

$$\nu = \frac{c}{\lambda} \tag{16–3}$$

where c is the velocity of light. Unfortunately, this gives us inordinately large numbers, and, since c is a constant and ν is therefore inversely proportional to λ, it has become the practice to use $1/\lambda$ instead of the frequency. The reciprocal of the wave-length is the number of wave-lengths in a centimeter or the number of waves in a centimeter. Hence $1/\lambda$ has come to be called the *wave-number.* Unfortunately, the wave-number is also represented by ν, so that $\nu = 1/\lambda$, where ν is the wave-number. In this book we shall distinguish between frequency and wave-number by priming the ν, so that it becomes ν' when it indicates wave-number. The relation between wave-number ν' and wave-length λ is then

$$\nu' = \frac{1}{\lambda} \tag{16–4}$$

and the relation between frequency ν, wave-number ν' and the velocity of light c is

$$\nu = c\nu' \tag{16–5}$$

In terms of the wave-number ν' Balmer's formula (16–2) becomes

$$\nu' = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \tag{16–6}$$

where R is an important constant known as the Rydberg constant and n is a whole number equal to or greater than 3. The value of the Rydberg constant is

$$R = 109{,}700 \text{ cm}^{-1} \tag{16–7}$$

When $n = 3$ we obtain the α line of the series. Since there are no lines on the long wave-length side of this line, it is known as the *head* of the series. When $n = \infty$ we obtain what is known as the *limit* of the series. This is shown as a broken line in Fig. 16.6. The wave-length of the head of the series is 6560 A and that of the limit is 3670 A. There is no line of shorter wave-length than that of the limit of the series.

16.4. Other Series Spectra.—Since Balmer discovered the series which bears his name, three other series spectra for hydrogen have been discovered. The Lyman series is in the ultraviolet, the Paschen series is in the infrared, and the Brackett series is in the far infrared. The formulas for these three series are

Lyman series

$$\nu' = R\left(\frac{1}{1^2} - \frac{1}{n^2}\right), n = 2, 3, 4 \cdots \tag{16-8}$$

Paschen series

$$\nu' = R\left(\frac{1}{3^2} - \frac{1}{n^2}\right), n = 4, 5, 6 \cdots \tag{16-9}$$

Brackett series

$$\nu' = R\left(\frac{1}{4^2} - \frac{1}{n^2}\right), n = 5, 6, 7 \cdots \tag{16-10}$$

where R is the Rydberg constant.

A series spectrum for ionized helium has been discovered which has the formula

$$\nu' = 4R\left(\frac{1}{3^2} - \frac{1}{n^2}\right), n = 4, 5, 6 \cdots \tag{16-11}$$

It seems, therefore, that we are able to express the wave-numbers of spectrum lines in terms of the Rydberg constant and of whole numbers.

By comparison of Figs. 16.1 and 16.6, it is seen that the K x-ray spectrum of an element contains three lines which are arranged in a somewhat similar way to the α, β and γ lines of the Balmer spectrum of hydrogen. The approximate statement of Moseley's law is that the wave-number (or frequency) of a line in an x-ray spectrum is proportional to the square of the atomic number of the element emitting the line. The helium spectrum whose formula is given in (16–11) contains the factor $4R$. But $4 = 2^2$ and 2 is the atomic number of helium. Further, the Lyman, Balmer, and Paschen series spectra of hydrogen may correspond to the K, L, and M x-ray spectra of elements of high atomic number. Does the α line of the K x-ray spectrum correspond to the α line of the Lyman series of hydrogen? If our idea is correct, the

K spectrum of any element should be given by

$$\nu' = Z^2R\left(\frac{1}{1^2} - \frac{1}{n^2}\right), n = 2, 3, 4 \cdots \qquad (16\text{–}12)$$

where R is the Rydberg constant and Z is the atomic number, the wave-number of the α line of the K spectrum being obtained when $n = 2$ in (16–12). Hence the wave-number of the α line of the K spectrum of molybdenum (At. No. 42) is

$$\nu' = 42^2 \times 109{,}700 \times \left(1 - \frac{1}{2^2}\right) = 1.45 \times 10^8 \text{ cm}^{-1}$$

This corresponds to a wave-length of 1/145 = 0.690 A. The wave-length of the $K\alpha$ line of molybdenum is 0.710 as seen in Table 14.2. This is an excellent agreement, considering that the α line of the Lyman spectrum of hydrogen has a wave-number of $109{,}700 \times (1 - \frac{1}{4}) = 82{,}200 \text{ cm}^{-1}$ or a wave-length of 1223 A and that the atomic numbers of hydrogen and molybdenum are 1 and 42, respectively. It is remarkable that a formula such as (16–12) should hold so well over such a large range.

If the L x-ray spectra correspond to the Balmer series spectrum of hydrogen, these spectra should be described by the formula

$$\nu' = Z^2R\left(\frac{1}{2^2} - \frac{1}{n^2}\right), n = 3, 4, 5 \cdots \qquad (16\text{–}13)$$

The α line of the L series should correspond to the α line or head of the Balmer series. The wave-number of the $L\alpha$ lines of molybdenum should therefore be

$$\nu' = 42^2 \times 109{,}700 \times \left(\frac{1}{2^2} - \frac{1}{3^2}\right) = 0.269 \times 10^{+8} \text{ cm}^{-1}$$

This corresponds to a wave-length of 1/0.269 = 3.72 A. The wave-length of the $L\alpha$ line of molybdenum is 5.40 A. The agreement in this case is not so good as in the previous case of the $K\alpha$ line of molybdenum. It will be remembered that, although there is only one excitation potential for the K x-ray spectrum of a given element, yet there are three excitation potentials for the L x-ray spectrum of a given element. This suggests that the L spectrum is made up of three separate and overlapping spectra which we may designate by L_1, L_2 and L_3. The reason for much of the complication of the optical spectra of certain elements is due to the overlapping of several series spectra. For instance the spectrum of lithium is represented by the spectrum *T* in Fig. 16.7. In this figure the lines are plotted against wave-number instead of against wavelength. There is at first sight no formula which gives the wave-numbers of the various lines. However, it is possible to pick out sets of lines which are arranged in a manner similar to the Balmer series. In the case of lithium there are four such sets as shown by *P*, *S*, *D*, and *F*. These are the component spectra which make up the total spectrum *T*.

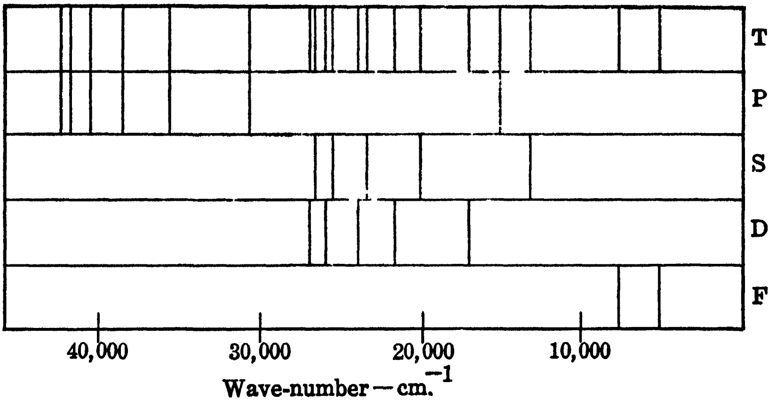


FIG. 16.7.—Optical spectrum of lithium.

Rydberg showed that, when the optical spectrum of a given element can be broken up into several component spectra each of which is similar in arrangement to the Balmer series, each of these component series can in many cases be expressed by a formula

$$\nu' = R\left\{\frac{1}{(a+m)^2} - \frac{1}{(b+n)^2}\right\} \qquad (16\text{–}14)$$

where *R* is the Rydberg constant, *m* and *n* are whole numbers, and *a* and *b* are constants, not necessarily whole numbers, for a given component series. As *n* is given different values while *m* is kept constant, the wave-numbers of different lines in a given component series are obtained.

16.5. Ritz Combination Principle.—The general series formula (16–14) can be written

$$\nu' = \frac{R}{(a+m)^2} - \frac{R}{(b+n)^2} \qquad (16\text{–}15)$$

or in the form

$$\nu' = A - B \tag{16-16}$$

This led Ritz to state the principle that the wave-number (or frequency) of a spectrum line is always given by the differences of two terms. It will be noticed that all of the formulas for series spectra which have been given in this chapter can be expressed in the form of (16–16).

16.6. Classical Atom Models.—Previous to 1913 various attempts had been made to devise theoretical atoms which would vibrate in overtones so as to give the various series spectra which are observed. These theoretical atoms—or models—were made up of positive and negative electricity. If such an atom is distorted, it vibrates and in so doing gives out radiation according to the classical electro-magnetic theory as described in Chapter VI. Thomson and Rutherford had each devised model atoms. The Thomson model atom consisted of a large sphere of positive electricity in which were embedded the negative electrons. The Thomson model was superseded by the Rutherford model, in which the positive electricity is concentrated in a massive nucleus at the center of the atom and the negative electrons revolve about the nucleus in a manner somewhat similar to the motion of the planets around the sun in the solar system. In the Rutherford theory, an atom of an element of atomic number Z contains Z electrons, each carrying a charge $-e'$, and a nucleus carrying a charge $+Ze'$, so that the atom as a whole carries no charge. Practically the entire mass of the atom is concentrated in the nucleus.

There is a theorem in electrostatics known as Earnshaw's theorem which states that a system of electrical charges cannot be in equilibrium under the action of the electrical forces between the charges alone. Accordingly, a system of electrons around a positive nucleus cannot be in equilibrium unless centrifugal force, which is not electrical, is brought into play by the electrons revolving about the nucleus. But, if an electron is revolving in a circular path about a massive positive nucleus, the electron is moving with acceleration v^2/r toward the nucleus, where v is the speed of the electron in its orbit and r is the radius of the orbit. We have seen in Sec. 6.4 that an electron moving with acceleration f should radiate energy at a rate of $2e'^2f^2/3c^3$ erg/sec according to the classical electromagnetic theory. Hence, the revolving electrons in the Rutherford atom should be continually radiating energy. Consequently, unless energy is continually supplied to the atom, the atom will lose energy and the electrons will travel more slowly and so will not have enough velocity to supply a centrifugal force (mv^2/r) equal to the attractive force of the nucleus. The electron will therefore be pulled in toward the nucleus and so will be speeded up and will move in an orbit of smaller radius. But in the orbit of smaller radius the acceleration will be greater and energy will be radiated at a still greater rate. The process is then repeated. The result of this process continually occurring is that the orbit will be a spiral. Each electron of an atom thus follows a spiral path into the nucleus and in this way the atom becomes annihilated. However, a remarkable property of atoms is that they persist and do not disappear. We are thus in a dilemma—we cannot arrange a system of positive and negative charges at rest so as to be in equilibrium and, if we supply centrifugal force by making the charges revolve about a center so as to produce equilibrium, this equilibrium is transient due to the radiation of energy by an accelerated charge. In either case the atom is finally annihilated.

16.7. Bohr's Theory.—In 1900 Planck announced the quantum theory of radiation. He was led to this theory because all attempts to explain the distribution of the energy among the wave-lengths of the continuous spectrum given out by a hot black body by means of the classical theory of light had failed. In the quantum theory it is supposed that radiation of frequency ν is emitted in quanta, the energy of each quantum being $h\nu$, where h is Planck's constant. In 1905 Einstein extended the idea of quanta to the radiation itself. Light, according to Einstein, consists of quanta or photons $h\nu$ which may or may not hit an atom. If an atom is hit by a quantum or photon, the photoelectric effect will occur when, in the equation

$$(\tfrac{1}{2})mv^2 = h\nu - W \tag{16-17}$$

W is less than $h\nu$. In 1913 Bohr of Copenhagen, while working in Rutherford's laboratory at Manchester, England, extended the quantum theory still further to the atom itself and was able to derive the value of the Rydberg constant and the formula for the Balmer series of hydrogen from his theory.

The Einstein photoelectric equation (16-17) may be written

$$\nu = \frac{(\tfrac{1}{2})mv^2 + W}{h} \tag{16-18}$$

Now $(\tfrac{1}{2})mv^2$ is the kinetic energy of the photoelectron and W is the work or energy required to take an electron out of a metal, so that (16–18) can be written in the form ν = energy/h. Hence, Bohr argues that, if ν' in (16–16) is replaced by the frequency ν, then A and B are each of the form energy/h. We therefore write (16–16) in the form

$$\nu = \left(\frac{W_1}{h}\right) - \left(\frac{W_2}{h}\right) \tag{16–19}$$

where W_1 and W_2 are energies. Multiplying (16–19) throughout by h, we obtain

$$h\nu = W_1 - W_2 \tag{16–20}$$

Bohr's First Assumption.—According to Bohr a given spectrum line is produced by numerous atoms each emitting a quantum $h\nu$, where ν is the frequency of the given spectrum line. In the process of emitting the quantum $h\nu$, the energy of an atom changes from W_1 to W_2, and the difference of the energies is emitted as a quantum $h\nu$ according to (16–20). Hence Bohr's first assumption is expressed by (16–20).

Bohr uses the Rutherford model of the atom. The simplest case is for hydrogen where a single electron revolves about a nucleus carrying a charge $+e'$. However, we shall consider the case where a single electron revolves about a nucleus whose charge is $+Ze'$. The attractive force is then $Ze' \times e'/r^2 = Ze'^2/r^2$ where r is the distance between the nucleus and the electron. Since almost the whole mass of the atom is concentrated in the nucleus, the electron revolves about the nucleus and the nucleus can to a high degree of approximation be considered as remaining stationary. In the simplest form of the Bohr theory the electron revolves in a circular orbit around the nucleus which is at the center of the orbit as shown in Fig. 16.8. Since the attractive force must equal the centrifugal force in order to produce equilibrium, we have

$$\frac{mv^2}{r} = \frac{Ze'^2}{r^2} \tag{16–21}$$

where v is the velocity of the electron in its orbit, e' is the charge of the electron in esu and m is its mass. According to the classical theory, the radius r may have any value, provided that v has the corresponding value as given by (16–21). Also, according to this theory, an electron revolving in an orbit as in Fig. 16.8 should radiate energy.

Bohr's Second Assumption.—Bohr boldly assumes that the electron may revolve in one of a set of particular orbits but not in any orbit.

He further assumes that the electron, although moving with acceleration v^2/r in one of these particular orbits, does not radiate energy. Bohr thus denies two consequences of the classical theory. Further, Bohr makes his second assumption more specific and states that the electrons revolve only in those circular orbits where the angular momentum of the electron about the nucleus is equal to an integral multiple of $h/2\pi$ where h is Planck's constant.

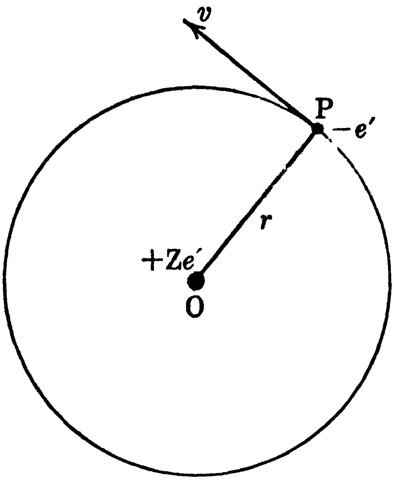


Fig. 16.8.

The momentum of a moving body is defined as the product of its mass and its velocity, while the angular momentum of a revolving body is defined as the product of its moment of inertia and its angular velocity. The moment of inertia of a mass m at P which is revolving about a center O as in Fig. 16.8 is mr^2. The angular velocity of the mass m at P about O is the number of radians through which OP turns in one second, so that the angular velocity is v/r. Hence the angular momentum is $mr^2 \times v/r = mvr$. Bohr's second assumption therefore requires that the electron revolve only in those orbits where

$$mvr = \frac{nh}{2\pi} \tag{16–22}$$

where n is a whole number.

Eliminating v from (16–21) and (16–22) we obtain

$$r = \frac{n^2h^2}{4\pi^2 Zme'^2} \tag{16–23}$$

Since $h^2/4\pi^2 Zme'^2$ is a constant for a given atom, it is seen that r is proportional to n^2, where n is a whole number. In the case of hydrogen, $Z = 1$, so that the radius of the smallest orbit is $1^2h^2/4\pi^2 me'^2$. Remembering that $h = 6.62 \times 10^{-27}$ erg sec, $m = 9.1 \times 10^{-28}$ gm, and $e' = 4.8 \times 10^{-10}$ esu, we therefore obtain for the radius of the first Bohr orbit in hydrogen

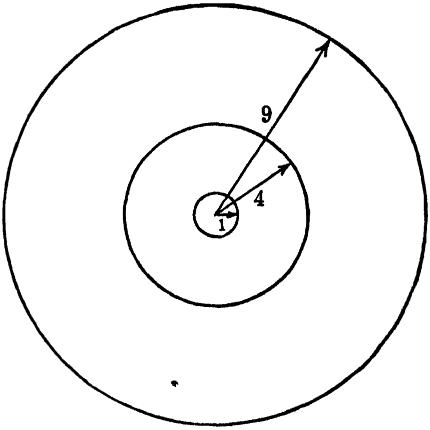


FIG. 16.9.—First three Bohr orbits in the hydrogen atom.

$$a_1 = 0.53 \times 10^{-8} \text{ cm} = 0.53 \text{ A} \tag{16–24}$$

The radii of the second, third and fourth Bohr orbits in hydrogen are $2^2 \times 0.53$, $3^2 \times 0.53$ and $4^2 \times 0.53$ A respectively. The first three Bohr orbits in hydrogen are drawn to scale in Fig. 16.9.

Now let us calculate the energy of an electron in any one of these orbits. The energy in a given orbit consists of both kinetic and potential energy. The velocity is found by substituting the value of r from (16–23) in (16–22), giving $v = 2\pi Ze'^2/nh$. Hence the kinetic energy is

$$(\tfrac{1}{2})mv^2 = \frac{2\pi^2 Z^2 me'^4}{n^2h^2} \tag{16–25}$$

The potential energy of a system of bodies is always referred to some standard configuration of the bodies which is said to have zero potential energy. In the case of an electron of charge $-e'$ and a nucleus of charge $+Ze'$ the potential energy is said to be zero when the two are at an infinite distance apart. Since, in allowing the electron to approach the nucleus till they are separated by a distance a, work is done by the attractive force, the potential energy of the system is less than zero by the amount of work done. Let the nucleus be at O and the electron at P in Fig. 16.10. Let $OP = r$ and $QP = dr$. If the electron moves from P to Q under the attractive force Ze'^2/r^2, the work done is

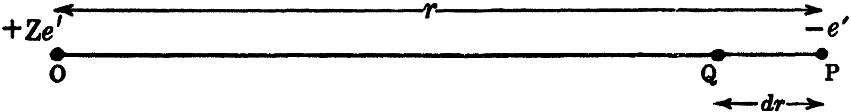


FIG. 16.10.

$$(Ze'/r^2) \times (-dr)$$

where the negative sign indicates that dr is in the direction of diminishing r. Let the diminution of potential energy as the electron goes from P to Q be $-dW$, so that

$$-dW = -\left(\frac{Ze'^2}{r^2}\right) dr \tag{16–26}$$

Integrating this from $r = \infty$ to $r = a$, we obtain

$$W = \int_\infty^a \left(\frac{Ze'^2}{r^2}\right) dr = Ze'^2\left[-\frac{1}{r}\right]_\infty^a = -\frac{Ze'^2}{a} \tag{16–27}$$

where a is a radius of a Bohr orbit. The potential energy of an electron in a Bohr orbit is thus $-Ze'^2/r$ where r is given by (16–23). The total energy of the electron in a Bohr orbit is

$$W_n = -\frac{Ze'^2}{r} + (½)mv^2 = -\frac{2\pi^2 Z^2 m e'^4}{n^2 h^2} \tag{16–28}$$

in virtue of (16–23) and (16–25).

The subscript n in W_n indicates that the energy is for the nth Bohr orbit. The total energy is negative because we have chosen to call the potential energy zero when the nucleus and electron are separated by an infinite distance.

In the Bohr theory the energy of an atom consisting of a nucleus and an electron can have only one of the set of discrete values given by (16–28). No intermediate values are possible. Occasionally however, the energy may suddenly change from a greater discrete value to a smaller discrete value. When this happens the energy of the atom suddenly changes from W_k to W_l where W_k is the energy of the electron in the kth orbit and W_l the energy of the electron in the lth orbit. For $W_k - W_l$ to be positive, k must be greater than l, since both W_k and W_l according to (16–28) are negative. The electron is said to fall from an outer orbit into an inner orbit. This sudden change in the energy of the atom is by Bohr's first assumption accompanied by the emission of a quantum $h\nu$, so that

$$h\nu = W_k - W_l = \frac{2\pi^2 Z^2 m e'^4}{h^2}\left\{\frac{1}{l^2} - \frac{1}{k^2}\right\} \tag{16–29}$$

If instead of the frequency ν we use the wave-number ν', (16–29) becomes

$$\nu' = \frac{2\pi^2 Z^2 m e'^4}{ch^3}\left\{\frac{1}{l^2} - \frac{1}{k^2}\right\} \tag{16–30}$$

Since l and k are whole numbers this formula is similar to the series formulas (16–6), (16–8), (16–9), and (16–10).

16.8. Agreement with Experiment.—For hydrogen, $Z = 1$, so that for this element

$$\frac{2\pi^2 1^2 m e'^4}{ch^3} = 109{,}700 \text{ cm}^{-1} \tag{16–31}$$

which is the value of the Rydberg constant. The Balmer series for hydrogen is obtained from (16–30) by making $l = 2$ and $k = 3, 4, 5,$ etc. for the various lines of the Balmer series. The Lyman series is obtained by making $l = 1$ and $k = 2, 3, 4,$ etc., successively. The Paschen series is obtained by making $l = 3$ and $k = 4, 5, 6,$ etc., successively.

The Balmer series is thus produced by electrons in different atoms falling from the 3rd, 4th, 5th, etc., Bohr orbits in different atoms to the 2nd orbit. It is to be noted that a given atom only emits one line while an electron falls from one orbit to another. One atom does not emit a whole spectrum at a given instant of time but only one line of that spectrum. The process of an electron falling from one orbit to another is known as an electron transition. The α line of the Balmer series is emitted by those atoms in which the electron transitions are occurring from the 3rd to the 2nd orbits; the β line is emitted by those atoms in which the transitions are from the 4th to the 2nd orbits; the γ line is emitted by those atoms in which the transitions are from the 5th to the 2nd orbits; and so on. The lines in the Balmer spectrum are not equally bright. The brightness of a line depends on the number of atoms in which the appropriate transition is occurring. For instance, if each of N_α atoms emits the quantum $h\nu_\alpha$ corresponding to the α line in one second, and if each of N_β atoms emits the quantum $h\nu_\beta$ in one second, the ratio of the energies in the α and β lines will be as $N_\alpha h\nu_\alpha$ is to $N_\beta h\nu_\beta$.

The Lyman series is produced by electron transitions from the 2nd, 3rd, 4th, etc., orbits in different atoms to the 1st orbit, and the Paschen series is produced by transitions from the 4th, 5th, 6th, etc., orbits in different atoms to the 3rd orbit. In Sec. 16.4 we have seen that a helium series spectrum is given by (16–11). In this formula the Rydberg constant is multiplied by 4, which is as it should be according to (16–30), since in that formula the Rydberg constant is multiplied by Z^2. The wave-number of the α line in the K x-ray spectrum of molybdenum is very nearly given by (16–30) when $Z = 42$, $l = 1$ and $k = 2$. We therefore see that the Bohr formula (16–30) holds over a wide range.

16.9. Energy Level Diagrams.—In Fig. 16.9 only three orbits are shown because the fourth and higher orbits are so large as to make it impracticable to represent them all on a page of this book. This is due to the fact that the radii of the orbits increase with the square of the quantum number n. If, however, the energies of the various orbits are calculated from (16–28), the energy values crowd together as the quantum number n is made large. Since the potential energy of a body with respect to the earth increases with the height of the level at which the body finds itself, it has become the practice to represent the different discrete energy states of an atom by lines drawn at different levels. These lines can always be conveniently drawn on a book page. The energies of the different orbits in the hydrogen atom are obtained by putting $Z = 1$ and $n = 1, 2, 3$, etc., in (16–28). We have seen that it is convenient to express the velocity and also the energy of an electron in electron-volts according to the relation

$$\text{Energy} = \frac{Ve'}{300} \tag{16–32}$$

where V is measured in electron-volts, e' in esu and the number 300 appears because 1 esu of potential equals 300 volts. The energy in electron-volts of the electron in the various discrete orbits in the hydrogen atom is therefore from (16–28) given by

$$V_n = -600\,\pi^2 \frac{me'^3}{n^2h^2} \tag{16–33}$$

Putting in the values of m, e', and h, we obtain

$$V_n = -\frac{13.6}{n^2} \tag{16–34}$$

as expressed in electron-volts. The negative sign implies that the energy is less than the energy when the hydrogen nucleus and the electron are separated by an infinite distance.

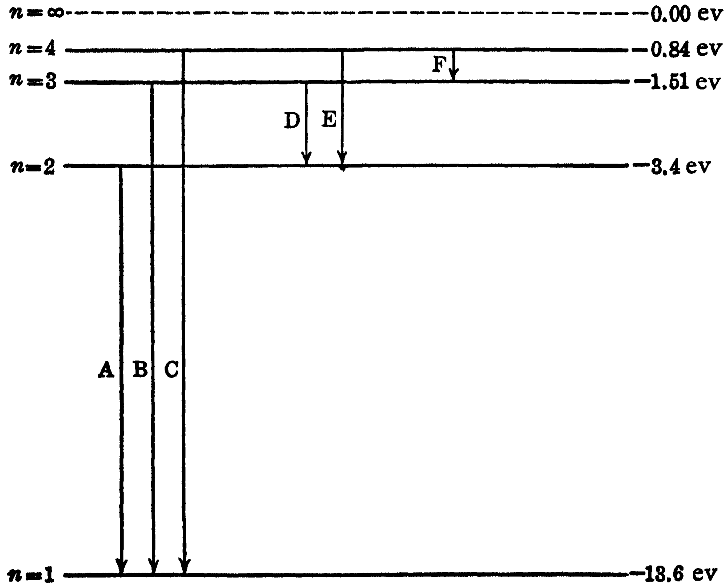


FIG. 16.11.—Energy level diagram for atomic hydrogen.

Let us draw a horizontal broken line to represent the level at which the energy of the atom is zero, as shown in Fig. 16.11. At a distance below the broken line proportional to 13.6 draw a line $n = 1$ as shown. For $n = 2$, $V_2 = -13.6/4 = -3.4$ ev. At a distance below the broken line proportional to -3.4 draw a line $n = 2$ as shown in Fig. 16.11.

Similarly the lines $n = 3$, $n = 4$ are drawn. Electron transitions are indicated by vertical arrows as shown by the arrows A, B, C, D, E, F, in Fig. 16.11. A diagram such as that shown in Fig. 16.11 is called an energy level diagram. Such diagrams have proved to be much more useful than diagrams such as Fig. 16.9.

The first three lines of the Lyman series are produced by the transitions represented by the arrows A, B, C, respectively. The energy which goes into the quantum when the transition A occurs corresponds therefore to $\{-3.40 - (-13.6)\}$ or 10.20 ev and this corresponds to a wave-length of

$$\lambda = \frac{12{,}345}{10.20} = 1210 \text{ A} \tag{16-35}$$

The first two lines of the Balmer series are produced by the transitions represented by the arrows D, E, respectively, while the first line of the Paschen series is produced by the transition represented by the arrow F. The wave-length of the second line of the Balmer series is $12{,}345/\{-0.84 - (-3.40)\} = 4820$ A.

The wave-numbers of the spectrum lines are proportional to the lengths of the corresponding arrows, because the wave-numbers are proportional to the corresponding energy changes which occur in the atom. Since the length of the arrow D equals the difference of the respective lengths of the arrows B and A in Fig. 16.11, the wave-number of the α line in the Balmer series should equal the difference of the respective wave-numbers of the β and α lines in the Lyman series. This is found to be the case experimentally. A similar relation is found to hold very nearly in x-rays, as shown in Table 16.1.

Since the wave-number of the $L\alpha$ x-ray spectrum line is so nearly equal to the difference of the wave-numbers of the $K\beta$ and $K\alpha$ lines respectively, and also since, according to Sec. 16.4 and (16-30), the wave-number of the molybdenum $K\alpha$ line can be so nearly calculated by the Bohr theory, it seems that the explanation of x-ray spectra must be along the lines of the Bohr theory. We shall return to the explanation of x-ray spectra in Chapter XVII.

16.10. Extension of Bohr's Theory.—We have derived the various discrete energy states of an atom consisting of a nucleus and an electron on the assumption that the electron moves in a circular orbit around the nucleus. It was soon found that elliptical orbits could be used and that except for small corrections due to relativity, the Bohr formula (16-30) could be derived for these orbits. The most recent development of the quantum theory is that the shape of the orbits is unknown and all that is necessary is that each electron in whatever orbit it may be moving must have an energy which is one of a set of discrete values such as given in (16–28).

TABLE 16.1

MOLYBDENUM X-RAY SPECTRUM

Line	Wave-number cm^{-1}
$K\beta$	1.585×10^8
$K\alpha$	1.409×10^8
$K\beta - K\alpha$	0.176×10^8
$L\alpha$	0.185×10^8

In the simple Bohr theory we have considered a single electron moving about a nucleus. However, according to the Rutherford theory of the atom, a neutral atom consists of a nucleus with a charge $+Ze$ surrounded by Z electrons, where Z is the atomic number. The formula (16–30) therefore only holds for an atom containing one electron. It therefore holds for the case of a neutral hydrogen atom. The formula holds only for those helium atoms which have lost one electron and only for those lithium (At. No. 3) atoms which have lost 2 electrons. Atoms which have lost electrons are said to be ionized, so that the Bohr formula holds for singly ionized helium and for doubly ionized lithium.

The extension of the Bohr theory to an atom consisting of a nucleus and more than one electron is extremely difficult. In astronomy there is what is known as the three-body problem. In this problem it is desired to find the orbits of each of three bodies which are moving under the action of the three mutual attracting gravitational forces. The exact solution of this problem is extremely difficult and has never been obtained, although approximate solutions for particular cases have been known for many years. The general problem for the case where there are more than three bodies remains unsolved. The neutral helium atom consists of two electrons moving under the attraction of the nucleus and under their own mutual repulsion. Here we have a three-body problem similar to that in astronomy but differing from it in that there is repulsion as well as attraction.

Fortunately, although the energy levels for atoms containing more than one electron cannot be calculated even approximately except with great difficulty, there are experimental methods for measuring the energy levels of atoms, as we shall see in Chapter XVIII.

16.11. Helium Spectrum.—A formula for a series spectrum of helium has been given in (16–11). When exact measurements are made it is found that the Rydberg constant for helium is slightly greater than that for hydrogen. In the Bohr theory as described in Sec. 16.7 we tacitly assumed that the mass of the nucleus was infinite. Actually, however, both the nucleus and the electron should be represented in Fig. 16.8 as revolving about their common center of gravity. The effect of this is to replace m in (16–28) and (16–30) by $mM/(M + m)$, where M is the mass of the nucleus. Thus the Rydberg constant for an electron revolving about a nucleus of mass M becomes

$$R_M = \frac{2\pi^2 e'^4}{ch^3}\cdot\frac{m}{(1 + m/M)} \qquad (16\text{–}36)$$

or

$$R_M \doteq R_\infty/(1 + m/M) \qquad (16\text{–}37)$$

Where R_∞ is the Rydberg constant for an atom in which the mass of the nucleus is infinite. Since the atomic weights of hydrogen and helium are very nearly 1 and 4 respectively, we have $R_{\text{He}} - R_{\text{H}} = 3R_\infty/(4 \times 1845) = 44.6$, which is the difference found experimentally.

CHAPTER XVI

PROBLEMS

1. Determine the wave-number and frequency corresponding to the following wave-lengths: 7000, 3000, 500, 20, 3, 0.71 A.

2. Determine the wave-numbers of the first 6 lines of the Lyman, the Balmer and the Paschen series spectra of hydrogen. Also determine the wave-lengths. Draw the respective spectra to scale with respect to wave-number and also with respect to wave-length. Also find the wave-number and wave-length of the limit of each spectrum and indicate these on the respective drawings.

3. Determine the wave-numbers of the β and α lines of the K x-ray spectrum and of the α line of the L x-ray spectrum of tungsten. Show that the difference of the wave-numbers of the $K\beta$ and $K\alpha$ lines is nearly but not quite equal to the wave-number of the $L\alpha$ line of tungsten as in the case of molybdenum.

4. The spectrum of ionized helium which is given by (16–11) of Chapter XVI is obtained from (16–30) by putting $Z = 2$, $l = 3$, and $k = 4, 5, 6 \cdots$. The energy of the electron in ionized helium is obtained by putting $Z = 2$ in (16–28).

Hence, obtain the formula for calculating the energy in electron-volts of the electron in the various discrete orbits in the ionized helium atom. Next calculate these energies for several values of n and then construct an energy level diagram drawn to scale for ionized helium.

5. What series spectra of higher wave-numbers than those given by (16–11) might be expected for ionized helium in view of the energy level diagram which has been constructed in Problem 4? Determine the wave-lengths of four consecutive lines of each of these expected series spectra from the energy level diagram of Problem 4.

CHAPTER XVII

THE QUANTUM THEORY

17.1. Doppler Effect.—In Sec. 3.10 we have given an explanation of the Doppler effect in terms of a wave theory. We now proceed to give an explanation of the effect for light in terms of a quantum or photon theory.

According to the relativity theory, energy has mass, so that when an atom is in a higher energy state it has a greater mass than when in a lower state. According to the Bohr theory, when an atom changes from an energy state W_1 to W_2, a quantum $h\nu$ is emitted according to (16–20). Combining these two ideas, the mass of the atom changes from M'_1 to M'_2 when a quantum is emitted, so that we have

$$h\nu = M'_1c^2 - M'_2c^2 \tag{17-1}$$

since on the relativity theory the energy of a mass M is Mc^2. However, when a quantum is emitted it is shot out in a particular direction like a bullet, and the rest of the atom recoils like a gun. Let us suppose that the atom is at rest before the emission of the quantum $h\nu$, so that $M'_1 = M_1$, the rest-mass of the atom before the emission of the quantum. However, M'_2 is not equal to M_2 the rest-mass after the emission of the quantum because, after the emission, the atom is recoiling with a speed βc, so that the relation between M'_2 and M_2 is

$$M'_2 = \frac{M_2}{\sqrt{1 - \beta^2}} \tag{17-2}$$

Hence (17–1) becomes

$$h\nu = M_1c^2 - \frac{M_2c^2}{\sqrt{1 - \beta^2}} \tag{17-3}$$

In Sec. 14.26 the momentum of a quantum was shown to be $h\nu/c$. Now the principle of the conservation of momentum requires that the atom after the emission of a quantum $h\nu$ must kick back with a momentum $h\nu/c$. Hence, using the relativity formula (13–31) for the momentum, we have

$$\frac{M_2\beta c}{\sqrt{1-\beta^2}} = \frac{h\nu}{c} \tag{17-4}$$

Eliminating β from (17-3) and (17-4), we obtain

$$h\nu = \frac{(M_1{}^2 - M_2{}^2)c^2}{2M_1} \tag{17-5}$$

In the previous paragraph we have supposed the atom to have a rest-mass M_1 and to be at rest before the quantum was emitted. We now suppose that the atom has a velocity $\beta_1 c$ before emitting the quantum and that this velocity is in the direction of the emission of the quantum. Also we suppose that as before the rest-mass of the atom changes from M_1 to M_2 on the emission of the quantum. However, the atom is initially moving with velocity $\beta_1 c$, so that its mass is $M_1/\sqrt{1-\beta_1{}^2}$ before the emission of the quantum and its energy is $M_1c^2/\sqrt{1-\beta_1{}^2}$. Hence, on the emission of the quantum which now has an energy $h\nu'$, we have

$$h\nu' = \frac{M_1c^2}{\sqrt{1-\beta_1{}^2}} - \frac{M_2c^2}{\sqrt{1-\beta_2{}^2}} \tag{17-6}$$

where $\beta_2 c$ is the velocity of the atom after emission of the quantum. Also the principle of the conservation of momentum requires that

$$\frac{M_1\beta_1 c}{\sqrt{1-\beta_1{}^2}} - \frac{M_2\beta_2 c}{\sqrt{1-\beta_2{}^2}} = \frac{h\nu'}{c} \tag{17-7}$$

Eliminating β_2 from (17-6) and (17-7), we have

$$h\nu' = \frac{(M_1{}^2 - M_2{}^2)c^2}{2M_1} \cdot \frac{\sqrt{1-\beta_1{}^2}}{1-\beta_1} \tag{17-8}$$

Dividing (17-8) by (17-5), we obtain

$$\frac{\nu'}{\nu} = \frac{\sqrt{1-\beta_1{}^2}}{1-\beta_1} \tag{17-9}$$

For initial velocities of the atom which do not approach that of light, β_1 is small and $\beta_1{}^2$ may be neglected, so that (17-9) becomes

$$\nu' = \frac{\nu}{(1-\beta_1)} = \frac{\nu c}{(c-v)} \tag{17-10}$$

where v is written for $\beta_1 c$ the initial velocity of the atom toward the observer. It is immediately seen that (17–10) corresponds to (3–53).

This theory of the Doppler effect was first stated by Schroedinger of Germany in 1922. This theory together with the theory of the Compton effect (published in 1923) shows how useful such ideas as a "bouncing quantum" and a "quantum shot out from an atom like a bullet from a gun" had become in 1923. The minds of physicists seemed definitely to be turning towards a corpuscular theory of light. However, as we have remarked in Secs. 14.13 and 14.24 no corpuscular theory can explain interference phenomena. We must therefore recognize the dual nature of light—sometimes it has properties associated with corpuscles and sometimes properties associated with waves. Since a corpuscle may be called a particle, Eddington has recently described light as consisting of "wavicles" (a combination of the words "wave" and "particle").

17.2. De Broglie Waves.—In 1924 Louis de Broglie of France made the bold step of extending the idea of the dual nature of light to matter itself. De Broglie assumes that any discrete amount of energy W has a frequency ν associated with it according to the quantum relation

$$W = h\nu \tag{17–11}$$

In the relativity theory we have seen that a mass has energy or an energy has mass. Hence, it would seem that a mass m which has an energy mc^2 would have an associated frequency ν given by

$$mc^2 = h\nu \tag{17–12}$$

In the case of a body moving with velocity βc the energy may be expressed in terms of the rest-mass m_o, so that

$$\frac{m_o c^2}{\sqrt{1 - \beta^2}} = h\nu \tag{17–13}$$

According to (3–47) the group velocity v of a set of waves traveling with wave or phase velocity u is given by

$$v = u - \lambda \frac{du}{d\lambda} \tag{17–14}$$

This gives v in terms of u and the wave-length λ. For our present purpose it is more convenient to express v in terms of u and the frequency ν. By the methods of the calculus it can be shown that, since $\nu\lambda = u$, the group velocity is given by

$$\frac{1}{v} = \frac{1}{u} - \frac{\nu}{u^2} \cdot \frac{du}{d\nu} \tag{17–15}$$

Let us now suppose that there is a wave or phase velocity u given by

$$u = \frac{c}{\beta} \tag{17–16}$$

where c is the velocity of light and β is less than unity. Differentiating (17–16), we obtain

$$\frac{du}{d\beta} = -\frac{c}{\beta^2} \tag{17–17}$$

Also differentiating (17–13), we obtain

$$\frac{d\nu}{d\beta} = \frac{m_o \beta c^2}{h(1 - \beta^2)^{3/2}} \tag{17–18}$$

Now

$$\frac{du}{d\nu} = \frac{du}{d\beta} \cdot \frac{d\beta}{d\nu} = -\frac{hc(1 - \beta^2)^{3/2}}{m_o \beta^3 c^2} \tag{17–19}$$

Substituting (17–19) in (17–15) and using (17–16) and (17–14), we obtain

$$v = \beta c \tag{17–20}$$

Hence the velocity of the particle of rest-mass m_o is the same as that of the group of the waves associated with it. Multiplying the numerator and denominator of the right side of (17–16) by c and transposing, we obtain

$$uv = c^2 \tag{17–21}$$

This equation means that the product of the wave velocity u and the group velocity v of the waves associated with a particle which is moving with velocity $v = \beta c$ is always equal to the square of the velocity of light. These associated waves are known as *de Broglie waves.*

Now for the associated waves the wave-length $\lambda = u/\nu$. Using the values of u and ν given by (17–16) and (17–13) respectively, we obtain

$$\lambda = \frac{h\sqrt{1 - \beta^2}}{m_o \beta c} \tag{17–22}$$

But $m_o/\sqrt{1-\beta^2}$ is the mass of a particle of rest-mass m_o which is traveling with velocity βc, so that we may replace $m_o/\sqrt{1-\beta^2}$ with m in (17–22). We then obtain de Broglie's famous equation

$$\lambda = \frac{h}{mv} \tag{17–23}$$

The meaning of this equation is that a particle of mass m which is moving with velocity v has associated with it de Broglie waves whose wave-length is given by (17–23).

For velocities small compared with that of light $mv = \sqrt{2m \text{ (kinetic energy)}}$. Then, since the kinetic energy is the difference between the total energy W and the potential energy V, we have $mv = \sqrt{2m(W - V)}$ and (17–23) becomes

$$\lambda = h/\sqrt{2m(W - V)} \tag{17–24}$$

17.3. Davisson-Germer Experiment.—The question now arises as to whether there is any experimental evidence for de Broglie's idea that a particle of matter should under the proper circumstances show properties which we ordinarily associate with wave motion. An electron is such a particle. Is it therefore to be expected that an electron or rather a stream of electrons may be diffracted in a way similar to the diffraction of a beam of light? According to de Broglie a stream of electrons moving with velocity v will under the proper circumstances behave in a manner somewhat similar to light of a wave-length λ given by (17–23). For instance, electrons which have fallen through a potential of 100 volts have a velocity of 5.9×10^8 cm/sec, so that, since $h = 6.6 \times 10^{-27}$ erg sec and $m = 9.1 \times 10^{-28}$ gm, the associated wave-length is 1.2×10^{-8} cm or 1.2 angstroms. This wave-length is in the region of x-ray wave-lengths, and it may be that electrons can be diffracted by crystals in a similar manner to x-rays.

In 1927 Davisson and Germer, of the Bell Telephone Laboratories, New York, shot electrons at a crystal of nickel and observed a sort of Laue spot effect. They found that the electrons acted as if they had a wave-length. Intense reflections took place at angles given by the Bragg formula

$$n\lambda = 2d \sin \theta$$

where d is the grating space of the crystal of nickel and λ is the de Broglie wave-length given by (17–23).

More recently G. P. Thomson has obtained diffraction of electrons from thin films of gold and other substances. He obtains photographs which are similar to the powdered crystal photographs obtained when x-rays are used. Such a photograph is shown in Fig. 17.1. The similarity of this photograph to that of Fig. 15.5 is obvious.

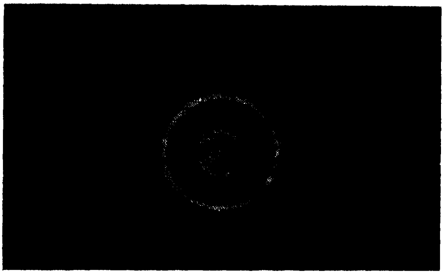


FIG. 17.1—Diffraction of Electrons by Gold Leaf (Thomson).

Thus, the present situation is that both light and streams of electrons act sometimes as if they are corpuscles and at other times as if they are systems of waves. Physicists tend now to believe in the dual nature of light waves and material particles.

17.4. Quantum Mechanics.—From 1912 until about 1920 the Bohr theory had remarkable success. However, beginning about 1920, it began to be realized that the quantum theory was not a complete success. In 1925–26 a fresh start was made from two different directions by Heisenberg and Schroedinger, both of Germany.

Heisenberg began with the idea that we have been putting too much into our atomic models. He states that our equations should contain only measurable quantities. We measure wave-lengths and intensities of spectrum lines, but we never measure radii of electron orbits in the atoms. Perhaps it is not even correct to speak of electrons moving in orbits because we can never measure these orbits—they are only figments of the imagination. In order to put Heisenberg's ideas into mathematical form, it is necessary to use a branch of mathematics known as matrix calculus, which is beyond the scope of this book.

Schroedinger, on the other hand, noted the dual nature of light. Sometimes light acts as if it were made up of corpuscles and at other times as if it were made up of waves. Schroedinger made use of de Broglie's ideas. He put these ideas into mathematical form and developed what is now known as "Wave Mechanics." In wave mechanics a particle of mass m moving with velocity v has associated with it a group of waves whose velocity is u, the relation between u, v and c being given by (17–21). In the case of light, u and v are each equal to c, but in the case of an electron $u > c$ while $v < c$. Schroedinger also showed that his ideas are mathematically equivalent to those of Heisenberg. Both the Heisenberg and the Schroedinger theories are included in the general term The New Quantum Theory.

From (17–24) the de Broglie wave-length associated with an electron having a total energy W and a potential energy V is

$$\lambda = h/\sqrt{2m(W - V)} \tag{17–25}$$

According to the views of de Broglie and Schroedinger, the de Broglie waves which are associated with the electrons of an atom produce sets of standing or stationary waves. In Sec. 16.9 we have spoken of energy levels. Bohr originally gave the name stationary states to these levels. According to de Broglie and Schroedinger's wave mechanics, a stationary state corresponds to something like a vibrational mode. In this view then the length of the path of an electron revolving about the nucleus of an atom must contain a whole number of wave-lengths of the associated de Broglie wave. For a circular path, therefore,

$$2\pi r/\lambda = n \tag{17–26}$$

Now we have seen in Sec. 16.7 that the potential energy of the electron in the hydrogen atom is $-e^2/r$ or $-2\pi e^2/n\lambda$ in virtue of (17–26). Substituting this value for V in (17–25), we obtain, after squaring and rearranging,

$$2nmW\lambda^2 + 4\pi mc^2\lambda - h^2n = 0 \tag{17–27}$$

This is a quadratic equation in λ. Since according to the theory only one de Broglie wave-length may correspond to an electron in an atom, the solution of (17–27) must give only one value of λ—in other words, (17–27) must be a perfect square. Whence

$$16\pi^2m^2e^4 + 8h^2n^2mW = 0$$

or

$$W_n = -2\pi^2me^4/h^2n^2 \tag{17–28}$$

which agrees with (16–28) obtained from the Bohr theory. Instead, however, of the picture of an electron revolving about a nucleus being as shown in Fig. 17.2a, the picture according to wave mechanics is more like that shown in Fig. 17.2b.

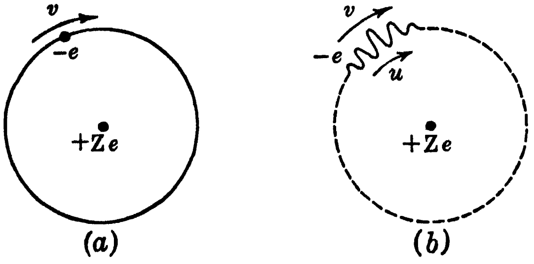


FIG. 17.2.—(a) Bohr's representation of an electron in an atom; (b) Representation of an electron in an atom according to wave-mechanics.

17.5. Fermi-Dirac Statistics.—In Chapter X we applied the Maxwell distribution function for the component velocities of the molecules of a gas to the electrons in a metal. It was pointed out, however, that the Maxwell distribution function, or Maxwell statistics, leads to a specific heat which is inconsistent with experiment. Recently Fermi, then of Rome, Italy, and Dirac, of England, have considered the case of an electron gas and have found that not only does the distribution function for an electron gas depart from the Maxwell distribution function, but that the distribution function for an ordinary gas such as helium also departs from the Maxwell function at temperatures close to the absolute zero. The Fermi-Dirac distribution function merges into the Maxwell distribution function when the temperature is high enough. For ordinary gases at ordinary temperatures, the Fermi-Dirac function is practically identical with the Maxwell function, but for the conduction electrons within a metal at ordinary temperatures there is still a distinct difference between the two functions and it is only at quite high temperatures that the Fermi-Dirac function for electrons approaches the Maxwell function.

We shall consider n free electrons in 1 cm^3. According to de Broglie each electron may be replaced by waves of wave-length

$$\lambda = h/mV \tag{17-29}$$

where V is the velocity of the electron in cm/sec. If these waves are enclosed in a space, standing or stationary waves are set up—in other words, vibrational modes are produced. In Chapter III we have seen that the number of vibrational modes in the wave-length range λ to

$\lambda + d\lambda$ is

$$8\pi d\lambda/\lambda^4 \tag{17-30}$$

per unit volume for transverse waves. It is an assumption of wave mechanics that a vibrational mode corresponds to each electron in the enclosed space. From (17–29)

$$d\lambda = -hdV/mV^2 \tag{17-31}$$

so that, in virtue of (17–29) and (17–31), (17–30) becomes

$$8\pi m^3V^2dV/h^3 \tag{17-32}$$

This is the number of electrons with speeds between V and $V + dV$. Just as in the Debye theory of specific heats we integrated over the total frequency range $\nu = 0$ to $\nu = \nu_m$ so here we integrate from $V = 0$ to $V = V_m$, where V_m is a maximum velocity. Since each velocity corresponds to a vibrational mode of a certain wave-length and each vibrational mode corresponds to an electron, the integral of (17–32) over the velocity range $V = 0$ to $V = V_m$ must equal the total number n of the electrons. Hence

$$\int_0^{V_m} 8\pi m^3V^2dV/h^3 = n \tag{17-33}$$

whence

$$V_m = (h/m) \cdot \sqrt[3]{3n/8\pi} \tag{17-34}$$

According to Fermi and Dirac, (17–32) gives the number of electrons with speeds between V and $V + dV$ up to a maximum speed V_m at the temperature of absolute zero. At this temperature the Fermi-Dirac distribution function is $F(V)$, where

$$F(V) = \text{constant} \times V^2 \tag{17-35}$$

for $V \leqq V_m$ and

$$F(V) = 0 \tag{17-36}$$

for $V \geqq V_m$.

Fermi and Dirac have also applied their theory to the distribution function for component velocities. In this case the number of electrons whose component velocities are between u and $u + du$ is $f(u)du$, where

$$f(u) = \text{constant } (V_m{}^2 - u^2) \tag{17-37}$$

for $u \leqq V_m$ and

$$f(u) = 0 \tag{17-38}$$

for $u \geqq V_m$. At all ordinary temperatures for the electrons in a metal the Fermi-Dirac distribution functions $F(V)$ and $f(u)$ are very nearly those at absolute zero as given by (17–35), (17–36), (17–37), and (17–38) respectively. The value of the critical velocity V_m is that correponding to an energy of 5 or 6 ev. Comparison of the Fermi-Dirac distribution functions with the corresponding Maxwell distribution functions is shown in Figs. 17.3 and 17.4. In each figure, Curve II is the Fermi-Dirac curve for electrons at absolute zero, it being assumed that V_m is 13.3×10^7 cm/sec (corresponding to 5 ev), while Curve I is the Maxwell distribution curve for electrons at 1000° Abs. Curves I and II in each figure apply to the same number of electrons so that the area under Curve I equals the area under Curve II. If the temperature of the electrons with the Maxwell distribution is reduced below 1000° Abs, the peak of Curve I in Fig. 17.3 moves to smaller values of V. As absolute zero is approached the peak approaches the axis of ordinates

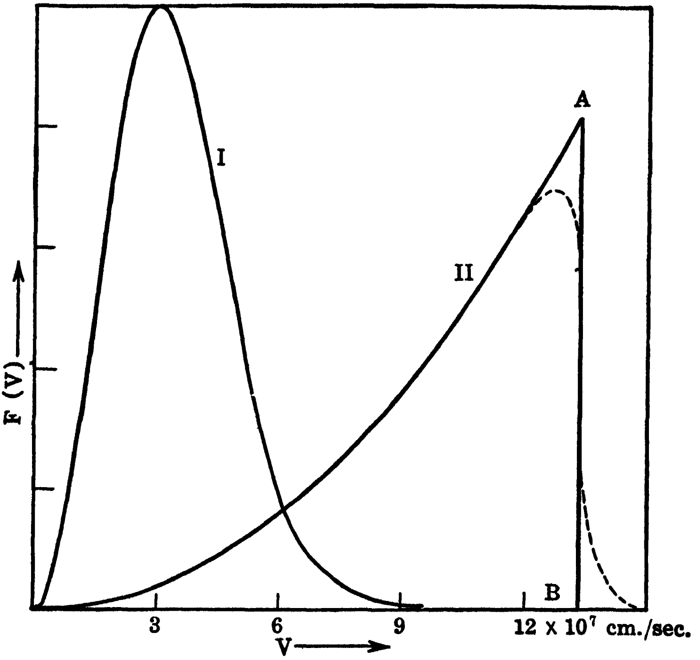


FIG. 17.3.—Total velocity distribution curves: I. Electrons at 1000° Abs (Maxwell; II. Electrons at 0° Abs (Fermi-Dirac); Broken curve—Electrons at a high temperature (Fermi-Dirac).

so that the Maxwell distribution curve differs vastly from the Fermi-Dirac curve. Also, as absolute zero is approached the peak of Curve I in Fig. 17.4 becomes narrower and higher, so that again the Maxwell distribution curve differs vastly from the Fermi-Dirac curve. At absolute zero, according to the Maxwell theory, the electrons all have zero velocity, while, according to the Fermi-Dirac theory, the electrons have velocities between zero and V_m. At 1000° Abs the Fermi-Dirac curves are very little changed from the curves at 0° Abs.

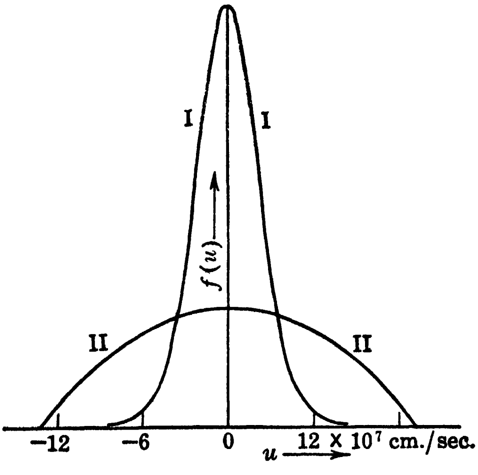


FIG. 17.4.—Component velocity distribution curves: I. Electrons at 100° Abs (Maxwell); II. Electrons at 0° Abs (Fermi-Dirac).

At high temperatures the Fermi-Dirac curves in Figs. 17.3 and 17.4 tend to approach the corresponding Maxwell curves as shown by the broken curves in these figures.

It is possible by means of the Fermi-Dirac theory to correlate the specific heat of a metal with its electrical resistance and its thermionic emission.

17.6. Quantum Numbers.—In the previous section we have stated that according to wave mechanics to each electron in a closed space there corresponds a vibrational mode of the de Broglie waves associated with the electron. We may look upon an atom as a closed space containing electrons since it is only under unusual circumstances that an electron can escape from an atom. We have seen in Secs. 14.12 and 14.18 that the number of electrons in an atom is Z, the atomic number.

With each of these electrons is associated a de Broglie vibrational mode of wave-length λ. In Chapter III when we were discussing the vibrational modes in a cube of side L we spoke of the solid cube as vibrating in the n_1th mode along the x-axis, in the n_2th mode along the y-axis, and in the n_3th mode along the z-axis. We abbreviated this by saying that the cube was vibrating in the (n_1, n_2, n_3)th mode. In the case of a vibrating sphere the problem is more difficult, the nodal surfaces instead of being planes, as they are for the cube, are more complicated surfaces. However, a sphere, like a cube, is a three-dimensional body and we must still use three numbers to describe the vibrational modes of a sphere. We may therefore speak of the sphere as vibrating in its (n_1, n_2, n_3)th mode. However, in the case of transverse waves we need to say whether the vibrations are, say, vertical or horizontal and for this we need another symbol n_4. This quantity n_4 is a symbol rather than a number. It can have only *two* meanings, n_4 = "vertical" or n_4 = "horizontal." Hence each vibrational mode produced by transverse waves in a three-dimensional body requires four "numbers" to describe it completely. An atom is more like a sphere than a cube and the problem of finding the modes which are set up by the de Broglie waves is difficult and beyond the scope of this book. However, since the de Broglie waves are transverse, four numbers are necessary. In the wave-mechanical theory of the atom, therefore, we speak of the (n, l, m_l, m_s)th vibrational mode. These quantities n, l, m_l, m_s are now known as quantum numbers. Since, according to Pauli, formerly of Germany and now of Princeton, one vibrational mode corresponds to one electron in the atom, we have Pauli's exclusion principle, which states that no two electrons in the same atom can have the same set of quantum numbers. The quantum numbers n, l, m_l, m_s are given names but we shall content ourselves with the names of two of them. The quantum number n is called the principal quantum number and m_s is known as the "spin" quantum number. The spin quantum number corresponds to the symbol which tells us in the case of transverse elastic waves in a solid whether the vibrations are vertical or horizontal. In quantum theory m_s can have the values $\pm\frac{1}{2}$.

In Chapter XVI and in Sec. 17.4 of this chapter, we have obtained the formula for the energy of an electron in a hydrogen atom. The quantity n in the formula is the principal quantum number. When the spectrum lines of hydrogen are examined with a spectroscope of high resolving power it is found that each line usually consists of two or more lines very close together. Also under the action of a magnetic or an electric field upon hydrogen atoms which are emitting a spectrum, the single lines are found to split into two or more lines. The spectrum lines are then said to possess a fine structure. Since we still think of the emission of a spectrum line as being due to an electron falling from one energy level to another, the energy levels themselves must have a fine structure. The exact formula for the energy of an energy level of an electron is a function of l, m_l, and m_s as well as n. However, the effect of l, m_l and m_s on the energy is usually small compared with that of n, the principal quantum number.

In the theory of wave mechanics it is shown that the quantum numbers are limited to values shown in Table 17.1.

TABLE 17.1

$$\begin{aligned} n &= 1, 2, 3, \cdots \\ l &= 0, 1, 2, \cdots (n-1) \\ m_l &= -l, -l+1, \cdots, 0, \cdots, l-1, l \\ m_s &= \pm\tfrac{1}{2} \end{aligned}$$

We have spoken of m_s as the spin quantum number. We can either think of an electron as a particle or as a vibrational mode. If we think of it as a de Broglie vibrational mode the quantum number m_s tells us which way the waves are polarized but if we prefer the particle view of an electron it is supposed that the electric charge on the electron is spinning about an axis. A spinning charged sphere produces a magnetic field and if placed in an external magnetic field the sphere will tend to line up the axis of spin in the direction of the external magnetic field. In the particle view of the electron, m_s tells us which way the axis of spin is lined up.

17.7. Rutherford-Bohr Atom.—We have introduced the reader to the Periodic Table of Chemical Elements in Sec. 14.18. The symbols of the chemical elements were first arranged by Mendeleef of Russia in a tabular form in 1870. The modern form of Mendeleef's periodic law is shown in Table 17.2. Until the end of 1940 it was supposed that uranium (At. No. 92) completed the periodic table and that there were no elements of higher atomic number than 92. In 1946 we know of elements 93, 94, 95, and 96.

We have already mentioned Rutherford's model of the atom in Sec. 16.6. In this model, Rutherford supposed that the atom of an

TABLE 17.2

PERIODIC TABLE

1	H 1.008, 1							He 4, 2
2	Li 7, 3	Be 9, 4	B 11, 5	C 12, 6	N 14, 7	O 16, 8	F 19, 9	Ne 20, 10
3	Na 23, 11	Mg 24, 12	Al 27, 13	Si 28, 14	P 31, 15	S 32, 16	Cl 35.5, 17	A 39.9, 18
4	K 39, 19	Ca 40, 20	Period	of 18 ele	ments		Br 80, 35	Kr 83, 36
5	Rb 85.4, 37	Sr 87 6, 38	Period	of 18 ele	ments		I 127, 53	Xe 130 2, 54
6	Cs 132 8, 55	Ba 137 4, 56	Period	of 32 ele	ments		— —, 85	Rn 222, 86
7	— —, 87	Ra 226, 88	Ac 229, 89	Th 232, 90	Pa 234, 91	U 238, 92	—, 93	—, 94

element of atomic number Z consists of a nucleus carrying a charge $+Ze$ around which revolve Z electrons each carrying a charge $-e$, where $e = 1.6 \times 10^{-20}$ emu. In the Bohr theory of hydrogen-like atoms, which are atoms consisting of a nucleus with charge $+Ze$ and one electron, we have seen that the electron cannot revolve in any orbit but only in one of a particular set of discrete orbits. This idea is now applied to the Rutherford atom. The Z electrons which revolve about the nucleus are separated out into groups and these groups are arranged to fit the periods of the Periodic Table.

According to the Rutherford-Bohr theory of the atom, the hydrogen atom consists of the hydrogen nucleus and one electron, and this electron in the normal hydrogen atom is revolving in the smallest Bohr orbit, the orbit in which the quantum number n equals unity. According to (16–24) the radius of this orbit is 0.53 A. Such hydrogen atoms are in their lowest energy state and so cannot radiate spectrum lines. In order for a hydrogen atom to emit a quantum the electron must be in an orbit which corresponds to a higher energy level than that corresponding to the normal orbit. The Rutherford-Bohr theory describes the atoms as they are in their normal states and not as they are when they are able to emit quanta. Substances do not of their own accord emit spectra.

The normal hydrogen atom is represented in Fig. 17.5. The electron revolves in the circle as shown. Since this circle can be in any plane, it is better to describe it as a shell. This is the shell corresponding to the Bohr orbit of lowest energy, and it is called the *K* shell. The normal helium (He) atom according to the Rutherford-Bohr theory is also represented in Fig. 17.5. It consists of a nucleus carrying a charge $+2e$ with two electrons each on the same shell. They are represented as being at the opposite ends of a diameter. This has no significance as the two electrons may be anywhere in the shell relative to each other. Both electrons are in the *K* shell. The next element is lithium (Li).

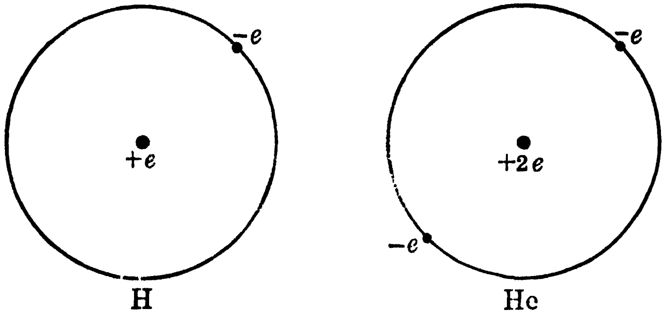


Fig. 17.5.—Rutherford-Bohr models of the hydrogen and helium atoms.

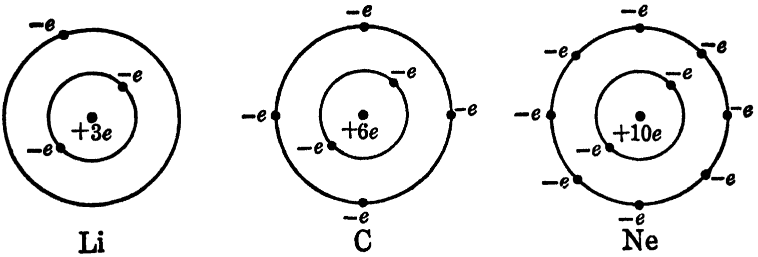
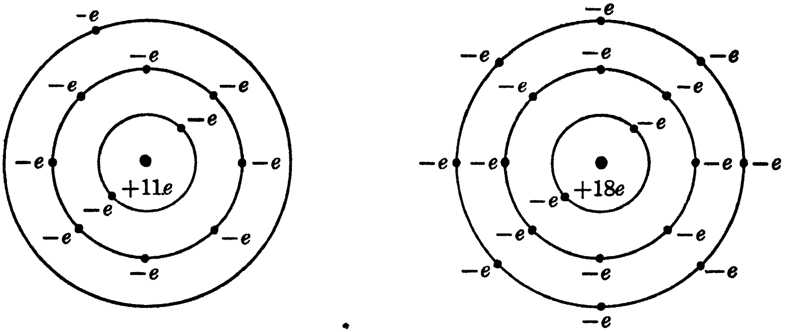


Fig. 17.6.—Rutherford-Bohr models of the lithium, carbon and neon atoms.

The lithium atom consists of a nucleus carrying a charge $+3e$ and two electrons in the *K* shell with the third electron in a new shell which is called the *L* shell. The lithium atom is represented in Fig. 17.6. As we go through the elements of the second period of the Periodic Table, electrons are added to the *L* shell until neon (Ne) is reached, when there are 8 electrons in the *L* shell, as shown in Fig. 17.6. The atom of carbon

(C) is also represented in Fig. 17.6. When we proceed from neon to sodium (Na)—that is, when we proceed from the second to the third period of the Periodic Table—the 11th electron takes its place in a new shell known as the *M* shell. The sodium atom is represented in Fig. 17.7. As we go through the elements of the third period, electrons are added to the *M* shell until argon is reached. The argon atom contains 2 electrons in the *K* shell, 8 electrons in the *L* shell and 8 electrons in the *M* shell, as represented in Fig. 17.7. Now as we proceed from argon to potassium, which is at the beginning of the fourth period, the 19th electron takes its place in a new shell known as the *N* shell. From here on the procedure becomes more complicated. For instance, an electron is sometimes added to an inner shell instead of always to the outermost shell as one proceeds along the period of the Periodic Table.



Na A

FIG. 17.7.—Rutherford-Bohr models of the sodium and argon atoms.

The *K* shell is complete with two electrons, the *L* shell is complete with 8 electrons, and the *M* shell with 18 electrons. In 1946, on account of the discovery of transuranic elements (At. No. >92), it is best to leave unstated the number of electrons which complete the shell outside the *M* shell.

The shells are determined by Bohr's quantum conditions. For the *K* shell the quantum number n equals one, for the *L* shell $n = 2$, for the *M* shell $n = 3$, and so on. The radii of the shells of the atoms represented in Figs, 17.5, 17.6, and 17.7 are not drawn to scale, and the electrons of a given shell are distributed at random on the shell and are not equally spaced as shown in the figures. According to the most recent view in physics, the electrons cannot be properly spoken of as being arranged in shells. Although a *K* electron can be at any distance from the nucleus, yet the two *K* electrons are most probably at a certain distance from the nucleus and this most probable distance corresponds to the radius of the *K* shell in the original form of the Rutherford-Bohr theory. Similarly the *L* electrons are most probably at a certain distance from the nucleus and this most probable distance corresponds to the radius of the *L* shell in the original form of the theory. Similarly the most probable distances of the *M* and *N* electrons from the nucleus correspond to the radii of the *M* and *N* shells in the original form of the theory. Accordingly, with these limitations in mind, we still speak of the *K, L, M*, etc., shells.

Instead of drawing diagrams such as those shown in Figs. 17.5, 17.6, and 17.7, we might represent the electrons of the various shells as being at different energy levels. In such a diagram, the two *K* electrons would have the lowest energy and be on the lowest energy level, the eight *L* electrons would be on the next to the lowest energy level and so on.

17.8. Quantum Theory of the Periodic Law.—In Sec. 17.6 we have stated that an electron in an atom may be viewed as a particle or as a de Broglie wave. From the wave aspect the electron corresponds to a vibrational mode. Since a vibrational mode is described by four quantum numbers n, l, m_l, m_s, Pauli has stated that no two electrons in the same atom can have the same set of quantum numbers. We now apply the rules in Table 17.1 and obtain Table 17.3.

TABLE 17.3

n	l	m_l	m_s
1	0	0	$-\frac{1}{2}$
1	0	0	$+\frac{1}{2}$
2	0	0	$-\frac{1}{2}$
2	0	0	$+\frac{1}{2}$
2	1	-1	$-\frac{1}{2}$
2	1	-1	$+\frac{1}{2}$
2	1	0	$-\frac{1}{2}$
2	1	0	$+\frac{1}{2}$
2	1	1	$-\frac{1}{2}$
2	1	1	$+\frac{1}{2}$

The *K, L, M*, etc. shells correspond to $n = 1, 2, 3$, etc. We thus see that for $n = 1$ there are only two possibilities and the *K* shell is therefore complete when it contains two electrons. For $n = 2$ there are eight possibilities and the shell is complete when it contains eight electrons. It can be shown that the nth shell is complete when it contains $2n^2$ electrons. If this rule is valid for the outer shells, we may say that the K, L, M, N shells are complete when they contain 2, 8, 18, 32 electrons respectively.

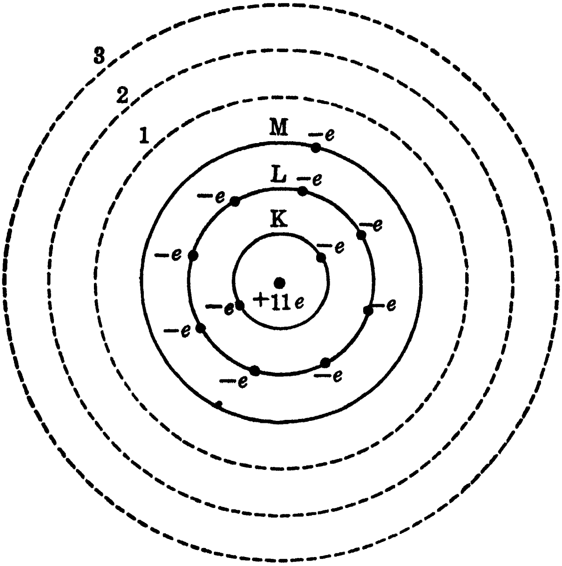


FIG. 17.8.—Bohr shells (or orbits) in the sodium atom.

17.9. Optical Spectra.—The Rutherford-Bohr theory applies to the normal atom as we have stated in Sec. 17.7. The optical spectrum of a substance is emitted only when some means is taken to excite the spectrum, such as placing the substance in a flame or in an electric discharge tube. When a gas is placed in a discharge tube, electrons hit the atoms of the gas. Owing to the impact of an electron from outside the atom, an electron in the atom may be knocked out of its normal shell or orbit into one of the outer shells. The electron after a while drops back into one of the inner shells and in doing so emits a quantum as explained in Chapter XVI. There are therefore optical shells or levels outside the outermost shell shown in Figs. 17.5, 17.6, 17.7. Thus, the complete system of shells of the sodium atom is represented in Fig. 17.8. The valence electron may be in the M shell as shown or in any one of the shells (or levels) which are represented by broken circles. The shells are not drawn to scale. If the valence electron is knocked out to shell 3, the electron will later return perhaps to shell 2, to shell 1 or back to its original shell. If the energy in shell 3 is W_3 and that in shell 2 is W_2 and so on, the quanta emitted due to the varoius electron transitions are

$$\left.\begin{aligned} h\nu &= W_3 - W_2 \\ h\nu &= W_3 - W_1 \\ h\nu &= W_3 - W_0 \end{aligned}\right\} \qquad (17\text{–}39)$$

where W_0 is the energy of the electron in its normal shell, the M shell, as shown in Fig. 17.8. Further, the electron, after having fallen from shell 3 to shell 2, may then fall from shell 2 to shell 1, emitting a quantum given by

$$h\nu = W_2 - W_1 \qquad (17\text{–}40)$$

We have shown three shells outside the M shell, but there may be more. These outside orbits, shells or levels are known as the optical shells or levels.

We have spoke of the valence electron being knocked out by the impact of an electron from outside the atom. However, if the impacting electron has sufficient energy, it may knock an electron from a deeper shell such as the L or the K shell. If in the case of the sodium atom represented in Fig. 17.8 an electron is knocked out of the L shell, it may go to the M shell or to one of the optical shells. The electron may then return to any inner shell provided that shell is not complete. When the electron is knocked out of the L shell, the electron in the M shell may fall into the vacancy left in the L shell.

If the impinging electron has still more energy, it may knock an electron from the K shell. In the case of an electron being knocked from the K shell of the sodium atom, this struck electron cannot go to the L shell because this shell is already complete, and so must go to the M or one of the optical shells. An electron may now fall from the L shell to fill the vacancy in the K shell or the electron in the M shell may fall to the K shell and so fill the vacancy in the K shell. If, however, an electron falls from the L shell to the K shell, a vacancy is left in the L shell, so that an electron from the M or an optical shell now falls to the L shell and fills the vacancy in that shell. With each of these transitions a quantum is emitted according to a relation similar to (17–40).

17.10. X-Ray Spectra.—The cathode rays which hit the target in an x-ray tube are impacting electrons. Traveling with a velocity of the order of 50,000 to 100,000 volts, they possess enough energy to penetrate into the atoms of the target and to knock out K, L and M electrons. In the case of molybdenum, whose atom contains $2K$, $8L$, $18M$, $13N$, and $1O$ electrons, the K, L and M shells are complete. A K electron must therefore be knocked out at least to the N shell, which is incomplete. If a K electron has been knocked out, an L electron may fall into the vacancy in the K shell, so that a quantum $h\nu_\alpha$ corresponding to the $K\alpha$ line is emitted according to the relation

$$h\nu_\alpha = W_L - W_K \tag{17-41}$$

where W_K and W_L are the energies of electrons in the K and L shells, respectively. If an electron falls from the M shell to fill the vacancy in the K shell, a quantum $h\nu_\beta$ corresponding to the $K\beta$ line is emitted according to

$$h\nu_\beta = W_M - W_K \tag{17-42}$$

where W_M is the energy of an electron in the M shell. Similarly, a quantum $h\nu_\gamma$ corresponding to the $K\gamma$ line is emitted when an electron falls from the N shell to fill the vacancy in the K shell. Thus, when owing to many impacts a K electron is knocked out of each of many molybdenum atoms, then in a fraction of the struck atoms an electron falls from the L shell to fill the vacant space in the K shell, in a second fraction of the struck atoms an electron falls from the M shell to fill the vacant space in the K shell, and in a third fraction of the struck atoms an electron falls from the N shell to fill the vacant space in the K shell. These various electron transitions produce the α, β and γ lines of the molybdenum K spectrum. Since the intensity of the lines diminishes in the order α, β, γ, the probability of the respective electron transitions diminishes in the order $L \rightarrow K$, $M \rightarrow K$, $N \rightarrow K$. There may be transitions such as $O \rightarrow K$, but the number of such transitions is relatively so small that no spectrum line on the short wave-length side of the γ line is observed.

CHAPTER XVII

PROBLEMS

1. A neutron (see Chapter XX) is an uncharged particle of mass 1.67×10^{-24} gm. What is the de Broglie wave-length of a neutron moving with a kinetic energy of 0.04 ev? At what glancing angle will a beam of such neutrons be

reflected in the first order from the (1, 0, 0) planes of a crystal of sylvine ($d_{100} =$ 3.14 A)?

2. Sylvine may be treated as a simple cubic crystal. Determine the energy and speed of each of the electrons which are reflected in the first order at a glancing angle of 10° from the (1, 1, 1) planes of sylvine.

3. Assume that a uranium atom ($Z = 92$) is stripped of all its electrons save one K electron. From the Bohr theory, calculate (*a*) the radius of the orbit of the electron, (*b*) the area of the circle bounded by this orbit, and (*c*) the kinetic energy of the electron in the orbit. For simplicity, make no corrections for the relativity change of mass at high speeds in this problem.

CHAPTER XVIII

CRITICAL POTENTIALS

18.1. Binding Energy.—Referring to (16–28) we see that the energy of the electron in the hydrogen atom when the electron is in the *n*th Bohr orbit is

$$W_n = -\frac{2\pi^2 me'^4}{n^2h^2} \tag{18–1}$$

If this energy is expressed in electron-volts as in (16–34) we have

$$W_n = -\frac{13.6}{n^2} \tag{18–2}$$

The meaning of the negative sign is that energy must be added to the electron in order to separate it completely from the nucleus. This energy which must be added to the electron in order to separate it completely from the nucleus is known as the *binding energy* of the electron. Another name for binding energy is work function. The binding energy of an electron in an atom is analogous to the kinetic energy which a body must possess in order to escape from the earth. Because of the way it is defined the binding energy is positive. The binding energy of an electron in the *n*th Bohr orbit of a hydrogen atom is E_n where

$$E_n = -W_n = \frac{2\pi^2 me'^4}{n^2h^2} \text{ ergs} \tag{18–3}$$

or expressed in electron-volts

$$E_n = \frac{13.6}{n^2} \tag{18–4}$$

Hence, if an energy equal to that obtained by an electron falling through 13.6 volts can be given to the electron in a hydrogen atom, the electron should be expelled and the hydrogen atom would then consist of a nucleus carrying a charge $+e$, but with no electron revolving about it. The atom as a whole then carries a positive charge. Such an atom is said either to be an ion or to be ionized.

A hydrogen atom can be ionized in two ways—(1) by knocking an electron out of the atom by means of another electron, or (2) by ejecting it by means of a quantum of light as in the photoelectric effect. We have studied the second case in Chapter XI. The threshold frequency of hydrogen should, according to the Bohr theory, be given by

$$h\nu = \frac{2\pi^2 m e'^4}{h^2} \tag{18-5}$$

or the threshold wave-length by

$$\lambda = \frac{h^3 c}{2\pi^2 m e'^4} \tag{18-6}$$

Using (11–13) and (18–4), we obtain

$$\lambda = \frac{12{,}345}{13.6} = \underset{\simeq}{908}\text{ A} \tag{18-7}$$

which is a wave-length in the far ultraviolet. Light of shorter wave-length than this will ionize a hydrogen atom, while light of longer wave-length will not ionize the atom.

18.2 Ionization Potential.—Let us consider the first way of ionizing a hydrogen atom. This may be done by hitting a hydrogen atom with an outside electron which has an energy of 13.6 ev or more. We shall for convenience call this outside electron the impinging electron in order to distinguish it from the atomic electron which was originally part of the atom and which after the impact is expelled from the atom. If the impinging electron possesses an energy greater than 13.6 ev, there is more than enough energy to knock the atomic electron out of the atom. The remaining energy is shared between the impinging and atomic electrons. Ionization potential is defined as that potential through which the impinging electron must fall in order to acquire sufficient energy to knock an electron out of an atom. In the case of the hydrogen atom, the ionization potential is 13.6 volts.

18.3. Radiation Potentials.—Again, referring to the Bohr theory, it is seen that the binding energy for the electron in the second Bohr orbit of the hydrogen atom is $13.6/2^2 = 3.4$ ev. The difference of the binding energies of the first and second Bohr orbits is therefore $13.6 - 3.4 = 10.2$ ev. Hence, if the electron in the normal or lowest orbit is given an energy equivalent to 10.2 ev, it can be raised to the second Bohr orbit. This can be done by allowing an outside electron with an energy of 10.2 ev or more to hit the electron in the hydrogen atom. In this case the hydrogen atom is not ionized for the electron is not knocked completely out of the atom. It might be said that the atom is incompletely or partially ionized. However, it has become the practice to say that the atom is in an *excited state* or that it is excited. When the hydrogen atom has its electron in the first Bohr orbit, the atom is said to be normal or in the normal state; when the electron is in the *second* Bohr orbit, the atom is said to be in the *first excited state;* when the electron is in the *third* Bohr orbit, the atom is in the *second excited state;* and so on. A radiation potential is defined as a potential through which the impinging electron must fall in order to acquire sufficient energy to knock an electron in the struck atom from the normal orbit to an outer orbit. In the case of hydrogen there are several radiating potentials but only one ionizing potential. The radiating potentials of hydrogen are $13.6 \times (1 - \frac{1}{2^2})$, $13.6 \times (1 - \frac{1}{3^2})$, $13.6 \times (1 - \frac{1}{4^2})$, etc., or 10.2, 12.1, 12.75, etc., volts. These are known as the first, second, third, etc., radiation potentials of hydrogen, respectively.

Radiation and ionization potentials are included in the general term *critical potentials.* Critical potentials are sometimes called *excitation potentials* when the potential is used to excite one or more spectrum lines. When, for instance, the impinging electron has an energy equivalent to 10.2 ev, it can knock the electron of a hydrogen atom to the second Bohr orbit. Later on this electron falls back to the first Bohr orbit and a quantum whose wave-length $\lambda = 12{,}345/10.2 = 1210$ A is emitted. Thus, by hitting the hydrogen atom with an outside electron which has fallen through a potential of 10.2 volts, the atom is caused to emit the spectrum line whose wave-length is 1210 A. Hence the potential of 10.2 volts is known as a radiating potential.

18.4 Application to X-Ray Spectra.—Although the ionization and radiation potentials can be calculated in the case of hydrogen, ionized helium and for doubly ionized lithium, these potentials cannot be calculated for other atoms. However, although these potentials cannot be calculated in general, they can be measured. We already have spoken of the excitation potentials for x-ray spectra in Sec. 14.22. The excitation potential for the K spectrum of molybdenum is 20,000 volts (see Table 14.4.) In our present view this means that when the impinging electrons have fallen through a potential of 20,000 volts or more they have enough energy to knock a K electron completely out of the molybdenum atom. The K spectrum is then excited as we have described in

Sec. 17.10. In this case the molybdenum atom is ionized, and 20,000 volts is the ionization potential for the K electrons. The ionization potential gives the binding energy, and so the binding energy of the K electron in the molybdenum atom is 20,000 ev. According to (16–28) and to Sec. 18.1 the binding energy of the electron in a hydrogen-like atom (an atom consisting of a nucleus with charge $+Ze$ and one electron) is

$$E_n = \frac{2\pi^2 m Z^2 e'^4}{n^2 h^2} \text{ ergs} \tag{18-8}$$

or expressed in electron-volts

$$E_n = 13.6 \times \frac{Z^2}{n^2} \checkmark \tag{18-9}$$

For the first Bohr orbit, $n = 1$, and for molybdenum $Z = 42$, so that the binding energy is 24,000 ev, which agrees at least as regards order of magnitude with the 20,000 ev obtained experimentally. However, the molybdenum atom is decidedly not hydrogen-like, since it contains 42 electrons, and it is no wonder that the above calculated value of 24,000 ev for the binding energy of a K electron in a molybdenum atom differs by 20 per cent from the experimental value of 20,000 ev.

In Sec. 14.22 we have mentioned the fact that there are three and five excitation potentials for the L and M x-ray spectra, respectively. The K, L, and M excitation potentials for molybdenum are shown in Table 18.1. Also two of the N excitation potentials are shown. The explanation of the three L excitation potentials is that instead of there being one L shell containing 8 electrons as we have described in Sec. 17.7, there are really three L shells. However, as the three L excitation potentials are fairly close together, the three L shells are close together, so that in the simplified theory of Sec. 17.7, we have shown the three L shells as one shell. Also, in the case of the five M shells, the excitation potentials are all at least of the same order of magnitude and of a dis-

TABLE 18.1

EXCITATION POTENTIALS OF MOLYBDENUM

	K	L_3	L_2	L_1	M_5	M_4	M_3	M_2	M_1	N_7	N_6
Volts	20,000	2870	2630	2520	510	413	402	236	233	65	41
Number of electrons	2	2	2	4	2	2	4	4	6	2	2

tinctly different order of magnitude from the L excitation potentials, so that again we were justified as representing the five M shells by one shell in the simplified theory of Sec. 17.7. These multiple L, M and N shells represent different energy levels. The three L and five M energy levels are sometimes called sublevels.

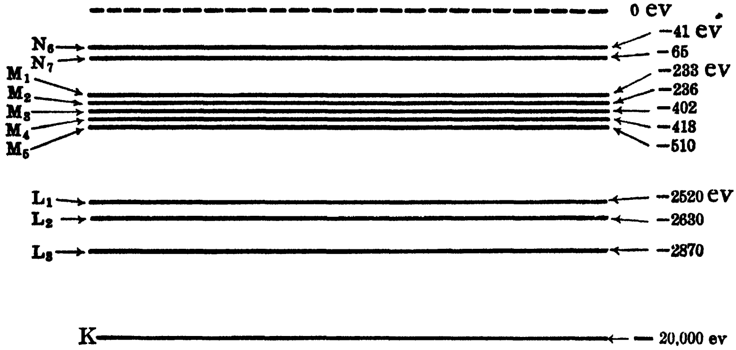


FIG. 18.1—Energy level diagram for molybdenum.

The excitation potentials shown in Table 18.1 are ionization potentials and are therefore measures of the binding energy of the electrons in the respective shells. Therefore an x-ray energy level diagram can be constructed as shown in Fig. 18.1. The figure is not drawn to scale. When an electron has been knocked out of the K level, one of the L electrons may fall into the vacant space in the K level. Hence, it would seem that we should observe three x-ray spectrum lines having the respective wave-lengths

$$12{,}345/(20{,}000 - 2870),$$
$$12{,}345/(20{,}000 - 2630),$$
$$12{,}345/(20{,}000 - 2520)$$

or, respectively, 0.721, 0.712, 0.708 A. In Table 14.2 the wave-length of the $K\alpha$ line of molybdenum is given as 0.710 A. However, if an x-ray spectroscope of high resolving power is used, this line is found to be a doublet whose components have the wave-lengths 0.712 and 0.708 A. The line does not contain a component of wave-length 0.721 A. This leads us to a very important principle which Bohr was compelled to state. This principle is known as the *Selection Principle.* The full theory underlying this principle is beyond the scope of this book, so we shall content ourselves with the following statement of the principle: *Although all spectrum lines are produced by electron transitions from the upper to the lower levels of an energy level diagram, yet not all transitions which can be drawn in an energy level diagram actually occur.* Thus, if there is a vacant space in the K level of molybdenum, an electron transition from the L_2 to the K level or from the L_1 to the K level may occur, but a transition from the L_3 to the K level cannot and does not occur. The Selection Principle therefore states that certain transitions are not allowed and that spectrum lines corresponding to such transitions are forbidden. The rules for determining those transitions which are forbidden is beyond the scope of this book. The $K\beta$ line of molybdenum is also a doublet and is produced by transitions from two of the M levels to the K level. The $K\gamma$ line is also a doublet and is produced by transitions from certain N levels to the K level.

The normal number of electrons in the various sublevels of molybdenum are shown in Table 18.1. Since there are twice as many L_1 electrons as L_2 electrons, it would seem that the chance of an L_1 electron falling into a vacant space in the K level should be twice as great as the chance of an L_2 electron falling into the vacant space. Hence it would seem that the $K\alpha_1$ line which corresponds to the $L_1 \rightarrow K$ transition should be twice as intense as the $K\alpha_2$ line which corresponds to the $L_2 \rightarrow K$ transition. This is found to be so experimentally, the shorter wave-length component of the $K\alpha$ doublet being twice as intense as the longer wave-length component.

18.5 Measurement of Critical Potentials.—The measurement of the x-ray critical potentials has been described in Sec. 14.22. The method of measurement of what might be called the optical critical potentials involves a different technique. The experimental arrangement for determining the optical potentials is shown diagrammatically in Fig. 18.2. A filament F is heated by the battery A. A battery B_1 is connected to the filament F and the grid (or wire gauze) G_1, so that there is a difference of potential of V_1 volts between F and G_1. The direction of the force acting on a negative electron between F and G_1 is in the direction of the arrow V_1. A battery B_2 is connected to G_1 and a second grid G_2, giving rise to difference of potential of V_2 volts between G_2 and G_1. The direction of the force on an electron between G_2 and G_1 is in the direction of the arrow V_2. A battery B_3 is connected between G_2 and ground. The plate P is connected to an electrometer E and K is a grounding key. The filament F, grids G_1, G_2, and the plate P are in a glass tube (not shown) which can be evacuated. The battery B_3 can be reversed, so that the direction of the force acting on an electron between G_2 and P can be reversed as indicated by the arrows V_3.

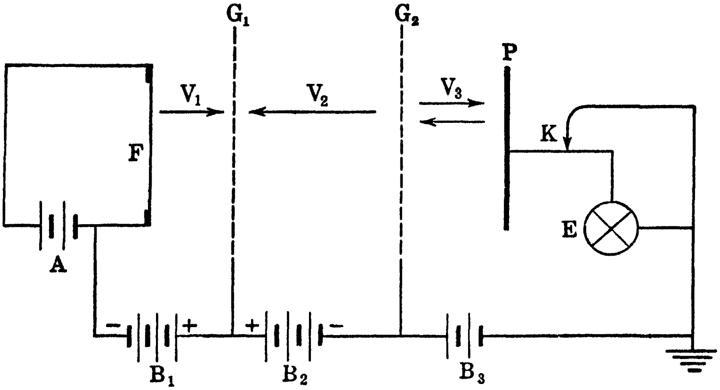


FIG. 18.2.—Apparatus for measuring critical potentials.

The glass tube (not shown) is filled with a gas, such as helium or oxygen, whose critical potentials it is desired to measure. The pressure is then reduced to about 1/10 mm of mercury. The filament is lighted so as to give a supply of electrons. Under the action of the potential V_1 the electrons arrive at G_1 with a speed v given by

$$(1/2)mv^2 = \frac{V_1 e'}{300} \tag{18-10}$$

where V_1 is measured in volts and $e' = 4.80 \times 10^{-10}$ esu. However, it is more convenient to express the energy in electron-volts rather than in ergs. Some of the electrons pass through the holes or spaces in the grid G_1 into the space between G_1 and G_2. If the electrons have sufficient energy, they are able to knock electrons out of the gas atoms between G_1 and G_2. If they do this, the atoms become ionized positively. These positive ions under the action of the electric field between G_1 and G_2 move toward the right, which is in the opposite direction to that indicated by the arrow V_2. If at first V_3 is made zero, some of these ions pass through the grid G_2 to the plate P. If, now, the grounding key K is raised, the ions give P a positive charge, and this is communicated to the electrometer, which then shows a certain rate of deflection depending on the rate at which ions arrive at P. Care must be taken that none of the electrons from the filament F arrive at P. This is accomplished by making $V_2 > V_1$, for then all electrons which come through the grid G_1 and are traveling to the right are turned back before they reach G_2.

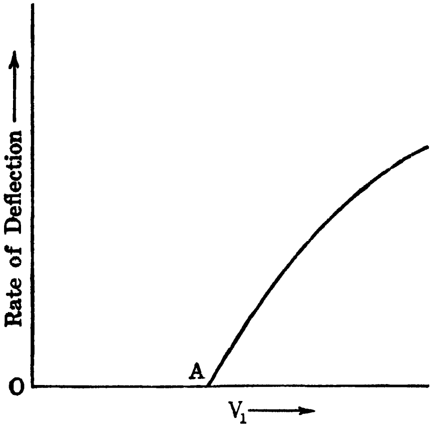


FIG. 18.3.—The critical potential is represented by OA.

Values of the rate of deflection of the electrometer for the corresponding values of the voltage V_1 are now obtained, and a curve plotted as in Fig. 18.3. It is found that for small values of V_1 (always keeping $V_2 > V_1$) the electrometer shows no deflection. After a certain value of V_1 represented by OA in Fig. 18.3 is reached, the electrometer begins to show a deflection, the rate of which increases as the difference between V_1 and the value represented by OA increases. The value in volts represented by OA is a critical potential. The problem now is to determine whether OA represents an ionization or a radiation potential.

18.6. Discrimination between Ionization and Radiation Potentials.—When an electron hits a gas atom with an energy equivalent to the voltage represented by OA, Fig. 18.3, the atom may either be ionized or excited. For instance, in the case of hydrogen, if the impinging electron has an energy of 10.2 volts, the electron in the hydrogen atom may be raised from the normal Bohr orbit to the second Bohr orbit. The atomic electron subsequently falls back to the normal orbit with the emission of a quantum corresponding to the wave-length 12,345/10.2 =

1210 angstroms. Now, light of this wave-length is very active photoelectrically, so that light of this wave-length which falls on the plate *P* ejects photoelectrons and the plate is left with a positive charge which is registered by the electrometer. Also, in the case of hydrogen, if the impinging electron has an energy of 13.6 ev, the atomic electron is knocked out of the atom and the atom becomes a positive ion. Some of the ions thus produced are pulled by the electric field between G_1 and G_2 over to the plate *P*. Hence, no matter whether *OA* represents a radiation or an ionization potential, the plate *P* builds up a positive charge when V_1 is greater than the potential represented by *OA*. So far we have assumed that the potential V_3 in Fig. 18.2 is zero. If, however, V_3 is not zero but is in a direction such that the electric force acting on an electron between G_2 and *P* is towards *P*, a photoelectron ejected from the plate *P* will be turned back into the plate *P* if the magnitude of V_3 is properly adjusted. If the positive charge on *P* is due to the photoelectric effect, a positive charge is built up on *P* when V_3 is in the direction of the lower arrow in Fig. 18.2. (the arrow represents the direction of the force acting on the electron), but a positive charge is not built up on *P* when V_3 is in the direction of the upper arrow. On the other hand, if the positive charge on *P* is due to positive ions, a positive charge is built up on *P* no matter what the direction of V_3, provided that $V_3 < V_2$.

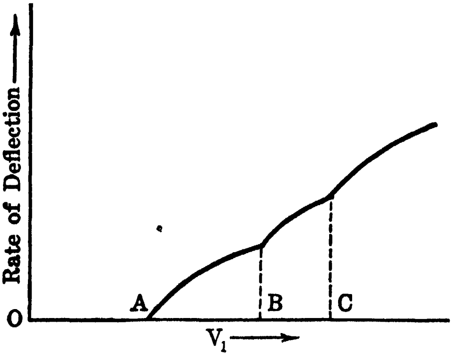


FIG. 18.4.—The abscissa of each kink in the curve represents a critical potential.

Whereas there is only one ionization potential for the normal hydrogen atom, there are several radiating potentials. Referring to the energy level diagram for hydrogen in Fig. 16.11 we see that the radiation potentials for the hydrogen atom are (13.6 − 3.4), (13.6 − 1.51), (13.6 − 0.84), etc. respectively, or 10.2, 12.09, 12.76, etc. volts, respectively. A curve such as shown in Fig. 18.3 is obtained for the first radiation potential when V_3 is in the proper direction or for the ionization potential when V_3 is in such a direction as to suppress the effect of the radiation potentials. If, however, the second and third radiation potentials are being measured, the curve in Fig. 18.3 has the form shown in Fig. 18.4. There are several kinks in the curve. The abscissa of each kink represents a radiation potential.

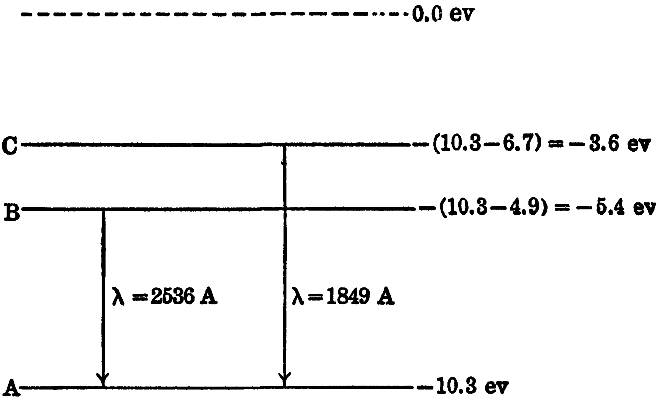


FIG. 18.5.—Energy level diagram for mercury vapor.

The radiation potentials found for mercury vapor are 4.9 and 6.7 volts, while the ionization potential is 10.3 volts. We can therefore construct the energy level diagram shown in Fig. 18.5. A transition from the *B* to the *A* level gives rise to the line whose wave-length is 12,345/4.9 = 2520 A. The wave-length of the actual spectrum line is 2536 A. The agreement is as good as can be expected because it is difficult to measure critical potentials to a greater accuracy than 0.1 volts. For perfect agreement the first radiating potential needs to be 4.87 volts. The transition *C* → *A* gives rise to the line whose wave-length is 12,345/6.7 = 1849 A. This line is observed in the mercury spectrum.

18.7 Resonance Radiation.—An electron in an atom can be raised from one energy level to a higher level by supplying it with an energy equal to the difference of the energies in the two levels. For instance, if the impinging electron has an energy of 10.2 ev, the electron within a hydrogen atom can be raised from the normal or first Bohr level to the second level. If, however, the impinging electron has an energy of 10.6 ev, which is more than enough to raise the atomic electron from the first to the second level, the atomic electron is raised to the second level and the impinging electron retains an energy of $10.6 - 10.2 = 0.4$ ev. We have seen in Chapter XI that a light quantum can eject an electron from an atom provided only that the frequency of the quantum is greater than the threshold frequency. The question now arises as to whether a light quantum can excite an atom without ionizing it. The answer is that the quantum can only excite but not ionize an atom if the energy of the quantum is exactly equal to the difference of the energy of the electron in the two levels. A quantum can raise the electron from the normal to the second level in the hydrogen atom if the quantum has an energy exactly equivalent to 10.2 ev—that is, if the wave-length of the light is $12{,}345/10.2 = 1210$ A. If the wave-length differs at all from 1210 A, the hydrogen atom is not excited. However, if the light does have a wave-length of 1210 A, the light is absorbed and the hydrogen atom is excited. Later on, the electron drops back to the original level and a quantum of wave-length 1210 A is emitted. Because this emitted light is of the same wave-length as the original light, it is called resonance radiation. If the resonance and the original radiation are of exactly the same wave-length, how can they be distinguished? They are distinguished by the fact that the original light is in a beam whose direction is under the control of the experimenter while the resonance radiation is equally intense in all directions. When the electron of a particular atom drops back from the excited level to the normal level, it has no memory of the direction from which the exciting quantum came, and so the emitted quantum is just as likely to be shot out in one direction as in any other direction.

One of the most famous cases of resonance radiation is that of the line 2536 A in the mercury spectrum. If this line is sent through mercury vapor, the line is very strongly absorbed. Consider the arrangement shown diagrammatically in Fig. 18.6. Light from a mercury arc lamp *A* is concentrated on a resonance lamp *B* by means of the quartz lens *L*. Both lamps are made of quartz so as to allow the passage of ultraviolet light. The lamp *B* contains mercury vapor at room temperature. It is found that the line 2536 A, which occurs in the spectrum of the light coming from *A*, causes the mercury vapor in *B* to give out light of this wave-length in all directions.

The name "resonance radiation" comes from the classical theory explanation of this phenomenon. According to the classical theory, the mercury atoms in *A* are set vibrating with a frequency corresponding to a wave-length of 2536 A. These cause the emission of electromagnetic waves of the same frequency. These waves cause resonance vibrations in the mercury atoms in *B* and these vibrations of the atoms in *B* cause the emission of waves of the wave-length 2536 A in all directions. Although this explanation of resonance radiation has been abandoned, the name "resonance radiation" has been retained.

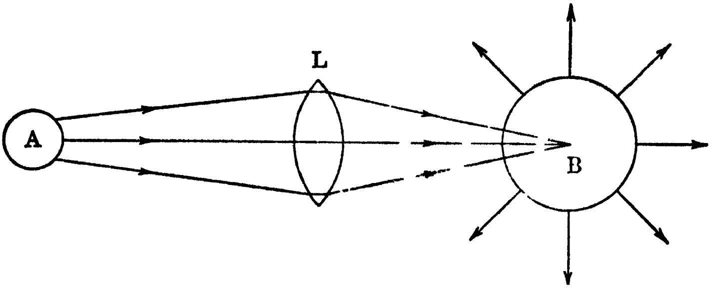


Fig. 18.6.—Experimental arrangement for showing resonance radiation.

18.8. Motion of Atomic Electrons.—We have seen in Sec. 14.26 that in the extended theory of the Compton effect, Jauncey's equation

$$\lambda' - \lambda = (h/mc) \cdot (1 - \cos \phi) - 2\lambda(v/c) \sin {}^{\phi}/_{2}. \qquad (18\text{–}11)$$

was obtained. In this v represents the component of the velocity of the electron in the direction of the bisector of the angle *ABC* in Fig. 14.31. Now, according to the Bohr theory of the atom as described in Chapter XVI, and also according to the modern form of the quantum theory as described in Chapter XVII, the electrons in an atom are moving about in various orbits. As an electron swings around in its orbit, the component of its velocity in a certain direction changes. Hence, if v in (18–11) is the component velocity of an atomic electron, v will change from positive to negative values, with the result that the change of wave-length of x-rays on scattering itself varies from one instant to another. The velocity of an electron in its orbit is so great and the radius of the orbit so small that the frequency of the changes of v from positive to negative is some 10^{15} to 10^{17} cycles per second. The result on the photographic film used for observing the Compton effect is that the Compton modified line α' (see Fig. 14.31) has a width greater than that of the unmodified line or that of the primary x-rays. The magnitude of the component velocity v cannot be greater than the speed V, of which v is a component. The values of v are therefore between $-V$ and V or vary over a range $2V$. Hence from (18–11) the width w of the Compton modified line is

$$w = 4\lambda(V/c) \sin {}^{\phi}/_{2} \tag{18–12}$$

while the center of the line is at a wave-length separation of

$$\lambda' - \lambda = (h/mc) \cdot (1 - \cos\phi) \tag{18–13}$$

from the wave-length of the primary rays. For scattering of Mo $K\alpha$ rays from carbon at an angle of 156°, DuMond found a width of 0.070 A. Putting $\lambda = 0.71$ A in (18–12) and solving for V, we obtain $V = 7.2 \times 10^8$ cm. In the Bohr theory of the atom, we obtained (16–23) for the radius of a circular orbit. Substituting this value for r in (16–22) we obtain

$$V = 2\pi Ze'^2/nh \tag{18–14}$$

for the speed of the electron in its orbit. Putting $Z = 6$ for carbon, we obtain $V = (13.1/n) \times 10^8$ cm/sec. For $n = 1$, V is 13.1×10^8 cm/sec and for $n = 2$ it is 6.55×10^8 cm/sec. The value obtained from DuMond's experiments is between these two values. The scattering is from both the K electrons ($n = 1$) and the L electrons ($n = 2$), so that this result is very satisfactory evidence for the motion of the electrons in an atom.

CHAPTER XVIII

PROBLEMS

1. Assume that the following atoms are stripped of all their electrons save one electron. Calculate the binding energies of a K electron of copper, silver, tungsten, lead. Compare with the binding energies given in Table 14.4. Note that for high values of Z, the percentage difference becomes less, being about 24 per cent for copper and 2.7 per cent for lead. For hydrogen the difference is zero.

2. Mercury vapor is monatomic. Calculate the root mean square speed of mercury atoms in mercury vapor at 400° C (boiling point = 357° C, atomic weight = 200.6). Assuming that the atoms are traveling in random directions with this speed, calculate the width of the absorption line in mercury vapor at 400° C. (Use the Doppler principle).

3. Silver $K\alpha$ x-rays are scattered from beryllium ($Z = 4$) at an angle of 160°. Calculate the velocity of a K electron and the width of the Compton modified line scattered by the K electrons.

CHAPTER XIX

RADIOACTIVITY AND ISOTOPES

19.1. Early Years of Radioactivity.—We have seen in Sec 7.6 that Plücker discovered cathode rays in 1859. However, it was not until 1897 that the true nature of cathode rays was recognized by J. J. Thomson. As a contrast, we saw in Sec. 14.1 that Roentgen's discovery of x-rays and his realization that these rays were something new and startling came suddenly on November 8, 1895. The realization of the true and startling nature of radioactivity took over six years as compared with the thirty-eight years required for the recognition of the electron. In brief, we can say that the discovery of x-rays was sudden, that of the electron slow, and that of radioactivity intermediate.

We have noted in Sec. 14.1 that in early 1896 J. J. Thomson and others believed that the emission of x-rays by a Crookes tube was connected with the greenish-yellow fluorescence or phosphorescence of the glass walls of the tube. It seemed worthwhile to search for penetrating radiations from substances that are phosphorescent under the action of x-rays or light. Balmain's paint, whose active ingredient is calcium sulfide, phosphoresces for several hours after exposure to light. Several phosphorescent substances were examined for rays which would penetrate black paper, but no certain results were obtained until February 24, 1896, when Henri Becquerel of France found that a uranium salt which had been exposed to strong sunlight emitted rays which could penetrate black paper and a thin sheet of aluminum and then blacken a photographic plate. However, Becquerel believed that he was observing a phenomenon of phosphorescence. Several days passed with overcast skies and in his impatience Becquerel developed a plate which had been enclosed in a plate-holder on the outside of which a button of uranium salt had been placed. The button and plate-holder had lain in a closed drawer for several days. Yet, on development, the plate under the button was quite as black as if the button had been exposed to direct sunlight. This remarkable result was obtained on March 1, 1896. Becquerel was immediately alert. He dissolved uranium nitrate in the dark. Solution of a phosphorescing salt removes its phosphorescence. Becquerel then recrystallized the uranium nitrate in the dark and found that the recrystallized salt emitted penetrating rays. As the weeks went by Becquerel tried various compounds of uranium. Finally, he tried uranium metal and found that it, like the compounds of uranium, emitted penetrating though invisible radiation. Nevertheless, Becquerel still believed that he was observing a phenomenon of phosphorescence. But the persistence of the invisible radiation puzzled him. The phosphorescence of Balmain's paint dies out after a while. On the other hand, the uranium rays persisted with undiminished intensity for a period of months and then of years. Becquerel attributed the persistent rays from uranium and its compounds to the proportion of uranium present. He said that these rays were a property of the element uranium.

Research on the peculiar rays from uranium stagnated during the latter part of 1896 and during the whole of 1897. The concept of phosphoresence seemed to be the stumbling block. Then, in the earlier part of 1898, Marie Curie, working in Paris, and Schmidt in Germany discovered that thorium compounds also emit penetrating, though invisible, rays which persist indefinitely. In the late 1890's uranium was the element of highest atomic weight and thorium was the next highest. It is strange from our present viewpoint that it had not previously occurred to anyone to test thorium for radioactivity, a term used by the Curies—Pierre and Marie—in 1898. However, the discovery of the radioactivity of thorium meant that the radioactivity of uranium was no longer a unique and anomalous property of uranium. Marie Curie stated that radioactivity is a property of the atoms of the radioactive substance. This concept was a distinct advance in the thinking of those days.

It was early noted by Becquerel that the penetrating rays from uranium and its compounds produces ionization in a gas similar to the ionization produced by x-rays. The Curies used the ionization method to measure radioactivity. They found that pitchblende, a metal containing uranium, was more radioactive than it should be according to the fraction of uranium present. The Curies proceeded to the decomposition of pitchblende and its separation, by chemical analysis, into its constituent elements, testing each preparation for its radioactivity. They obtained "radiferous" bismuth and barium sulfides, which were both strongly radioactive. Finally, they obtained radiferous bismuth sulfide which was 400 times as radioactive, mass for mass, as uranium.

The Curies claimed that they had found a new element, which they called "polonium" in honor of the country of Mme. Curie's birth. However, in 1898 the existence of a new element in small quantities was accepted only if its spark spectrum could be obtained. No such spectrum was presented and the Curies' claim was received with reservation.

The Curies then worked on the radiferous barium chloride which they obtained from the sulfide by chemical means. By fractional precipitation of the chloride from solution in water by alcohol, they found an increasing radioactivity in the precipitate. Finally at the end of 1898 they obtained a fraction possessing a radioactivity of 900 times that of uranium. Since spectroscopic analysis of this fraction showed a new spectrum line, it was conceded that the Curies had discovered a new element. They called this new element *radium* because of its high radioactivity. After more than three years of the most exacting and tedious effort, Mme. Curie obtained a sample of radium chloride spectroscopically free from barium. The atomic weight of the radium in this sample was found to be 225, which is quite different from 137, the atomic weight of barium.

In 1899 Rutherford, then of McGill University, Canada, turned his attention to the radiations emitted from uranium. He measured the absorption of the rays in aluminum by means of an ionization chamber. He assumed an exponential law of absorption such as given in (14–7) and found that the rays from a powdered uranium compound consisted of a soft component with an absorption coefficient $\mu = 1600 \text{ cm}^{-1}$ and a hard component with $\mu = 15 \text{ cm}^{-1}$. Rutherford called the soft component α-rays and the hard component β-rays. In 1900 Villard of France found exceedingly penetrating rays from a sample of a radium salt and these came to be known as γ-rays.

In 1900 Rutherford announced the isolation of thorium emanation. This is a radioactive gas at ordinary temperatures. Previous to 1900 the persistence of the rays from radioactive substances had been the outstanding property of these substances. But the rays from thorium emanation and, later, from radium emanation were found to die out as time passed. Rutherford also discovered radioactive substances which were deposited from the thorium emanation on solid substances and particularly on negatively charged metals if air was present. The radioactivity of the deposit obtained after long exposure to the thorium emanation decays exponentially with the time after removal from the thorium emanation according to

$$I = I_0 e^{-\lambda t} \tag{19-1}$$

where I_0 is the activity soon after removal, I is the activity as measured at a time t after the measurement of I_0, and λ is a constant known as the radioactive constant. A curve which represents (19–1) is known as a *decay* curve. Rutherford found that for short exposures of a negatively charged metal rod to thorium emanation, the activity of the rod increased with the time according to

$$I = I_0(1 - e^{-\lambda t}) \tag{19-2}$$

A curve which represents (19–2) is known as a recovery curve. In 1900 Rutherford obtained $\lambda = 0.063\ \text{hr}^{-1}$ in both (19–1) and (19–2).

After examining various decay and recovery curves, Rutherford and Soddy were led in 1902–03 to announce the theory of radioactive transformations. The true nature of radioactivity had at last been recognized. In Sec. 14.1 we have noted that according to the prevalent opinion held by physicists in early 1895 the atom was indestructible, impenetrable and immutable. The recognition of the electron in 1897 had shattered the concept of impenetrable atoms; Rutherford and Soddy's epochal theory of radioactive transformations shattered the concept of indestructible and immutable atoms.

Previous to 1903 the general public had little interest in radioactivity. Then early in 1903 Pierre Curie and Laborde discovered the heating effect of radium. It seemed at the time that radium provided an inexhaustible source of energy. Most extravagant claims were made for the miraculous powers of radium. Numerous articles appeared in the press and popular journals of the day. It was even suggested in 1903 that by means of the energy of radium the whole earth might be blown up and the end of the world brought about.

19.2. Effects Produced by α-, β-, and γ-Rays.—The α-rays are very easily absorbed in matter, being unable to penetrate more than a thickness of 0.1 mm of aluminum or of a few centimeters of air. The α-rays possess a very intense ionizing power. Provided that the α-rays can penetrate to a fluorescent screen or photographic plate, the fluorescent or photographic effect of the rays is intense. The β-rays are much less easily absorbed than the α-rays, being able to penetrate a thickness of 1 mm of aluminum. The ionizing, fluorescent, and photographic effects of γ-rays are much less than those of β-rays. Recent measurement of the heating effect of radium gives a rate of emission of 130 cal/hr from a gram of radium in equilibrium with its transformation products.

19.3. Nature of α-Rays.—We have mentioned in the previous section that α-rays produce fluorescence in certain substances. If the fluorescence is examined by means of a low-power microscope, it is found to consist of a series of scintillations which are produced by the impacts of the α-rays upon the fluorescent screen. These scintillations give optical evidence that the α-rays consist of discrete particles. We therefore speak of α-particles. Since the impact of each α-particle on the screen can be observed it is possible to count the number of α-particles which are emitted in a given time by a sample of radioactive substance in a small cone whose axis is in a certain direction.

If the radioactive substance is placed in a vacuum and the α-particles are allowed to enter a hollow metal vessel also in the vacuum, a positive charge is built up on the vessel. If the vessel is connected to an electroscope this charge can be measured. From previous observations of the scintillations produced by the α-particles on a fluorescent screen, the number of α-particles entering the vessel in unit time can be found. If then a charge Q is built up on the vessel in t seconds, the charge entering the vessel per second is Q/t. Since the number of α-particles entering the vessel per second is known, the charge on each α-particle can be found. This charge is 9.60×10^{-10} esu or $+2e'$, where e' is the charge on the hydrogen ion in electrolysis and $-e'$ is the charge on the electron.

If a beam of α-rays is passed down a long evacuated tube, and electric and magnetic fields are applied, it is found that the rays are deflected. The deflections produced by the electric and magnetic fields are respectively in the opposite directions to those produced by these fields on a stream of cathode rays as described in Chapter VII. This shows that the α-particles are positively charged. From the magnetic and electrostatic deflections, the velocity and the ratio of the charge to the mass of an α-particle can be found. The velocity of the particles varies, being of the order of 2×10^9 cm/sec, and the ratio of the charge to the mass is 4820 emu/gm. In Chapter VII, we have seen that the ratio of the charge to the mass of the hydrogen ion in electrolysis is 9568 emu/gm or almost twice the ratio for the α-particle. However, we saw in the previous paragraph that the charge on an α-particle is twice the charge on the hydrogen ion in electrolysis, so that the mass of an α-particle must be four times the mass of the hydrogen ion. Since we have seen in

Chapter VII that the mass of the hydrogen ion is 1840 times the mass of an electron, the mass of an α-particle is 4 × 1840 = 7360 times the mass of an electron. The mass of an α-particle is 6.64×10^{-24} gm. The atomic weight of helium is 4 and it therefore seems that α-particles might be ionized atoms of helium. This prediction was verified by Rutherford, who allowed the α-particles from a radioactive substance to penetrate through a very thin-walled glass tube into a discharge tube. At first, when an electric discharge was sent through the tube, none of the spectrum lines of helium was observed. After a few hours, sufficient α-particles had penetrated into the discharge tube and the spectrum of helium became visible, the spectrum becoming brighter as more α-particles penetrated into the discharge tube. The α-particle is therefore an atom of helium which has lost two negative charges and so carries a positive charge of $2e$. According to Sec. 17.7, the helium atom consists of a nucleus carrying a charge $+2e$ and two electrons, each carrying a charge $-e$. Hence an α-particle is the nucleus of the helium atom.

19.4. Range of α-Rays.—In 1899 Rutherford had distinguished α-rays from β-rays by means of their absorption coefficients in aluminum. Until 1904 it was assumed that all rays—α-, β-, γ-, and x-rays—were absorbed exponentially. The rays might be heterogenous but still the fundamental law of absorption was exponential. However, in 1904 it occurred to W. H. Bragg, then of Adelaide, Australia, that α-particles may be absorbed according to a law which is not exponential. He showed that, after the α-particles have proceeded a certain distance in air, they lose their ionizing power. This certain distance is known as the range of the α-particles. The range of α-particles depends upon the particular radioactive substance from which the particles are emitted. The range of α-particles can be shown by means of an apparatus shown diagrammatically in Fig. 19.1. A thin layer of radioactive substance is placed at A at the bottom of a metal box. The α-rays emerge through the window B and form a cone as shown. They enter a very shallow ionization chamber consisting of the metal plate C and a wire gauze electrode D, which allows α-particles to enter the ionization chamber. The particles ionize the gas in the chamber and produce a deflection of an electroscope. The distance between D and the radioactive substance at A can be varied by means of a screw. If the rate of deflection of the electroscope is plotted against the distance of D from A, a curve such as shown in Fig. 19.2 is obtained. The ionization is at first almost independent of the distance, for small distances. At greater distance, it increases to a sharp maximum and then suddenly drops to zero. The distance *OA*, Fig. 19.2, is the range of the particles.

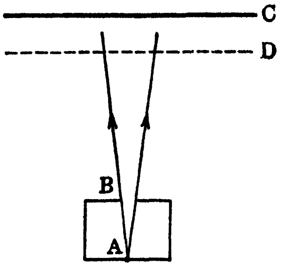


Fig. 19.1.—Apparatus for showing the range of α-particles.

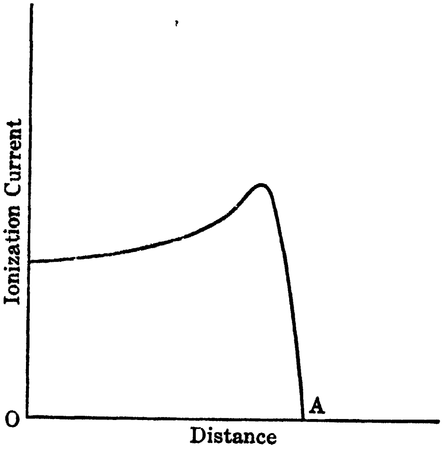


Fig. 19.2.—The length of *OA* gives the range of α-particles.

If a thin layer of radium is taken and allowed to stand for some time and the experiment to determine the range performed, a curve similar to Fig. 19.3 is obtained. The curve on close analysis is seen to be a composite of four curves similar to Fig. 19.2. There seem to be four distinct types of α-particles differing from each other in range—the ranges in air being 3.5, 4.2, 4.8 and 7.1 cm respectively. It may be remarked here that the range of any particular type of α-particle varies inversely as the pressure of the gas through which it passes. The ranges which we have just given are for air at atmospheric pressure. As the α-particles move through air they cause ionization by knocking electrons out of the atoms of the air.

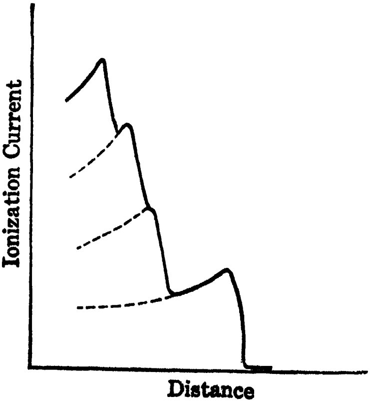


Fig. 19.3.—Composite curve given by α-particles of four different ranges.

An α-particle thus leaves a trail of positive and negative ions. As the particle proceeds along its path, it loses speed owing to the energy lost to the electrons which are knocked out of the atoms of the gas. Finally its speed reaches a certain critical value below which the α-particle is unable to knock electrons out of atoms.

The phenomena which we have just described may be classed as range absorption phenomena in contrast with exponential absorption phenomena which we studied under x-rays.

19.5. Nature of β-Rays.—If an electric field is applied to a beam of β-rays moving in an evacuated tube, the β-rays are deflected in the same direction as cathode rays in a discharge tube. Likewise, if a magnetic field is applied, the rays are deflected again in the same direction as cathode rays. Because the β-rays are deflected in a similar way to cathode rays by electric and magnetic fields, it was immediately believed that the β-rays consist of a stream of negatively charged particles. From measurements of the electrostatic and magnetic deflections, the velocity and the ratio of the charge to the mass of β-particles can be calculated according to the method described in Chapter VII. It is found that the velocity of the β-rays varies from 0.3 to 0.98 times the velocity of light and that the ratio of the charge to the mass is not constant as with cathode rays but has a maximum value of 1.76×10^7 emu/gm which is the value of e/m for cathode rays. However, we have shown in Chapter XIII that the relativity theory requires that the mass m of a body moving with a speed approaching that of light increase above the rest-mass m_0 according to the formula

$$m = \frac{m_0}{\sqrt{1 - \beta^2}} \tag{19-3}$$

where β is the ratio of the velocity of the body to that of light. Hence e/m for the β-particles is given by

$$\begin{aligned} \frac{e}{m} &= \left(\frac{e}{m_0}\right)\sqrt{1 - \beta^2} \\ &= 1.76 \times 10^7 \times \sqrt{1 - \beta^2}\ \text{emu/gm} \end{aligned} \tag{19-4}$$

In 1909 Bucherer of Germany, using a method where the β-particles were subjected simultaneously to electric and magnetic fields at right angles to each other, was able to sort out the particles according to their velocities and then to measure the e/m for those particles of a given velocity. He found that the e/m for the particles varied with the velocity according to (19–4). For instance, Bucherer found experimentally that the value of e/m for the particles of highest velocity, for which $\beta = 0.98$, was 0.35×10^7 emu/gm, while, according to (19–4) the value of e/m should be $1.77 \times 10^7 \times \sqrt{1 - .98^2} = 0.35 \times 10^7$, which is in exact agreement with the experimental value.

The β-particles do not exhibit a definite range in air as do the α-particles. This is due to the fact that they are high speed electrons and in knocking electrons out of atoms as they ionize these atoms they are deviated from their paths. A slowly moving β-particle has a mass which is about the same as that of an electron while a fast moving electron has a mass which may be as much as five times the rest-mass of an electron. In either case, the mass of the β-particle is not much different from the mass of an electron in the atom and so when an impact occurs the β-particle is deviated. In the case of an α-particle, however, the mass of this particle is $4 \times 1840 = 7360$ times as great as the rest-mass of an electron. The mass of the α-particles is thus so great that the particle knocks electrons out of its way without being deviated to any great extent. The α-particle travels in a straight line through air, while a β-particle follows an irregular and devious path.

19.6. Nature of γ-Rays.—When very strong electric and magnetic fields are applied to a beam of γ-rays, no deflection of the beam can be found. The γ-rays are therefore believed to be of the same nature as x-rays, differing from them in wave-length. As we have noted in Sec. 14.12, Stokes suggested that the sudden stoppage of cathode particles by the target of an x-ray tube gives rise to the production of x-rays. Some of the γ-rays may be produced by the sudden stoppage of β-particles in the radioactive material itself. In the early years of radioactivity it was believed that the sudden emission of β-particles from radioactive atoms produces γ-rays owing to the high acceleration of the β-particles. This idea required that β-ray emission is always accompanied by γ-ray emission. However, this is not always the case and another explanation of γ-rays must be sought.

The γ-rays approximately follow the exponential absorption law which we have previously noted for x-rays. Thus, if various thicknesses of iron are placed between a source of radioactive material which emits γ-rays and an ionization chamber, it is found that the intensity I of the rays which penetrate a thickness t of iron is given by

$$I = I_0 e^{-\mu t} \qquad (19\text{–}5)$$

where I_0 is the original intensity of the γ-rays and μ is the absorption coefficient.

The wave-length of γ-rays is measured by passing a beam of rays through a slab of rocksalt crystal as shown in Fig. 19.4. Reflection of the γ-rays proceeding from the source R occurs from the crystal planes at A and B within the crystal. The position of A is determined by the angle θ which is itself determined by Bragg's law

$$n\lambda = 2d \sin \theta \tag{19-6}$$

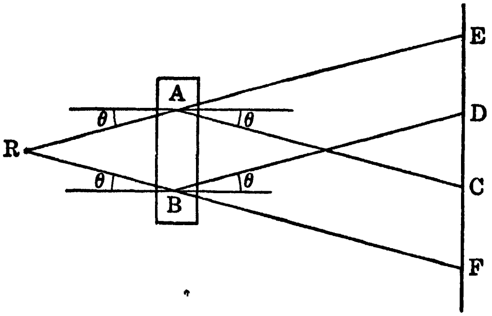


FIG. 19.4.—Experimental arrangement for determining the wave-length of γ-rays.

The reflected rays fall upon a photographic film at C and D and at these positions black lines appear on the developed film. It is seen that γ-rays from R enter the crystal at angles greater and less than that particular value of θ which satisfies Bragg's law. These rays pass through the crystal and cause a general blackening of the film, so that the black lines at C and D appear against a field of lesser blackening. However, a ray passing in the particular direction RA suffers an apparent extra absorption in penetrating through the crystal to E. This apparent extra absorption is due to the fact that some γ-rays have been reflected out of the beam RA into the beam AC. Hence γ-rays of less intensity arrive at E than at neighboring portions of the photographic film so that a white line instead of a dark line appears at E. This white line appears against a background of general blackening. Similarly a white line occurs at F. Obviously the angle EAC is twice the angle θ which appears in the Bragg equation. This method shows that the wave-length of γ-rays varies from about .005 angstrom to 0.4 angstrom. From Secs. 8.5 and 11.9, we may express the energy of a γ-ray quantum by means of the relation

$$\frac{\text{Energy of a quantum}}{\text{in electron-volts}} = \frac{12{,}345}{\lambda \text{ (in angstroms)}} \tag{19-7}$$

The energy of γ-rays from naturally radio-active substances varies from 31,000 to 2,500,000 ev.

Other methods for determining the energy of γ-rays are now in use. The method just described cannot be used for γ-rays whose energies are higher than 2.5×10^6 ev.

19.7. The Million-Electron-Volt.—The energies of α- and β-particles and of γ-ray quanta are so often of the order of a million electron-volts that a new unit of energy, the million-electron-volt, has come into widespread use in nuclear physics. We shall use the symbol "mev" for this new unit of energy. Since, according to Sec. 8.5 one electron-volt equals 1.6×10^{-12} erg, we have

$$1 \text{ mev} = 1.6 \times 10^{-6} \text{ erg} \tag{19-8}$$

Further, on account of the Einstein energy-mass relation $E = mc^2$,

$$1 \text{ mev is the equivalent of } 1.77 \times 10^{-27} \text{ gm} \tag{19-9}$$

The energy of the rest-mass of an electron is thus 0.51 mev, while the kinetic energy of an α-particle traveling with a velocity of 2×10^9 cm/sec is 8.25 mev. The equivalent energy of a mass equal to unit atomic weight (1.66×10^{-24} gm) is 934 mev or one mev is the energy equivalent of an atomic weight of .00107. Also in virtue of (19–7)

$$1 \text{ mev is the energy of a quantum of } \lambda = .0123 \text{ A} \tag{19-10}$$

The relation of the kinetic energy W of α-particles as they are emitted from the atoms of a radioactive substance to the range of these particles in air has been found experimentally. This relation is approximately expressed by

$$W(\text{in mev}) = 4.9(Rp/T)^{2/3} \tag{19-11}$$

where R is the range in cm of air, p the pressure of the air in cm of mercury, and T the temperature of the air in degrees absolute. This relation holds fairly well up to a range of 25 cm.

19.8. Radioactive Transformations.—According to recent determinations the atomic weight of radium is 226. Its chemical properties place it in the same column of the periodic table as calcium, strontium and barium. Its atomic number is 88. It is a silver white metal with a melting point about 700° C.

If a fresh sample of radium is enclosed in a hermetically sealed glass tube such that the α-particles given out by the radium can penetrate the tube and the rate at which α-particles are emitted is determined, it is found that the rate increases with time, finally approaching a value which is four times the rate which was found when the radium was initially placed in the tube. Also initially it is found that the radium emits only α-rays and it is not until some time later that the emission of β- and γ-rays appears.

On the other hand if a speck of radium is placed in a constant current of air, it is found that the α-ray activity, as the rate of emission of α-particles may be called, remains at the initial value. Furthermore, if the radium is placed in a long tube and a slow stream of air is passed along the tube, it is found that the air in parts of the tube remote from the radium appears to be radioactive since it emits α-particles. However, this apparent radioactivity of the air can be removed by passing the air through a U-tube which is immersed in liquid air. It thus appears that a gas is given off by the radium and that this gas is condensed out of the air at a low temperature. Further, it appears that this gas is radioactive since, on allowing the U-tube to come to room temperature, it is found that the air in the U-tube shows α-ray activity. This gas which is given off by radium is variously known as radium emanation, radon and niton. A sample of radium emits this gas at a constant rate. The optical spectrum of the emanation has been found and it is quite distinct from the spectrum of radium and also from the spectra of all other elements. The emanation is thus a definite chemical element. Since it is a gas its molecular weight can be determined and it is found to be about 222. It is chemically inert and thus resembles the inert gases helium, neon and argon. It is therefore placed in the same column in the periodic table as these gases. The atomic number of radon is 86.

If a supply of radon is drawn off from a sample of radium and enclosed in a sealed thin-walled glass tube, it is found that the α-ray activity of the gas in the tube increases and then begins to decrease. After a few hours the α-ray activity decreases according to a very definite law as represented by

$$I = I_o e^{-\lambda t} \tag{19–12}$$

where I_0 is the α-ray activity after a few hours, I is the activity as measured at a time t after the measurement of I_0, and λ is a constant known as the radioactive constant. This constant must not be confused with wave-length which is also represented by the symbol λ. The value of λ in the above case is 0.18 $(\text{day})^{-1}$. This means that the α-ray activity of radon drops to half value (i.e., $I = I_0/2$) in a time given by $\lambda t = 0.693$. This time is

$$= \frac{0.693}{0.18} = 3.85 \text{ days}$$

As in the case of a fresh sample of radium, it is found that a fresh sample of radon initially shows no β-ray activity, but, if the emanation is sealed in a tube, the tube will after a while show β- and γ-ray activity. If, after the radon has been in the glass tube for several hours, it is blown out of the tube, it is found that the tube shows α-ray activity. It seems in this case that something else is produced while radon emits α-particles. It was found that this something else is deposited upon surfaces and especially on negatively charged metal surfaces. This something is known as the active deposit.

As we have noted in Sec. 19.1, Rutherford and Soddy were led in 1902–03 to put forward the transformation theory of radioactivity. According to this theory an atom of a radioactive element is transformed into an atom of a different element when it emits an α- or a β-particle. According to Sec. 14.18, the concept of atomic number began to come into use in 1912. Previous to 1912, the Rutherford and Soddy transformation was vague as to the relation between a parent and a daughter atom in a radioactive process. In 1912 Rutherford and Soddy removed this vagueness of their theory. If the atom emits an α-particle, its atomic weight is reduced by four (the atomic weight of helium) and its atomic number by two. If the atom emits a β-particle, its atomic weight remains practically constant since the mass of an electron is so much smaller than that of a hydrogen atom and its atomic number is increased by one. Thus the chemical properties of the transformed atom are quite different from those of the original atom. It may be that the transformed atom is also radioactive, in which case another transformation may occur. The theory of successive transformations explains the observed facts concerning radium. A radioactive series is shown diagrammatically in Table 19.1. Thus, when an atom of radium (Ra) emits an α-particle, the remaining atom becomes an atom of radon (Rn). The atom of radon in turn emits an α-particle, becoming an atom of RaA and so on, as shown in Table 19.1.

If radium is kept isolated from its transformation products as by placing it in a stream of air so that the radon is blown away, it is found that only α-rays are emitted. The transformation from Ra to Rn therefore is accompanied by the emission of an α-particle. As shown in Table 19.1, the transformation of RaB to RaC is accompanied by the emission of a β-particle. According to Rutherford and Soddy the α-ray activity of a radioactive element which emits α-rays is proportional to the number of atoms of the element present. As the emission of each α-particle accompanies the transformation of an atom of the element, the rate at which atoms of the element are being transformed into atoms of a different element is proportional to the number of atoms of the element present. If n is the total number of atoms present at a time t and a number dn of the atoms are transformed in an interval dt after the time t, then the number of atoms of the original element at a time $t + dt$ is $n - dn$. The rate of loss of the atoms is thus dn/dt, and this rate of loss is proportional to n, so that

TABLE 19.1

Element	Ra $\overset{\alpha}{\nearrow}\!\rightarrow$	Rn $\overset{\alpha}{\nearrow}\!\rightarrow$	RaA $\overset{\alpha}{\nearrow}\!\rightarrow$	RaB $\overset{\beta}{\nearrow}\!\rightarrow$	RaC $\overset{\beta}{\nearrow}\!\rightarrow$	RaC′ $\overset{\alpha}{\nearrow}\!\rightarrow$
Atomic weight..	226	222	218	214	214	214
Atomic number.	88	86	84	82	83	84
Half life........	1600 yrs	3.83 days	3 min	26.8 min	19.7 min	10^{-6} sec

$$\frac{dn}{dt} = -\lambda n \tag{19–13}$$

where λ is a constant and the negative sign appears because there is a loss. Rearranging (19–13), we have

$$\frac{dn}{n} = -\lambda dt$$

Then, integrating both sides, we obtain

$$\log n - \log n_0 = -\lambda t$$

or

$$n = n_o e^{-\lambda t} \tag{19-14}$$

where n_o is the number of atoms present at $t = 0$. The constant λ is called the radioactive or transformation constant. The relation of λ to the half-value period T is

$$T = \frac{0.693}{\lambda} \tag{19-15}$$

In 1946 it became the custom to speak of the half life of a radioactive element rather than its half-value period The half lives of various radioactive elements are shown in Table 19.1. Thus, a gram of radium slowly disintegrates so that in 1600 years there is only half a gram of radium left. Radon disintegrates more rapidly, only half of it remaining after 3.83 days. RaA has a half life of 3 minutes.

The heating effect which we mentioned in Sec. 19.2 is due to the absorption of the α-, β-, and γ-rays in the walls of the tube and in the water of the calorimeter. The value of 130 calories per hour is for 1 gm of radium in equilibrium with its transformation products. The heating effect of 1 gm of radium isolated from its transformation products is about 25 calories per hour.

19.9. Radioactive Elements and the Periodic Table.—In Chapter XVII we have described the periodic table of the chemical elements. The sixth and seventh rows of this table contain radium and its transformation products. Rows 6 and 7 of the periodic table are shown in more detail in Table 19.2. It is seen that space Number 84 is occupied by RaC′ and RaA—the atomic weight of RaC′ being 214 and that of RaA being 218. Each of course has the same atomic number since they are both in the same space of the periodic table. RaC′ and RaA have

TABLE 19.2

6			Pb 207, 82 RaB 214, 82	RaC 214, 83	RaC′ 214, 84 RaA 218, 84	85	Rn 222, 86
7	87	Ra 226, 88					

the same chemical properties and are therefore the same element. However, atoms of RaA are heavier than atoms of Ra′C. It appears then that the atoms of element Number 84 do not all have the same weight. The two substances RaC′ and RaA are known as *isotopes.* It is seen also that space Number 82 is occupied by lead (Pb) with an atomic weight of 207 and RaB of atomic weight 214. Substances *A*, *B*, *C* . . . are isotopes of an element when they have the same chemical properties as the element but differ from each other in their respective atomic weights. All atoms of a given isotope have equal weights. An element as it occurs in nature is usually a mixture of isotopes. The proportion in which the isotopes are mixed determines the atomic weight as measured by chemical means. Thus, lead obtained from different localities is found to have slightly different atomic weights, as determined by chemical means. These atomic weights vary from 206.08 to 207.69.

A series of transformation products as shown in Table 19.1 is called a radioactive series. The particular series shown is known as the radium series. The end of this series is RaG with atomic weight 206 and atomic number 82. Lead is thus the end product of the radium series. Radium itself is a transformation product, the original parent element in the radium series being uranium with atomic weight 238 and atomic number 92.

19.10. International Atomic Weights.—In Sec. 7.8 we have mentioned that originally the atomic weight of hydrogen was taken as unity. However, it was found in the course of time that, if the atomic weights of the elements are expressed in terms of $O = 16$ instead of $H = 1$, the atomic weights of many elements are whole numbers or very nearly whole numbers. It is nowadays customary to use a system of atomic weights based upon $O = 16$ rather than a system based upon $H = 1$. The atomic weights based upon $O = 16$ are known as the international atomic weights. These are the weights given in tables. In such tables hydrogen has an atomic weight of 1.0080. However, even in the international system of atomic weights, chlorine has an atomic weight of 35.45, which is distinctly not a whole number, being practically half way between 35 and 36. It has recently been found that there are isotopes of chlorine having the atomic weights 35 and 37. The chemical atomic weight of 35.45 is therefore an average of the weights 35 and 37, mixed in the proportion of 3.44 atoms of Cl^{35} to one atom of Cl^{37}.

19.11. Positive Rays and Isotopes.—In Chapter VII we have described the electric discharge through gases at low pressures. At a pressure of about 1 mm of mercury, the Crookes dark space appears in front of the cathode. It is found that, if the cathode is perforated, streams of luminosity appear at the back of the cathode, each stream proceeding from a perforation in the cathode. If these streams of luminosity are exposed to electric and magnetic fields they are found to consist of positive charges and so these streams are called positive rays.

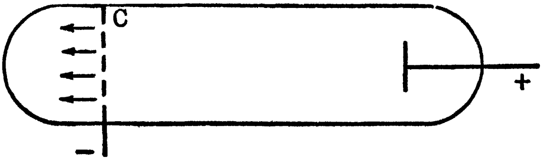


FIG. 19.5.—Positive ray tube.

The perforated cathode is represented at *C* in Fig. 19.5. The particles travel in the directions of the arrows. The electrostatic and magnetic deflections of the streams show that the ratio of the charge to the mass equals that for the same ions in electrolysis. For the accurate determination of the ratio of the charge to the mass of the positive particles, the cathode is made of thick metal with a long but very narrow hole bored in the cathode. The positive rays emerging from the hole pass into a highly evacuated space where they can be acted upon by strong electrostatic and magnetic fields. As described in Chapter VII, the e/m for these positively charged particles can be calculated from the electrostatic and magnetic deflections of the rays.

Applying this method, J. J. Thomson found the masses of a large number of particles. The particles depend upon the nature of the gas in the tube. In 1912 J. J. Thomson found neon gas to give two different values of e/m. If e is the charge on the hydrogen ion in electrolysis or 1.6×10^{-20} emu, the value of m can be found. The two values of e/m for neon gave values of 20 and 22 for m in atomic weight units. (An atomic weight unit has an actual mass of 1.66×10^{-24} gm).

In 1918 Dempster of Chicago and in 1919 Aston of Cambridge, England resumed the investigation of these charged particles. An apparatus for measuring the mass of these charged particles is called a mass spectrograph. An improved type of mass spectrograph was built by Bainbridge of Harvard in 1933. This spectrograph is shown diagrammatically in Fig. 19.6. The positive ions (charged particles) are emitted by the source S. They are then accelerated by a potential difference V between S and the hole in the plate P. On passing through a hole in P the particles have a velocity given by

$$(\tfrac{1}{2})mv^2 = Ve$$

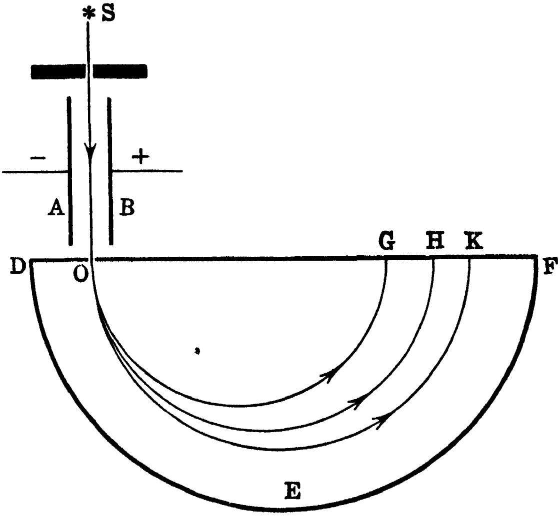


FIG. 19.6.—Bainbridge mass spectrograph.

where e is the charge on the particle and m its mass. The stream of particles now passes between the plates A and B of the "velocity selector." These plates are maintained at a potential difference V_1 and at the same time a magnetic field of strength H_1 is applied to the space between the plates A and B. The magnetic field is perpendicular to the plane of the diagram in Fig. 19.6. Only those particles enter the semi-circular chamber DEF whose path through the velocity selector is a straight line. The path of the positive ions as shown in Fig. 19.6 is in an evacuated space.

The force due to the electrostatic field on a positively charged particle between A and B is toward A and of magnitude V_1e/d, where d is the distance apart of the plates A and B. Also if the magnetic field is from the reader into the plane of the diagram the force due to this field is towards B and of magnitude H_1ev. The path of the particles through the selector is a straight line when these two forces are equal—that is, when

$$V_1e/d = H_1ev$$

or

$$v = V_1/H_1d \tag{19–16}$$

Although only particles with this velocity get through to the semicircular chamber, yet particles of different masses can get through. So particles having different masses but the same velocity enter the semicircular chamber DEF. The magnetic field H_1 is applied to this space also and causes the paths of the particles to be bent into circles as shown. Equating the magnetic force H_1ev to the centrifugal force, we have

$$mv^2/r = H_1ev \tag{19–17}$$

Eliminating v from (19–16) and (19–17), we obtain

$$m = \frac{redH_1^2}{V_1} \tag{19–18}$$

It is seen that for a given charge e the mass of the ion is proportional to r the radius of the circle. A photographic film is placed at GK inside the semicircular chamber. On development of the film lines appear at G, H, K, Fig. 19.6. The diameters of the circular paths giving rise to these lines are OG, OH and OK respectively. Hence the masses of the corresponding particles can be found. The lines G, H, K constitute a mass spectrum. The mass spectrum of the element germanium is shown in Fig. 19.7. The numbers refer to the atomic weights of the isotopes of germanium. The width of the lines gives a measure of the relative abundance of the isotopes. From this relative abundance an average atomic weight is found and this agrees with the chemical atomic weight of germanium.

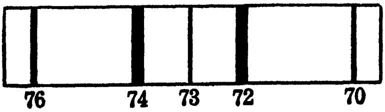


FIG. 19.7.—Mass spectrum of germanium.

The source of positive ions shown as S in Fig. 19.6 is usually a pinhole leading into a side tube. For ions of a gas the hole is in the cathode of a discharge tube containing the gas under investigation; for ions of a solid, a salt of the element is bombarded with a strong current of electrons in the side tube and positive ions produced. Positive ions pass through the hole into the mass spectrograph.

19.12. Mass Number.—On the basis of 16 for the atomic weight of oxygen, Aston at first concluded that the masses of all isotopes with the exception of hydrogen could be represented by whole numbers. This is known as Aston's whole number rule. More recent and more accurate work has shown that this is not quite an exact statement of fact. However, the actual masses differ so little from being whole numbers that it is convenient to distinguish between isotopes of the same element by using the nearest whole numbers. This is done in Fig. 19.7. For instance, we distinguish the isotopes Ge^{72} and Ge^{74}. The numbers 72 and 74 are known as mass numbers, although on the basis of $O = 16$ the masses of Ge^{72} and Ge^{74} differ very slightly from 72 and 74.

19.13. Heavy Hydrogen or Deuterium.—The discovery of the isotope H^2 of hydrogen was the result of very exact and careful measurement. At first it was believed that there was only one isotope of oxygen—namely, O^{16}. This seemed very fortunate since the chemical atomic weight of oxygen had been taken as 16. A mass spectrum of hydrogen obtained in the late 1920's showed that the exact mass of H^1 was 1.0077. Since the accepted value of the chemical atomic weight of H^1 is also 1.0077, it was concluded that there was only one isotope of hydrogen. However, in 1929, it was found that oxygen was a mixture of the isotope O^{16} with slight traces of O^{17} and O^{18}. Hence, if we take $O^{16} = 16$, the chemical atomic weight of oxygen will be slightly greater than 16. On this new basis the chemical atomic weight of hydrogen must be somewhat greater than 1.0077 and this means that hydrogen must carry a trace of an isotope heavier than H^1. In 1931 Birge and Menzel of the University of California estimated the relative abundance of isotope H^2 to H^1 as one part in 4500. In 1931 it was not possible to detect an isotope of such small relative abundance, and some method of concentrating the heavier isotope was needed. From theory it was expected that liquid H^2 would have a somewhat higher boiling point than liquid H^1 and hence it should be possible to cause some separation from a mixture of the two isotopes by the method of fractional distillation used so much in chemistry. If chemical liquid hydrogen is boiled the residue is richer in the heavier isotope. In 1932 Urey, of Columbia University, New York, subjected such a residue to spectral analysis. The Balmer series lines of H^1 were obtained as usual, but each line was accompanied by a faint line on its short wave-length side. From Sec. 16.11 the Rydberg constant for an atom with a nucleus of mass M is related to that for an atom with a nucleus of infinite mass by

$$R_M = R_\infty M/(M + m) \tag{19–19}$$

where m is the mass of the electron. Hence

$$\frac{R_2}{R_1} = \frac{M_2(M_1 + m)}{M_1(M_2 + m)} \tag{19–20}$$

where the subscripts 1 and 2 refer to H^1 and H^2 respectively. If λ_1 is the wave-length of a certain line in the spectrum of H^1 and λ_2 that of the corresponding line in the spectrum of H^2, it is seen from Sec. 16.11 that $\lambda_1/\lambda_2 = R_2/R_1$. Replacing the left side of (19–20) by λ_1/λ_2 and then solving for M_2, we obtain

$$M_2 = \frac{M_1 m \lambda_1}{m\lambda_1 - (M_1 + m)\,(\lambda_1 - \lambda_2)} \tag{19–21}$$

From the measured difference $(\lambda_1 - \lambda_2)$ between a Balmer line and its faint companion, Urey found that the faint lines came from atoms with nuclei of mass 2 in atomic weight units. This discovery was announced in 1932.

The isotope H^2 is of such great interest and importance that at first it was given the special name deuterium and the special chemical symbol D. The more recent practice is to call it hydrogen-2 with the symbol H^2. The symbol H is used both for H^1 and for chemical hydrogen, which consists of a mixture of H^1 and H^2. The boiling point of liquid H^1 is 20.4° Abs, while that of liquid H^2 is 23.5° Abs.

If ordinary water containing an electrolyte such as sodium hydroxide (NaOH) is electrolyzed until only a small portion of the original liquid remains, the proportion of molecules containing deuterium atoms is quite large. In one case over 600 gallons of water containing an electrolyte were electrolyzed over a period of four months. At the end of the period 80 cm^3 of solution remained. Ninety per cent of all the hydrogen atoms in the residue were found to be of the heavy variety.

It is convenient to use the terms physical atomic weight and chemical

FIG. 19.8.—Harold Clayton Urey.

atomic weight as applied to an atom. The physical atomic weight of an atom is referred to the oxygen isotope $O^{16} = 16$, while the chemical atomic weight of an atom is referred to chemical oxygen $O = 16$, where O represents the mixture of O^{16}, O^{17} and O^{18} found in nature. By means of the mass spectrograph the physical atomic weight of H^2 atoms has been found to be 2.01472. It is customary to call the nucleus of $_1H^2$ a *deuteron.*

CHAPTER XIX

Problems

1. A β-particle is moving with a speed of 0.95 times that of light. Find the energy of the β-particle as expressed in mev.

2. When β-particles are suddenly stopped, γ-rays are produced in a manner similar to the production of x-rays by the sudden stopping of cathode particles in an x-ray tube. β-particles of velocity 0.90 that of light fall upon a target of heavy metal and produce secondary γ-rays. Determine the shortest wave-length of these γ-rays, using the Duane-Hunt relation.

3. Find the time it will take a gram of radium to disintegrate to 0.2 gm of radium.

4. We begin with a given quantity of RaA. Find the time for the quantity to diminish to one-tenth due to radioactive disintegration.

5. What is the range of 1.5 mev α-particles in air at a pressure of 50 cm of mercury and a temperature of 20° C?

6. In the Bainbridge velocity selector, the distance apart of the plate is 0.5 cm, the potential between the plates is 600 volts, and the strength of the magnetic field is 5000 gauss. (*a*) Find the velocity of the charged particles which pass through the selector. (*b*) Determine the atomic weight of these particles when the radius of the path in the Bainbridge spectroscope is 8.5 cm, the magnetic field being 5000 gauss. Assume that the particles carry a charge $+e$. (*c*) Determine the radius when the particles have a physical atomic weight of 16.

7. Given that the wave-length of the α line of the Balmer series for H^1 is 6564.62 A, find the wave-length of the α line of the Balmer series for H^2.

CHAPTER XX

NUCLEAR PHYSICS

20.1. The Nuclear Atom.—The general view held by physicists during the middle of the first decade of the twentieth century was that the mass of an atom was comprised mostly of the electrons in the atom. For instance, since the mass of the helium atom is 7360 times that of an electron, it was concluded that the helium atom contained something like 7360 electrons. But these electrons would give the atom a tremendous negative charge, so that it was necessary to compensate this charge by having the electrons distributed in a sphere of positive electricity whose total charge was equal to the total negative charge of the electrons. However, in 1904 Nagaoka of Tokyo, Japan, suggested another view of the atom. Maxwell had explained the rings of the planet Saturn by supposing that each ring consists of many small objects, perhaps stones, revolving about the planet. There are gravitational forces of attraction between any two objects in a ring and between each object and the planet. Bringing into play the centrifugal force due to the revolution of the objects in the ring about the planet, Maxwell was able to show that the ring was stable. Nagaoka applied Maxwell's theory of Saturn's rings to a ring of electrons revolving about a central positively charged nucleus. In this case the electrons repel each other but are attracted to the central nucleus with a positive charge. Nagaoka's suggestion indicates that the minds of physicists in 1904 were beginning to turn away from the many electron atom. As we have seen in Secs. 14.8 and 14.12, Barkla's work on the scattering of x-rays by light elements together with Thomson's theory of the scattering of x-rays indicated toward the end of the first decade of the present century that the number of electrons in an atom was roughly equal to half the atomic weight, except in the case of hydrogen where the number of electrons equals the atomic weight. The number of electrons in the atom of helium was thus $4/2 = 2$ and not 7360. Evidence against the many electron atom became overwhelming. But, with the disappearance of the many electron atom, how was the mass of the atom to be explained? Rutherford turned to the idea that the mass was almost wholly concentrated in the nucleus.

20.2. Scattering of α-Rays.—In 1911 Rutherford, then of Manchester, England, tested the nuclear theory of the atom by scattering α-particles from thin gold leaf. We have seen in Sec. 19.4 that α-rays have a range in air and that they usually follow a straight path in air. However, although most α-particles follow a straight path, a few of them suffer sudden bends, so that, if a beam of α-rays is passed through thin gold leaf and the scintillations produced by the beam on a fluorescent screen are observed, it is found that, although most of the scintillations occur at a central spot on the fluorescent screen, occasional scintillations occur at varying distances from the central spot. The radioactive source is at *R*, Fig. 20.1. A diaphragm is placed at *S*. The α-rays pass through the leaf *L* and arrive at the fluorescent screen *F*. Most of the α-particles hit at *O*, but a few hit at *P*. The arrangement in Fig. 20.1 may be enclosed in an evacuated glass vessel. Certain particles are scattered as represented by *AP*, Fig. 20.1. Usually when an α-particle ploughs through an atom, it knocks electrons out of the Bohr orbits of the atom. Occasionally, however, the α-particle approaches close to the nucleus and perhaps a collision takes place. In this case the mass of the α-particle is less than that of the nucleus, so that in any collision or close collision the α-particle is knocked out of its path.

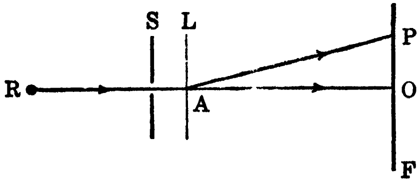


FIG. 20.1.—Experimental arrangement for showing the scattering of α-particles.

On the assumption that both the nucleus and the α-particle have positive charges and that the inverse square law of electrostatics is valid down to such short distances, Rutherford showed that, in the case of a massive nucleus such as that of gold (At. Wt. 197), the α-particles followed paths as shown in Fig. 20.2. The closer the approach of the α-particle to the nucleus the greater the scattering angle θ. Each of the above paths is a hyperbola with the heavy nucleus at a focus. Close approaches to the nucleus are very rare so that the number of scintillations occurring per unit time at a small area at *P*, Fig. 20.1, decreases very rapidly, as the distance OP increases. From the way in which the scintillation rate for a small area varies with the distance of that small area from the central spot O, Rutherford showed that the inverse square law of force between the charged α-particle and the charged nucleus holds down to distances as small as 3.2×10^{-12} cm for gold nuclei. Hence, the radius of a gold nucleus is not greater than 3.2×10^{-12} cm. Using copper and silver foils, the upper limits for the radii of copper and silver nuclei are 1.2×10^{-12} and 2×10^{-12} cm respectively.

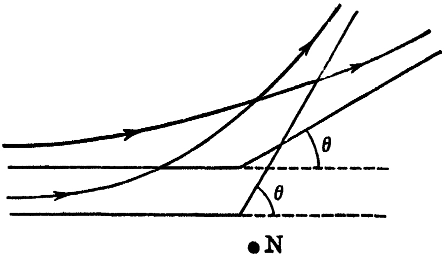


FIG. 20.2.—Showing the scattering of α-particles by a nucleus N—the closer the approach of the particle to the nucleus the greater the scattering angle θ.

In 1912 the concept of atomic number and the results of Rutherford's experiments on the scattering of α-particles and Barkla's experiments on the scattering of x-rays began to influence physicists toward the ideas that the mass of the nucleus of an atom equals the atomic weight of the atom on the basis of O = 16 and that this nucleus is surrounded by a number of electrons equal to the atomic number. The time was ripe for the Bohr theory which we have described in Chapter XVI. This theory was announced in 1913.

✓**20.3. Wilson Cloud Tracks.**—If a gas is suddenly expanded it cools. In the case of ordinary air which contains dust particles and which is saturated with moisture, a sudden expansion causes some of the moisture to condense on to the dust particles. These dust particles form the nuclei of minute rain drops, and together these rain drops make a cloud. An expansion of any size produces a cloud in dusty air saturated with moisture. If the dust particles are removed from the air, a small expansion of air saturated with moisture will not produce a cloud. If, however, the dust free air contains positive and negative ions and if the expansion, as measured by the ratio of the volumes of the air before and after the expansion, is above a certain critical value, small droplets are formed on the ions. We have seen that as α-rays plough their way through air they leave behind them a trail of positive and negative ions. If, immediately after the production of this trail in saturated air, the air is expanded beyond the critical value, droplets form upon the ions in the trail and form what is known as an α-ray track. This has the appearance of a thin white thread, which can be seen by the naked eye. The droplets in the track disperse in a second or so and the track disappears. In order to obtain a permanent record of the tracks, Wilson arranged to take a photographic snapshot of the track immediately after its production. The α-ray tracks are in general found to be straight. Occasionally, however, there is a bend in the track corresponding to a close approach of the α-particle to the nucleus of an atom of one of the air molecules. An example of α-ray tracks is shown in Fig. 20.3. Tracks are also made by β-rays and by the secondary cathode rays produced by x-rays. These tracks are not straight like the α-ray tracks, but are irregular owing to the fact that, when an impinging electron knocks another electron out of an atom, the impinging electron is usually knocked out of its own path.

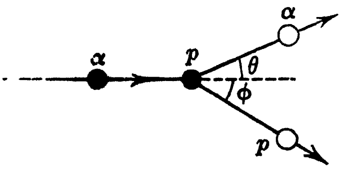


FIG. 20.3.—Collision of nuclear particles.

20.4. Bombardment with α-Particles.—The range of α-particles can be studied by moving a fluorescent screen away from the thin layer of radioactive material and counting the scintillations. There is a sudden drop in the count at the range. In 1914 Marsden, at Manchester, England, did this experiment with the α-rays from radium C passing through hydrogen. He found a sudden drop in the count at the range, but beyond the range there were very definite scintillations. These long-range scintillations are caused by H-particles produced by the collision of α-particles with hydrogen atoms. These H-particles are now known as protons.

In Fig. 20.3 the black circles represent particles before collision and the white circles particles after collision. The kinetic energy of the impinging α-particle is known. After the impact the kinetic energies of the recoiling proton and the deflected α-particles can be determined from the lengths of the respective fork tracks. It is found in this case that the total kinetic energies before and after impact are equal, so that the impact is elastic. Sometimes the α-particle makes a head-on collision with the hydrogen nucleus and the proton recoils in the direction of the original α-ray track. In this case $\phi = \theta = 0$ in Fig. 20.3. Conservation of kinetic energy gives

$$(1/2)m_\alpha v_{\alpha1}^2 = (1/2)m_\alpha v_{\alpha2}^2 + (1/2)m_p v_{p2}^2 \tag{20-1}$$

and conservation of momentum gives

$$m_\alpha v_{\alpha1} = m_\alpha v_{\alpha2} + m_p v_{p2} \tag{20-2}$$

The subscripts α and p refer to the α-particle and proton, respectively, and the subscripts 1 and 2 to the respective quantity before and after collision. The kinetic energy of the recoiling proton is

$$W_{p2} = (1/2)m_p v_{p2}^2 \tag{20-3}$$

and that of the impinging α-particle is

$$W_{\alpha1} = (1/2)m_\alpha v_{\alpha1}^2 \tag{20-4}$$

Eliminating $v_{\alpha1}$, $v_{\alpha2}$ and v_{p2} from the above four equations, we obtain

$$W_{p2} = \frac{4m_p m_\alpha}{(m_\alpha + m_p)^2} \cdot W_{\alpha1} \tag{20-5}$$

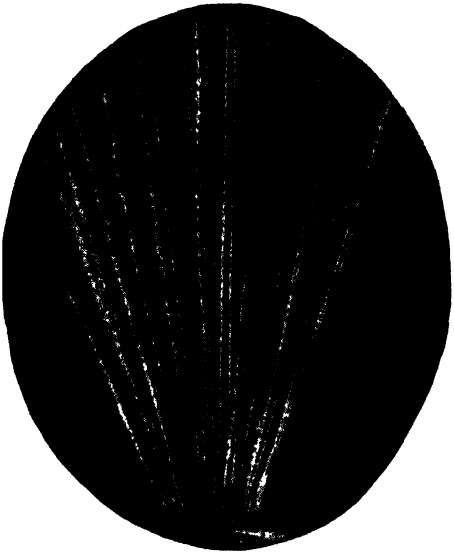
Since $m_\alpha = 4m_p$, the energy of the recoiling proton in a head-on collision is 0.64 times the kinetic energy of the α-particle before the collision. The length of the proton track shows this to be the case.

In 1919 Rutherford, using Marsden's method, looked for long-range scintillations when air, carbon dioxide, and nitrogen were substituted for the hydrogen. He found long-range scintillations when air or nitrogen, but not carbon dioxide, was used. Finally, these scintillations were traced to the presence of nitrogen in the gas. To Rutherford's surprise the long-range scintillations had the appearance of H-particle scintillations, which are not so bright as α-particle scintillations. After taking care to exclude any hydrogen from his apparatus these scintillations persisted so long as nitrogen was present. Rutherford made the bold suggestion that transmutation of one element into another had occurred. According to our present views, a process described by the nuclear equation

$${}_7N^{14} + {}_2He^4 \rightarrow {}_1H^1 + {}_8O^{17} \tag{20-6}$$

had occurred. In this equation, the subscript to the left of the chemical symbol represents the atomic number and the superscript to the right represents the mass number. In a nuclear equation the mass and atomic numbers must separately balance. Thus, the sum of the mass numbers on each side of (20–6) is 18 and that of the atomic numbers is 9.

Visual evidence for the process described by (20–6) is given by the Wilson cloud-track picture shown in Fig. 20.4. Very rarely a fork appears in an α-ray track as shown in Fig. 20.4. In this case it seems that the α-particle approached a nucleus of nitrogen so closely that a collision took place. We must distinguish two kinds of collision—elastic and inelastic. In an elastic collision, the nitrogen nucleus recoils, just as when one billiard ball is struck by another. An elastic collision is one in which the total kinetic energy of the particles before collision equals that of the particles after collision. In an inelastic collision this relation no longer holds.



(*Courtesy Harkins and Schuh and Physical Review*)

FIG. 20.4.—α-ray tracks in nitrogen. The long leg of the forked track was made by the ejected proton and the short leg by an oxygen atom of atomic weight 17. Probability of forked track as shown is 8 in 1,000,000 α-ray tracks.

It is seen from Fig. 20.4 that the longer leg of the fork is thinner than the α-ray track below the fork. This is characteristic of proton tracks. The thinness of these tracks as compared with the α-ray tracks is evidence for the view that protons ionize the gas through which they pass less heavily than do α-particles. It has been found that deuterons, which are the nuclei of ${}_1H^2$ atoms, can also produce cloud tracks. These are slightly thicker than proton tracks and distinctly thinner than α-ray tracks. In (19-11) a formula for the range of an α-ray track was given in terms of the energy of the α-particle at the beginning of the track. Similar formulas have been found for the ranges of protons and deuterons. In all three cases the relation is of the form

$$W(\text{in mev}) = k(Rp/T)^{2/3} \tag{20–7}$$

where R is the range in cm of air, p the pressure of the air in cm of mercury, T the absolute temperature, and k is a constant for each kind of particle. The values of the constant k are shown in Table 20.1.

TABLE 20.1

VALUES OF *k*

Protons	1.24
Deuterons	1.57
α-particles............	4.90

Formula (20–7) together with the values of k given in Table 20.1 holds fairly well up to a range of 25 cm.

We now return to Fig. 20.4. The longer leg of the fork is a proton track. In the original photograph the shorter leg of the fork is thicker than an α-ray track, implying that it is produced by a nucleus heavier than an α-particle. A study of the lengths and directions of the fork indicates that the collision is inelastic—the total kinetic energy of the particles after the collision being about 0.5 mev less than that of the α-particle before the collision. It is therefore believed that the α-particle is absorbed into the hydrogen nucleus from which the proton is expelled and a new nucleus—that of O^{17}—is produced.

20.5. Isotopic Masses.—The following table gives the stable isotopes up to atomic number 14 whose exact masses have been measured with the mass spectrograph. The values given in the fourth column of Table 20.1 are in *mass units*. The abbreviation for these units is "mu." Although the electron (e) and the neutron (n) are not strictly "elements," their masses on the basis of $O^{16} = 16$ are included in the table.

TABLE 20.2

Element	Atomic Number	Mass Number	Exact Mass $O^{16} = 16$
e	−1	0	0.00055
H	1	1	1.00812
n	0	1	1.00893
H	1	2	2.01472
He	2	3	3.01701
	2	4	4.00388
Li	3	6	6.01690
	3	7	7.01804
Be	4	8	8.00777
	4	9	9.01497
B	5	10	10.01605
	5	11	11.01286
C	6	12	12.00398
	6	13	13.00766
N	7	14	14.00750
	7	15	15.00489
O	8	16	16.00000
	8	17	17.00450
	8	18	18.00470

Note: The masses in this table are for the neutral atoms with their orbital electrons.

20.6. Lawrence's Cyclotron.—Until 1932 nuclear disintegration had only been produced by α-particles having a kinetic energy greater than 3 mev. About this time theory indicated that high-speed protons should be very efficient in producing nuclear disintegration. It was at first thought that the only way to obtain 1 mev protons was to have one million volts available. Van der Graaff therefore constructed a large electrostatic machine which was housed in an airship hangar both because of its size and because of the huge voltages developed. However, in 1932 Lawrence, of the University of California, discovered that by a certain device a stream of protons could be speeded up to have an energy of 1.2 mev per proton even though only 10,000 volts were available. This device is known as the cyclotron.

The cyclotron essentially consists of a hollow cylindrical metal box which is cut in two as shown in Fig. 20.5a. An alternating voltage of frequency f is connected to the two halves. There is a source of protons at S, Fig. 20.5b. The whole apparatus shown is in a vacuum. Let V be the maximum potential difference between the half-boxes. Then $(\frac{1}{2})mv_1^2 = Ve$, where e and m are the charge and mass of the proton and v_1 is the velocity given to it when it falls through a potential V. The proton passes into the inside of the half-box on the left of Fig. 20.5b and while in this box it is under the action of zero electric field and so travels with a velocity v_1. However, a steady magnetic field of strength H is applied in a direction perpendicular to the plane of 20.5b. This bends the path of the proton into a circle of radius r_1. Equating centrifugal force to the force acting on the moving proton due to the magnetic field, we have

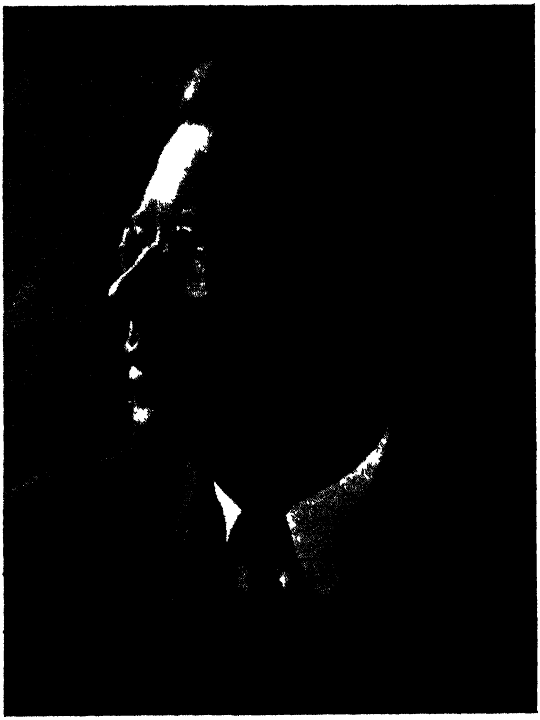


FIG. 20.5A.—Ernest Orlando Lawrence.

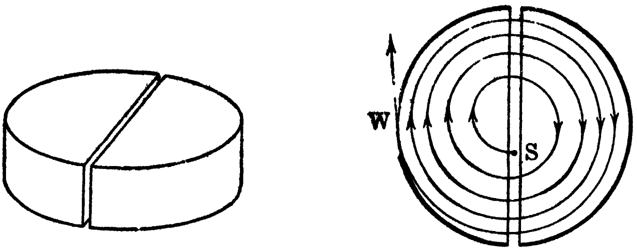
$$mv_1^2/r_1 = Hev_1$$

or

$$v_1 = r_1He/m \tag{20-8}$$

The time taken to traverse the semicircle of radius r_1 is $\pi r_1/v_1$ and this from (20–8) is $\pi m/eH$. Having traversed the semicircle in the left half-box, the proton jumps across to the right box. If now at the time of this jump the potential between the two half-boxes is reversed to its negative maximum the kinetic energy is increased by Ve, so that

$$\begin{aligned}(\tfrac{1}{2})mv_2^2 &= Ve + (\tfrac{1}{2})mv_1^2 \\ &= 2Ve\end{aligned} \tag{20–9}$$



(a) (b)

Fig. 20.5.—Lawrence's cyclotron.

The proton now travels in the right half-box with a velocity v_2 and in a semicircle of radius r_2. However, the time spent in the right half-box is $\pi m/eH$ which is the same as that spent in the left half-box. If the voltage is again reversed when the proton jumps from the left to the right half-box there is another increase of energy, Ve. In the sth semicircle the radius is

$$r_s = (1/H)\sqrt{2smV/e} \tag{20–10}$$

The radii of the semicircles increases as $\sqrt{s}$. If half the period of oscillation of the alternating voltage is $\pi m/eH$, the protons continue to build up speed. Finally one of the semicircles becomes so large that the proton hits the thin window, *W*, Fig. 20.5b, and passes through. This proton may now be used for disintegration experiments. The success of the cyclotron depends upon the fortunate fact that the time spent in each semicircle is the same irrespective of the speed with which the proton is traveling. It is only necessary to keep the frequency f of the alternating voltage and strength H of the magnetic field so that

$$f = He/2\pi m \tag{20–11}$$

This formula holds only as long as m is a constant—that is, as long as the speed of the flying particle is much less than that of light.

Other particles, such as deuterons (the nuclei of ${}_1H^2$) and α-particles, can be accelerated in a cyclotron, e and m in (20–10) and (20–11) then referring to these other particles. The energy of the particle in the sth semicircle is

$$W = sVe \tag{20–12}$$

Eliminating s and H from the above three equations, we have

$$W \doteq 2\pi^2 m f^2 r_s^2 \tag{20–13}$$

For a given cyclotron, the radius r_s of the largest semicircle and the frequency f are fixed. Hence W varies as the mass m of the charged particles whirling in the cyclotron. A cyclotron which delivers 8 mev protons will deliver 16 mev deuterons and 32 mev α-particles.

The semicircular boxes of Fig. 20.5 are now known as *dees* because of their resemblance to the letter "D." It is customary to describe a cyclotron by the diameter of its pole pieces. Lawrence's first cyclotron had 4-inch pole pieces. Then came a cyclotron that had 11-inch pole pieces and delivered 1.2 mev protons. This was followed by cyclotrons with 27.5-, 37-, and 60-inch pole pieces. The last weighed 200 tons, took about 200 kilowatts of power, and delivered 16 mev deuterons (8 mev protons) in a stream of about 500 μa (1 microampere = 10^{-6} amp). This was in 1942. In 1946 Lawrence was using a 184-inch cyclotron.

The Washington University cyclotron has pole pieces 45 in. in diameter, weighs 100 tons, takes 150 kw of power, and delivers 350 μa of 12 mev deuterons. It should be noted that the diameter of the dees is somewhat less than that of the pole pieces. In the Washington University cyclotron, the radius of the largest semicircle within the dees is 19.25 in. and the frequency of operation is 10.6 megacycles. Eq. (20–11) may be written in the form

$$W(\text{in mev}) = 128\, \mu f_m^2 r_i^2 \tag{20–14}$$

where μ is the mass number of the particles emerging from the cyclotron, f_m is the frequency in megacycles, and r_i is the radius of the largest semicircle within the dees in inches. Putting in values for the Washington University cyclotron, we obtain 11.6 mev for deuterons. Formula (20–14) holds as long as the speed of the flying particles is much less than that of light.

20.7. Bombardment with Protons and Deuterons.—Rutherford's experiments on the scattering of α-particles showed that the inverse square law between an α-particle and a nucleus of gold breaks down at a distance of about 3×10^{-12} cm. It may be assumed that a closer approach than this is necessary for the disintegration or transmutation of a nucleus. The force of repulsion between an α-particle and a nucleus is greater than that between a proton and a nucleus since the charge on the proton is one-half that on the α-particle. Before 1932, disintegrations had not been obtained with α-particles of energy less than 3 mev. In 1932 it seemed that because of the proton's smaller charge disintegration might be produced by protons of energies much less than 3 mev. Cockcroft and Walton of Cambridge, England, immediately set out to test this idea. In order to obtain protons of energy approaching 1 mev they used an ingenious device based on the following principle: If three capacitances are charged to, say, 250 kv in parallel and then by a switching system are connected in series (in cascade) the voltage available is 750 kv, provided that there has been no leakage during the switching process. By this method Cockcroft and Walton were able to obtain 0.70 mev protons.

Crockcroft and Walton bombarded a thin film of lithium metal with protons of varying energy. They used the method of scintillations to determine what particles, if any, were emitted by the lithium. No particles other than protons were found when the energy of the protons was less than 0.125 mev. At this energy, however, scintillations having the brightness of α-particle scintillations appeared and the range of the α-particles was found to be 8.4 cm in air. Furthermore, Cockcroft and Walton obtained evidence that the α-particles were emitted in pairs, one forward and the other backward in regard to the original direction of the bombarding protons. The process which occurs is described by the nuclear equation

$$_{3}\mathrm{Li}^{7} + {}_{1}\mathrm{H}^{1} \rightarrow 2\,{}_{2}\mathrm{He}^{4} + w \qquad (20\text{–}15)$$

where w is the increase in kinetic energy. From (19–11), assuming p = 76 cm, T = 290° Abs, and using R = 8.4 cm, we obtain W = 8.3 mev. But there are two α-particles, so that the increase in kinetic energy is $2 \times 8.3 - 0.125 = 16.5$ mev.

Disintegration or transmutation is produced when certain substances are bombarded with deuterons. Such a case is shown by the nuclear equation

$$_{10}\mathrm{Ne}^{20} + {}_{1}\mathrm{H}^{2} \rightarrow {}_{9}\mathrm{F}^{18} + {}_{2}\mathrm{He}^{4} \qquad (20\text{–}16)$$

In this case an isotope, $_{9}\mathrm{F}^{18}$, of ordinary fluorine, $_{9}\mathrm{F}^{19}$, is produced.

20.8. Einstein Mass-Energy Law.—Perhaps the best experimental support for the law

$$\text{Mass} = \text{energy}/c^2 \tag{20–17}$$

occurs in nuclear physics. According to Sec. 20.7, Cockcroft and Walton obtained $w = 16.5$ mev in the nuclear process described by (20–15). According to Table 20.2, the exact masses of the nuclei involved are ${}_1H^1 = 1.00812$, ${}_3Li^7 = 7.01804$, and ${}_2He^4 = 4.00388$ mu. The total mass on the left side of (20–15) is 8.02616 mu and that on the right side is 8.00776 mu. There is thus a loss of mass of 0.01840 mu. From Sec. 19.7 this loss of mass is equivalent to an energy of $0.01840 \times 934 = 17.2$ mev. According to Einstein's law this must appear as a gain in kinetic energy. It should be noted that the masses used are for neutral atoms with their orbital electrons. Actually, the nuclei are completely stripped of the orbital electrons, so that the mass on the left side of (20–15) is $8.02616 - 4 \times 0.00055$, while the mass on the right side is $8.00776 - 4 \times 0.00055$, where 0.00055 is the mass of the electron in mu. Since the same mass is subtracted from both sides of the equation, it is customary to neglect the masses of the electrons. When the inaccuracy of (19–11) and the experimental error in the range obtained by Cockcroft and Walton are taken into account, the value of 17.2 mev agrees well with the value of 16.5 mev. This result gives experimental evidence of the validity of Einstein's law (20–18). So in 1932 the possibility of releasing atomic (more correctly, nuclear) energy was shown. Since 1932 many experimental results have proved beyond doubt the validity of Einstein's law.

20.9. The Geiger-Müller Tube.—We have described an ionization chamber in Secs. 14.5 and 14.6. In Sec. 14.6 a graph of the ionization current as registered by the electroscope against the voltage across the chamber is given. The current increases to the saturation value as the voltage increases. The potential to produce saturation is usually of the order of 50 volts. Suppose, however, that the voltage is very greatly increased, say, to the order of 50,000 volts, the gas in the ionization chamber being at atmospheric pressure. The graph now has the form shown in Fig. 20.6. At increasingly high voltages the curve takes a sudden turn upward as shown by AB. At B the voltage is so high that a spark is just about to jump and the current just about to become larger by several powers of ten. An ionization chamber which is operated at a voltage just below the sparking potential is very sensitive to ionizing agents which enter it. When, for instance, an α-particle enters the chamber a track of ions is produced and these positive and negative ions move under the high electric field which exists in the chamber toward their respective electrodes.

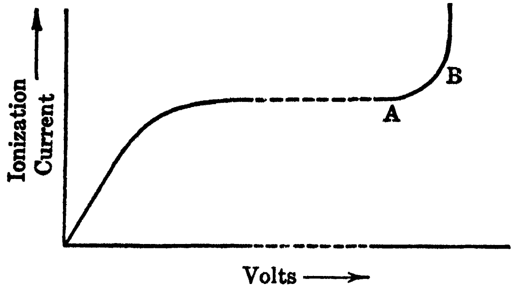


FIG. 20.6.—*AB* shows the variation of the current through an ionization chamber with the voltage as the sparking potential is approached.

A Geiger-Müller tube is an ionization chamber which is on the verge of a corona discharge. An example of a Geiger-Müller tube or counter

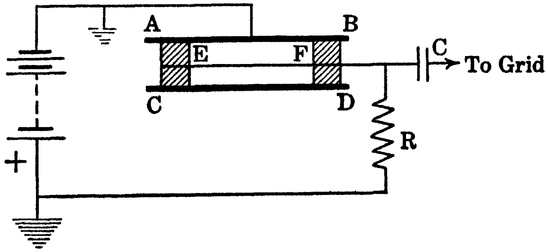


FIG. 20.7.—Geiger-Müller tube and connections.

and its connections is shown in Fig. 20.7. *ABCD* is a metal tube of length 10 cm and diameter 2 cm. *EF* is a 5-mil (1 mil = 1/1000 inch) tungsten wire supported by gas-tight insulators. The tube contains nitrogen at a pressure of about 5 cm of mercury and the potential between the metal cylinder and the wire is about 1200 volts. The wire *EF* is connected through the resistance *R* to the positive pole of the battery. The central wire of a Geiger-Müller counter is always the ***anode.***

From (5–31), the difference of potential between a point *P* and the inner cylinder of two coaxial cylinders is

$$V_r = \frac{V}{\log \frac{b}{a}} \cdot \log \frac{r}{a} \tag{20–18}$$

where V is the difference of potential between the inner cylinder of radius a and the outer cylinder of radius b, and r is the distance of the point $P(a < r < b)$ from the axis of the cylinders. If V is in volts, V_r is in volts. For the Geiger-Müller counter just described $V_r = 677$ volts when r is 1.0635 mm. The radius of the wire is 0.0635 mm so that the potential drop in the first $1.0635 - 0.0635 = 1.00$ mm from the wire is 677 volts, while that in the next $10.00 - 1.0635 = 8.9365$ mm is also 677 volts. The electric field near the wire is thus very strong. In the ionization process, an electron is knocked out of an atom or molecule. In a strong electric field the electron is pulled over to the anode before it has time to attach itself to an atom or molecule and make a negative ion. An electron must have an energy of about 35 ev to produce a pair of ions in nitrogen. In the region within 1 mm of the wire the electron collects an energy sufficient to produce a maximum of $677/35 = 19$ pairs of ions. But this is the greatest possibility. Actually, less than this number is produced. If, say, 10 pairs are produced, the negative ion in each case is an electron and this electron gathers kinetic energy as it moves under the electric field. Some of these electrons are produced at 1 mm from the wire and some close to the wire. For simplicity we assume a distance of 0.5 mm. For further simplicity we suppose this region to be broken up into 5 concentric zones each of thickness 0.1 mm. An electron enters the fifth zone, counting outward from the wire. In this zone the electron produces another electron so that 2 electrons enter the fourth zone. Following this line of reasoning we find that 2^2 electrons enter the third zone, 2^3 the second zone, 2^4 the first zone, and $2^5 = 32$ electrons hit the central wire. This multiplying process is known as a *Townsend avalanche*, named for Townsend of Oxford, England.

So far, we have not considered the role of the positive ions which are produced. Since the Geiger-Müller counter contains nitrogen, the positive ion is usually N_2^+ or perhaps N^+. In the first case it has a mass $2 \times 14 \times 1840 = 51{,}520$ and in the second case 25,760 times that of the electron. Hence the positive ions move very sluggishly under the action of the electric field. After a fraction of a microsecond the wire is surrounded by a positive ion sheath of sufficient strength to inhibit the "avalanche" process in the region near the wire. The discharge is then said to be quenched. However, the positive ions move, though slowly, to the metal cylinder of the Geiger-Müller counter. When they hit the cylinder they are apt to produce secondary electrons. If such a secondary electron is produced it moves towards the central wire and another Townsend avalanche results. This is not desirable. We now introduce the term *threshold voltage*. A Geiger-Müller counter will not operate unless the voltage between the central wire and the cylinder is above the threshold voltage. The central wire becomes charged negatively during the Townsend avalanche. This charge is neutralized by a conventional current flowing upward through the resistance R in Fig. 20.7. There is a voltage drop in R and the voltage between the wire and the cylinder becomes less than the threshold voltage. If this condition continues until the last positive ion has reached the cylinder, no further Townsend avalanches start during this period. After the last positive ion has reached the cylinder the negative charge on the wire is neutralized and the counter is set to record the entrance of a second α-particle or other ionizing agent into the counter.

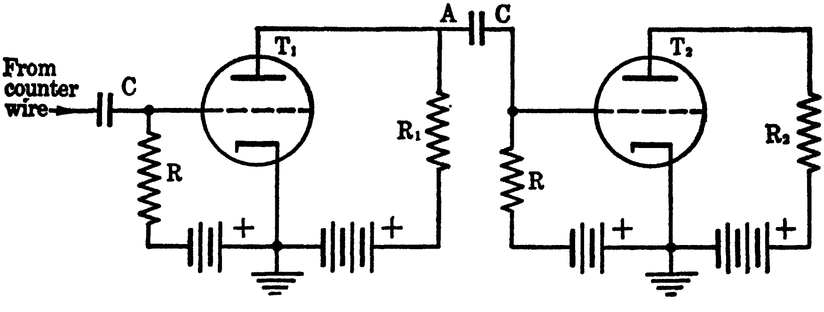


FIG. 20.8.

The wire is connected through a condenser to the grid of a triode. When a Townsend avalanche occurs in the Geiger-Müller counter, a negative pulse appears on the grid of the triode. This causes a sudden drop in the plate current of the triode circuit. For many purposes a sudden increase in the plate current is more desirable. A negative pulse is changed to a positive pulse by the method shown in Fig. 20.8. The negative pulse from the wire causes a sudden decrease of the plate current through R_1; this causes a sudden increase in the potential of A;

and this then gives a positive pulse to the grid of the triode on the right. This in turn causes a sudden increase in the plate current through R_2. If R_2 is replaced by a loud-speaker, a Townsend avalanche in the Geiger-Müller counter is registered by a click in the loud-speaker. It has become the practice nowadays to abbreviate "Geiger-Müller counter" to "Geiger counter" even though the original Geiger counter differed somewhat from the present Geiger-Müller counter. Hereafter, we shall refer to the device as a Geiger counter.

In the Geiger counter just described the time of transit for the last positive ion to the metal cylinder is about 100 μ sec. The time constant RC of the resistance and capacitance shown in Fig. 20.7 must not be less than 100 μ sec. Consequently, the α-particles or other ionizing agents must enter the Geiger counter at intervals greater than 100 μ sec in order to produce separate pulses on the grid of the first triode in Fig. 20.8. The Geiger counter can deliver somewhat less than 10,000 pulses per second. Pulses of this frequency are fed through a series of devices to a mechanical recorder which usually cannot register more than 100 separate events per second. The series of devices is a *scaling* circuit with several stages. In each stage every other input pulse is passed on to the succeeding stage. We therefore speak of a scale of 2, a scale of 4, a scale of 8, and so on. A scale of 64 is common. In a scale of 64, every 64th pulse is passed on to the mechanical recorder. If the Geiger counter delivers 6400 pulses per second, the mechanical recorder will register 100 counts per second. The description of a scaling circuit is beyond the scope of this book.

In recent practice, nitrogen rather than air is used in Geiger counters since oxygen, which has an affinity for electrons, tends to spoil the avalanche process. Counters containing 90 per cent argon and 10 per cent alcohol vapor are also widely used.

20.10. The Neutron.—In 1930 Bothe and Becker of Germany bombarded targets of lithium, beryllium, and boron with the 5.2 mev α-particles of polonium (RaF) and found that an unusually penetrating radiation was produced. These results were confirmed in 1931 by Joliot and Irene Curie in France. Joliot and Curie's experiment is shown diagrammatically in Fig. 20.9. A slab of lead is placed between the Be target and a Geiger counter, G-M, or some other kind of ionization chamber. The Be target emits "rays" which penetrate several centimeters of lead to actuate the counter. In January 1932, Joliot and Curie placed different screens S between the counter and the lead.

Nothing unusual happened until they used a screen of paraffin. Instead of absorbing the "rays" coming from the Be, the paraffin caused the counts in the Geiger counter to increase greatly. Chadwick of Cambridge, England, immediately seized upon the results of Joliot and Curie and began experiments of his own. Perhaps one of the best experiments was made by Feather, also of Cambridge, England, who replaced the Geiger tube in Fig. 20.9 by a Wilson cloud chamber and found that the increase of counts produced by the screen of paraffin was due to swiftly moving protons. The α-particles fall upon the Be target and produce "rays" which pass through the lead. These "rays" then fall upon the paraffin which is rich in hydrogen, and eject high-speed protons which are detected by the counter or cloud chamber. The question then is: What is the nature of these "rays"? There were two hypotheses. Because the "rays" can pass through such a great thickness of lead they are either γ-rays or particles of matter which carry no charge. These chargeless particles are called neutrons. The protons which were ejected from the paraffin in the forward direction of the "rays" in Fig. 20.9 have a kinetic energy of 5.65 mev. If a screen containing nitrogen is placed at S it is found that the "rays" eject N nuclei with an energy of 1.61 mev in the forward direction. If one of these recoiling atoms is due to the Compton effect of a γ-ray quantum we have for a head-on collision

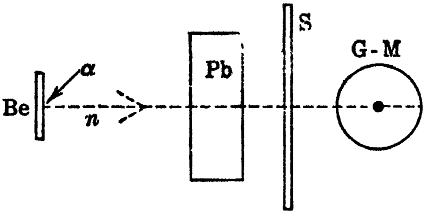


FIG. 20.9.—Production of neutrons.

$$h\nu_1 = h\nu_2 + (\tfrac{1}{2})mv_2^2 \tag{20-19}$$

from the conservation of energy, and

$$h\nu_1/c = -(h\nu_2/c) + mv_2 \tag{20-20}$$

from the conservation of momentum. The quantum $h\nu_2$ bounces from the nucleus of mass m in the backward direction and v_2 is the velocity of recoil of the nucleus. The kinetic energy of the recoiling nucleus is

$$W_2 = (\tfrac{1}{2})mv_2^2 \qquad (20\text{–}21)$$

Then, eliminating ν_2 and v_2 from the above three equations, we obtain

$$h\nu_1 = (W_2/2) + \sqrt{mc^2W_2/2} \qquad (20\text{–}22)$$

for the energy of the γ-ray quantum which ejected the nucleus with a kinetic energy W_2 from the screen S, Fig. 20.9. The value of mc^2 in (20–22) is $934A$ mev where A is the atomic weight of the recoiling nucleus. Hence (20–22) may be written in the form

$$\underset{\text{in mev}}{h\nu_1} = (W_2/2) + \sqrt{934AW_2/2} \qquad (20\text{–}23)$$

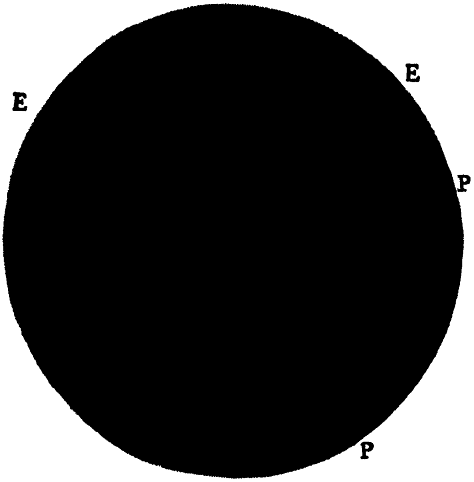
where W_2 is in mev. Hence, for recoiling protons we put $A = 1$, $W_2 = 5.65$, mev and obtain $h\nu_1 = 54.1$ mev. For the recoiling nitrogen nuclei $A = 14$, $W_2 = 1.61$ mev and we obtain $h\nu_1 = 103.5$ mev. But the same agency produces both the recoiling nitrogen nuclei and the recoiling protons. The agency therefore should have the same energy in both cases. Hence the assumption of γ-rays and the Compton effect is invalid.

We now try the idea of particles of mass m_1 and velocity v_1 making elastic impacts with hydrogen or nitrogen nuclei. The problem is similar to that of Fig. 20.3 and Sec. 20.4. The energies of the recoiling and impinging particles are related by (20–5). Let us first put $W_{p2} = 5.65$ mev and $m_p = 1$ for the recoiling protons and then put $W_{p2} = 1.61$ mev and $m_p = 14$ for the recoiling nitrogen nuclei. We then have two simultaneous equations which we can solve for $W_{\alpha 1}$ and $m_{\alpha 1}$ which refer to the impinging particle. The kinetic energy turns out to be 5.7 mev and the mass of the particle to be 1.16 mu. When the "rays" are assumed to consist of particles these particles have the same energy whether they eject nitrogen nuclei or protons from the screen S in Fig. 20.9. The above argument when applied to the experimental facts enabled Chadwick to announce the discovery of the neutron in February 1932.

The nuclear reaction which occurs when the α-particles strike the beryllium target in Fig. 20.9 is represented by

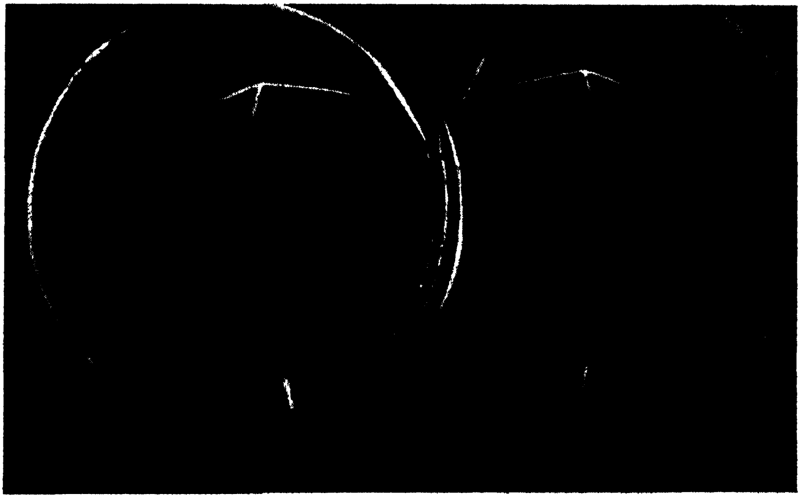
$${}_4\mathrm{Be}^9 + {}_2\mathrm{He}^4 \rightarrow {}_0n^1 + {}_6\mathrm{C}^{12} \qquad (20\text{–}24)$$

where ${}_0n^1$ represents a particle carrying zero charge and having a mass number unity. The other product of the reaction is a carbon nucleus.



Courtesy of Carl D. Anderson and The Physical Review

FIG. 20.10.—Cloud Tracks. A negative photograph, the tracks being black. At the top a track *EE* due to a 120 mev electron; at the right a track *PP* due to a 130 mev proton. The electron track is thinner than the proton track. The bending of the tracks is produced by a 17,000 gauss magnetic field.



Courtesy of T. W. Bonner and W. M. Brubaker and The Physical Review

FIG. 20.11.—Cloud Tracks. A positive photograph, the tracks being white. A stereoscopic picture of a disintegration of nitrogen into two α-particles and a Li nucleus. This disintegration was produced by a 12.9 mev neutron.

20.11. Mass of the Neutron.—Although neutrons are produced most plentifully when beryllium is bombarded by α-particles, lithium, boron, sodium, and several other elements will serve as target material for the production of neutrons. With boron for target material, the nuclear reaction is

$$_5B^{11} + {_2He^4} \rightarrow {_7N^{14}} + {_0n^1} + w \qquad (20\text{–}25)$$

Data obtained from the mass spectrograph give $_2He^4 = 4.00388$, $_5B^{11} = 11.01286$ and $_7N^{14} = 14.00750$. The increase w in the kinetic energy is somewhat difficult to estimate because the expelled neutrons have no ionizing properties and do not cause cloud tracks in a Wilson cloud chamber. However, the neutrons occasionally make head-on collisions with the protons in a paraffin screen. We know that the mass of the proton nearly equals that of the neutron and we know that when one body makes an elastic head-on collision with a second body of equal mass the second recoils with all of the kinetic energy of the first body. The kinetic energy of the recoiling proton can be measured from the length of its cloud track. In this way Chadwick estimated w as -1.5 mev. By the Einstein mass-energy law this is equivalent to $-1.5/934 = -.00161$ mu. Then, solving (20–26) for $_0n^1$, we obtain 1.00763 mu for the neutron. The most recent value is 1.00893 mu.

Neutrons, as well as protons, deuterons, and α-particles, produce transmutation. An example is given by the nuclear equation

$$_8O^{16} + {_0n^1} \rightarrow {_2He^4} + {_6C^{13}} \qquad (20\text{–}26)$$

It should be noted that the collision represented by (20–26) is inelastic. When neutrons knock protons out of paraffin as in Sec. 20.10 the collision is elastic. Transmutation does not occur when the collision is elastic, that is, when there is no change in kinetic energy.

20.12. Artificial Radioactivity.—In Chapter XIX we have described the experimental facts concerning natural radioactivity—the radioactivity which occurs in certain elements found in nature. Nowadays we think of the radium as exploding according to the nuclear reaction

$$_{88}Ra^{226} \rightarrow {_2He^4} + {_{86}Rn^{222}} \qquad (20\text{–}27)$$

The nuclear reactions we have previously discussed in the present chapter are the result of the collision of an agent with a nucleus. In (20–27) the radium nucleus disintegrates into an α-particle and a nucleus of radon. However, all the radium nuclei do not instantaneously disintegrate in the fashion of (20–27) but only a fraction of the nuclei disintegrate in a given interval of time. In the above reaction half the radium nuclei disintegrate in 1600 years. In some radioactive transformations a fast-moving electron or β-particle is ejected as in

$$_{82}\mathrm{RaB}^{214} \rightarrow {}_{83}\mathrm{RaC}^{214} + {}_{-1}e^{0} \qquad (20\text{–}28)$$

The symbol $_{-1}e^{0}$ represents a particle carrying a charge -1 and having zero mass number. The electron is such a particle. Half the RaB nuclei break up in 27 minutes.

In 1934 Joliot and Curie discovered the phenomenon of artificial radioactivity. Anderson's discovery of the positron two years before will be described in Chapter XXII. Suffice it to say here that the positron has the same mass as the electron but that it carries a positive charge numerically equal to the negative charge on the electron. For the purposes of nuclear physics the positron is represented by $_{+1}e^{0}$. In Joliot and Curie's experiment targets of boron, magnesium, and aluminum were bombarded with α-particles. They found that these targets were able to register effects in an ionization chamber or Geiger counter for several minutes after the α-particles had ceased to bombard the target. The targets became radioactive with the activity dying down according to the exponential law of natural radioactivity. This new phenomenon is known as artificial radioactivity. However, the rays given out by the artificially radioactive target in Joliot and Curie's experiment were neither α-, β-, nor γ-rays. Instead, the rays were found to be swiftly moving positrons which have the same ability to ionize a gas as ordinary fast moving electrons. When α-particles collide with boron nuclei the reaction

$$_{5}\mathrm{B}^{10} + {}_{2}\mathrm{He}^{4} \rightarrow {}_{0}n^{1} + {}_{7}\mathrm{N}^{13} \qquad (20\text{–}29)$$

may occur. However, the $_{7}\mathrm{N}^{13}$ nucleus is not stable but breaks up sometime later according to the radioactive transformation

$$_{7}\mathrm{N}^{13} \rightarrow {}_{+1}e^{0} + {}_{6}\mathrm{C}^{13} \qquad (20\text{–}30)$$

The half-life of $_{7}\mathrm{N}^{13}$ nuclei is several minutes. This substance is called radio-nitrogen.

Artificially radioactive nuclei which eject ordinary electrons have been discovered since the original experiment of Joliot and Curie. Radio-phosphorus, $_{15}\mathrm{P}^{32}$, is produced by bombarding ordinary phosphorus with deuterons according to

$$_{15}P^{31} + {}_1H^2 \rightarrow {}_1H^1 + {}_{15}P^{32} \tag{20-31}$$

This radio-phosphorus disintegrates with the ejection of an electron according to

$$_{15}P^{32} \rightarrow {}_{-1}e^0 + {}_{16}S^{32} \tag{20-32}$$

Radio-phosphorus is used in the treatment of leukemia, a disease in which there is an excessive number of white cells in the blood. The half life of $_{15}P^{32}$ is about 14 days. One millicurie (mc) of $_{15}P^{32}$ is given by mouth to the patient in the form of sodium phosphate once a week. Phosphorus, whether ordinary or radioactive, goes mostly to the liver, spleen, and bone-marrow. The radio-phosphorus treatment of leukemia has lengthened the life of the patient beyond his life expectancy in many cases.

At present a *curie* of a radioactive substance is defined as that amount of the substance which gives the same number of radioactive disintegrations per second as a gram of radium. The number of atoms in 1 gm of Ra is $6.03 \times 10^{23}/226 = 2.66 \times 10^{21}$. The number of disintegrations of 1 gm of Ra in 1 sec is $2.66 \times 10^{21} \times 0.693/P$, where P is the half life of Ra in seconds. Recent determinations give the half life of radium as 1580 years, so that the number of disintegrations is therefore 3.70×10^{10} sec^{-1}. It is usual to measure the strength of a radioactive substance in millicuries (1 mc = 0.001 c). A millicurie of a substance is that amount which gives 3.70×10^7 disintegrations per second.

CHAPTER XX

PROBLEMS

1. If the maximum voltage between the semicircular boxes of Lawrence's cyclotron is 5000 volts, what is the radius of the 225th semicircular orbit of the protons, given that the strength of the magnetic field is 4000 gauss? What is the energy of these protons?

2. Consider the nuclear reaction represented by Eq. (20–15). Assume that the helium nuclei have equal kinetic energies and that their lines of travel make equal angles with the line of travel of the bombarding proton, the angles being on opposite sides of this latter line. Find the kinetic energy of each helium nucleus when the kinetic energy of the proton is 1.2 mev and find the angle that the path of each helium nucleus makes with the path of the bombarding proton.

3. When neutrons from a certain source bombard oxygen and produce the reaction of Eq. (20–26) it is found that the sum of the kinetic energies of the helium and carbon nuclei which are produced is 5.50 mev. Determine the kinetic energy of the neutron producing the reaction.

4. A reaction of the type of Eq. (20–30) may be considered as an explosion. Since the mass of the expelled electron is so much smaller than that of the carbon, almost the whole of the kinetic energy of the particles after the explosion is possessed by the electron. Given that the kinetic energy of the electron is 4.5 mev determine the isotopic weight of $_7N^{13}$.

CHAPTER XXI

NUCLEAR FISSION

21.1. Bombardment with Neutrons.—In 1934 it occurred to Fermi, then of the University of Rome, Italy, that neutrons, because of their lack of charge should be effective in penetrating nuclei, especially those of high atomic number which repel protons and α-particles strongly. A radium salt and powdered beryllium enclosed in a gas-tight brass tube is a powerful source of neutrons. The radium is soon in equilibrium with its α-emitting daughter substance radon, and this with succeeding α-emitting daughter substances of the third and later generations. The purpose of the gas-tight brass tube is to prevent the radon from escaping. The α-particles from the radon have an energy of 5.3 mev and on hitting the beryllium nuclei produce a reaction described by

$$ {}_2\mathrm{He}^4 + {}_4\mathrm{Be}^9 \rightarrow {}_6\mathrm{C}^{12} + {}_0n^1 + w \qquad (21\text{–}1) $$

where w is the increase of the kinetic energy of the particles after the collision above that of the particles before the collision. The loss of rest mass is 0.00600 mu, corresponding to a positive value of w of 5.6 mev. The conservation of momentum requires that the neutron carry the greater part of this energy, so that the neutron from the reaction has a kinetic energy of somewhat less than $(5.3 + 5.6) = 10.9$ mev. A neutron with an energy of this order of magnitude is called a "fast" neutron.

Using the neutrons from such a source, Fermi found that in many cases the nucleus of the bombarded atom captured a neutron with the production of an unstable nucleus which then attained stability by emitting an electron. When uranium was used as the target, a process which could be described by

$$ {}_{92}\mathrm{U}^{238} + {}_0n^1 \rightarrow {}_{92}\mathrm{U}^{239} \qquad (21\text{–}2) $$

was obtained. This process was then followed by a second process which could be described by

$$ {}_{92}\mathrm{U}^{239} \rightarrow {}_{-1}e^0 + {}_{93}\mathrm{X}^{239} \qquad (21\text{–}3) $$

where ~~X represents~~ the new element. After prolonged exposure to the neutron source, Fermi obtained a mixture of β-ray emitting substances with half-lives of 10 sec, 40 sec, 13 min, and 90 min. There was a possibility that the new substances might have atomic numbers lower than that of uranium, but this was ruled out at that time. The new substances were placed in the 93rd, 94th, 95th, and 96th positions in the periodic table. These elements were called transuranic elements; they did not exist in nature. As time went by, the original four half-lives increased to nine and elements up to 97 were imagined, some of them being isotopes of others.

21.2. The Discovery of Fission.—In 1938 Irene Curie and Savitch in Paris found still another radioactive substance in uranium which had been bombarded with neutrons. This substance had a half-life of 3.5 hours. The bombarded uranium salt was dissolved and a salt of lanthanum ($Z = 57$) added. The lanthanum salt was then precipitated by chemical means and it was found that the radioactive substance came down with the lanthanum. Curie and Savitch could find no chemical means of separating the radioactive substance from lanthanum. They were, however, unwilling to conclude that the substance was an isotope of lanthanum; they preferred to place it in the already crowded territory beyond uranium.

Hahn and Strassmann of Germany repeated Curie and Savitch's results and, in addition, found that if a barium salt is used as a carrier in place of the lanthanum salt another radioactive substance comes down with the barium and cannot be separated from the barium by any known chemical means. This new substance had a half-life of 86 min. Moreover, at that time, it was known that an isotope of barium had a half-life of about 86 min. Still, Hahn and Strassmann preferred to believe that the new substance was an isotope of radium which had chemical properties similar to but not the same as those of barium. After many tedious fractional precipitations and crystallizations, the isotope of "radium" still persisted in coming down with the barium. Finally, it occurred to Hahn and Strassmann in early 1939 that, since the new substance always behaved chemically like barium, perhaps it was barium. This world-shaking idea meant that something new and startling had been discovered. It was suggested that the total process consists of that described by (21-2) and an entirely new process perhaps described by

$$_{92}U^{239} \rightarrow {}_{56}Ba + {}_{36}Kr + x\,{}_{0}n^{1} \qquad (21\text{-}4)$$

where x was unknown at the time. It should be noted that the mass numbers of Ba and Kr (krypton) are omitted. The isotope of Ba which has a half-life of 86 minutes has a mass number of 139. In 1939 the heaviest known isotope of Kr had a mass number of 88. In 1944 Seaborg of the University of California published a table of isotopes showing that the heaviest isotopes of Ba and Kr had mass numbers, known with any degree of certainty, of 140 and 89. The best that we can say is that the mass numbers and the value of x on the right side of (21–4) are uncertain.

In 1939 the production of lanthanum, as observed but not recognized by Curie and Savitch in 1938, could be explained by the process

$$_{92}U^{239} \rightarrow {}_{57}La + {}_{35}Br + x{}_0n^1 \tag{21–5}$$

As in the case of (21–4), the mass numbers and the value of x on the right side of (21–5) are uncertain. The symbol Br represents bromine. Further, the Ba shown in (21–4) may not be a primary product of the process. From Seaborg's tables, $_{55}Cs^{139}$ is a β-ray emitter with a half-life of 7 min. The process is described by

$$_{55}Cs^{139} \rightarrow {}_{-1}e^0 + {}_{56}Ba^{140} \tag{21–6}$$

The $_{56}Ba^{140}$ thus produced has a half-life of 86 min, so that the 86-min product may not be a primary product. Furthermore, $_{55}Cs^{139}$ is a daughter substance of $_{54}Xe^{139}$ which is a β-ray emitter with a half-life of less than 0.5 min. The symbols Cs and Xe represent cesium and xenon, respectively.

Previous to 1939, all observations had led to the view that the recoiling heavy nucleus had ejected a particle such as an α-particle, a deuteron, a proton, or an electron. The mass of each of these ejected particles is much less than that of the recoiling nucleus. But this new process involved the splitting of a nucleus into two fragments each having very roughly the same mass. The process is not unlike that of cell division in biology. Since the division of a cell into roughly two equal parts is called fission, Lise Meitner and Frisch suggested the term *fission* to describe processes like those of (21–4) and (21–5). Much of the credit for the discovery of fission is due to Meitner, a German woman physicist born in Austria. Just before the idea of fission was published, she fled from Germany to Denmark, where she communicated her ideas to Bohr, the Danish physicist. Meitner finally took up residence in Sweden.

21.3. Mass Surplus and Mass Deficit.—In 1939 it was known that the elements at or near the beginning and the end of the periodic table have atomic masses which are slightly greater than their respective mass numbers. These elements are said to have a mass surplus. ~~Over~~ a wide region about the middle of the periodic table the atomic masses are slightly less than the corresponding mass numbers. These elements are said ~~to have a mass deficit.~~ The atomic masses which may be involved in (21–2) and (21–4) are shown in Table 21.1.

TABLE 21.1
Atomic Masses

U^{238}	238.14	Kr^{84}	83.94
Ba^{138}	137.92	n^1	1.01

These values are for the most abundant isotopes of U, Ba, and Kr and are correct to the second decimal place. The values are on the basis of $O^{16} = 16.00000$. They are taken from a table published by Cork of the University of Michigan in 1947.

21.4. Energy Released by Fission.—The total mass on the left of (21–2) is 239.15 mu whereas that on the right of (21–4) is 239.03 when an appropriate value of x is inserted in (21–4) to balance the equation and it is assumed that the most abundant isotopes are involved. This indicates a loss of 0.12 mu. According to Einstein, this loss of rest mass must be compensated by a gain of kinetic energy as given by

$$\Delta E = c^2 \Delta m \tag{21–7}$$

where Δm is the loss of rest mass. If Δm is in mu we have

$$\Delta E \text{ (in mev)} = 934 \Delta m \text{ (in mu)} \tag{21–8}$$

Hence, the energy released by a single fission process is $934 \times 0.12 = 112$ mev approximately. The energy released depends upon the particular process involved and this in turn depends upon the primary fission products of the process. In 1940 Turner of Princeton University gave a tentative value of about 200 mev for the average energy released per fission for the various possible fission processes.

We shall now compare Turner's value with some other energies that are familiar to us. A steel rifle bullet traveling with a speed of 2500 ft/sec has a kinetic energy per atom of iron of 0.17 ev $= 1.7 \times 10^{-7}$ mev. The energy released per fission of one nucleus is thus $200/(1.7 \times$

$10^{-7}) = 1.18 \times 10^9$ = more than *one billion* times that of an atom in a rifle bullet in flight. One of the most energetic of chemical reactions is that of hydrogen burning in oxygen as represented by the chemical process

$$2H_2 + O_2 \rightarrow 2H_2O + w \tag{21-9}$$

where ~~w is the energy released~~. If a gram of hydrogen is burned, the energy released in the form of heat is about 30 kg-cal. A kilogram-calorie is 1000 calories. The heat released per atom of hydrogen works out as about 2.1×10^{-12} erg or about 1.3 ev. The average energy released by fission of a single nucleus is 1.54×10^8 = 154 million times that released per atom of hydrogen burned in oxygen.

The first indications of the release of atomic (more correctly, nuclear) energy were obtained by Frisch in Copenhagen and Joliot in Paris in early 1939. Frisch used a uranium-lined ionization chamber through the walls of which neutrons were shot. Forgetting for the moment the neutrons on the right sides of (21–4) and (21–5), we note that the two fragments of roughly equal mass will divide the kinetic energy of 200 mev between them. Frisch suggested that perhaps each fragment would have a kinetic energy of about 70 mev. These nuclear fragments are highly charged and as they plough their way through the gas of the ionization chamber they will produce a trail of ion-pairs. An energy of 35 ev is required for each ion-pair, so that there is a total of 70 mev/35 ev = 2×10^6 ion-pairs for each fragment. The fragments will travel in opposite directions, one going into the gas of the chamber and the other into the walls. The central electrode is made the cathode and is connected to the grid of a triode as shown in Fig. 21.1. When a fission fragment enters the chamber, a maximum charge of $2 \times 10^6 \times 4.8 \times 10^{-10} = 9.6 \times 10^{-4}$ esu = 3.2×10^{-13} coulomb appears on the cathode and grid. A 7 mev α-particle would give rise to 2×10^5 ion-pairs and a charge of 3.2×10^{-14} coulomb on the electrode and grid. These charges are positive pulses and they drive the grid of the triode less negative. The bias of the grid can be set so that the recording device S registers fragment pulses but not α-particle pulses. Further, by finding the negative bias of the grid just necessary to prevent S from registering it is possible to find the size of the charge on the central electrode if the capacitance of the central electrode and grid is known. Assuming a capacitance of $2\mu\mu f$, we obtain a potential pulse of 0.16 volt. This may be amplified by several stages so as to operate the recording device. Frisch found a charge corresponding to 2×10^6 positive ions. The purpose of R is to drain off the charge and to set the ionization chamber ready for the next pulse of ions. The recording device may be a cathode ray oscilloscope.

In regard to (21–7) and (21–8), we have spoken of a loss of rest mass. In Secs. 13.9 and 13.11 it was stated that conservation of mass is a cardinal principle of the relativity theory. According to this theory an object of rest mass m_0 moving with velocity v has a mass $m_0/\sqrt{1 - \beta^2}$ where $\beta = v/c$. In the process of fission the total mass is the same after fission as before. Even if γ-rays are emitted the total mass is unaltered since the mass of a photon of energy $h\nu$ is $h\nu/c^2$. Energy has mass and mass is a form of energy. Perhaps mass as we ordinarily think of it is a form of potential energy. We are accustomed to thinking of kinetic energy as kinetic energy and not as mass. Yet, a box containing a flywheel is heavier when the flywheel is rotating than when not. Likewise a copper ball is heavier when hot than when cold.

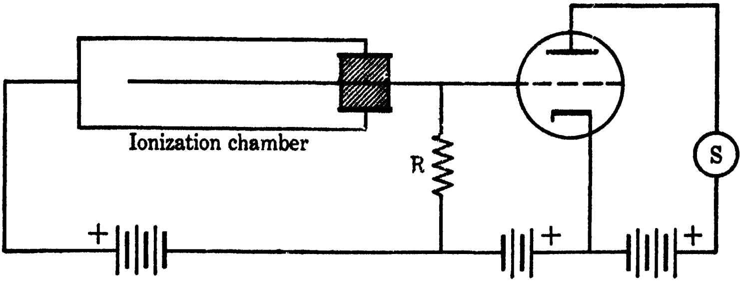


FIG. 21.1.—Apparatus for measuring pulses of charge produced by fission fragments in an ionization chamber.

21.5. Collision Cross Section.—So far we have not discussed the probability of a collision of a neutron with a nucleus occurring. If the original line of flight of the center ~~of a neutron comes within~~ a certain distance b of the center of a nucleus and if the neutron on approaching the nucleus is captured or deflected elastically, the area πb^2 is known as the collision cross section for the process involved. We speak of elastic-collision, inelastic-collision, capture, and fission cross sections. If a neutron passes through a thin layer of thickness dt containing dN_1 nuclei, the area of the layer being S, the probability of a collision occurring is $\pi b^2 dN_1/S$, which is the ratio of the total area of circles of radius b to the area S. But

$$dN_1 = (S\rho N dt)/A \tag{21-10}$$

where ρ is the density of the substance, N is Avogadro's number, and A is the atomic weight. Hence the probability of a neutron colliding with a nucleus in the layer is $(\sigma\rho N dt)/A$, where σ has been written for the cross section πb^2. If N_2 neutrons enter the layer the number lost by collision is

$$dN_2 = -\mu N_2 dt \tag{21-11}$$

where

$$\mu = \sigma\rho N/A \tag{21-12}$$

Eq. (21-11) is similar to (14-2) for the absorption of x-rays. If μ is a constant, we can integrate (21-11), obtaining

$$N_2 = N_{20} \exp(-\mu t) \tag{21-13}$$

where N_2 is ~~now the number~~ of neutrons penetrating through a layer of thickness t, and N_{20} is the number entering the layer.

In the case of x-rays we note from Table 14.5 that μ/ρ for x-rays of wave-length 1A in copper is 133.0. Subtracting the mass scattering coefficient of 0.2, we find $\tau/\rho = 132.8$ for the photoelectric absorption. From (21-12)

$$\sigma = (\tau/\rho) \cdot (A/N) \tag{21-14}$$

Putting in values for copper, we obtain $\sigma = 1.39 \times 10^{-20}$ cm^2 for the collision photoelectric cross section of x-rays of $\lambda = 1.00$A in copper. The distance $b = 6.65 \times 10^{-11}$ cm. When the line of flight of the center of a photon of $\lambda = 1.00$A is closer to the center of a copper atom than 6.65×10^{-11} cm, photoelectric absorption of the photon occurs. Nuclear physicists have come to use the *barn* as a convenient unit of cross section. It is defined by

$$1 \text{ barn} = 1 \times 10^{-24} \text{ cm}^2 \tag{21-15}$$

The term barn is said to have arisen from the common description of a poor marksman as being one who could not hit the side of a barn The cross section for the absorption process just described is 13,900 barns. The smallest cross section given by Cork is of the order of 10^{-8} barn.

21.6. Uranium-235.—Soon after fission was discovered in early 1939 Bohr suggested that it was probable that thermal neutrons are effective in producing fission of the uranium isotope $_{92}U^{235}$. A *thermal neutron* is a neutron which is moving with a kinetic energy equal to the average kinetic energy of translation of a gas molecule at room temperature. According to Secs. 8.4 and 8.5, this energy is $(3/2)kT$, where k is Boltzman's constant and T is the absolute temperature. For $T = 293°$ Abs and with $k = 1.38 \times 10^{-16}$ erg deg^{-1}, the energy is 6.06×10^{-14} erg = 0.038 ev. Bohr's suggestion was quickly found to be correct. Moreover it was found that the probability of fission of U^{235} by thermal neutrons is quite large. Natural uranium consists of 99.3 per cent of U^{238} and 0.7 per cent of U^{235}. The fission of U^{235} is described by some reaction like

$$_{92}U^{235} + {}_0n^1 \rightarrow {}_{56}Ba + {}_{36}Kr + x\,{}_0n^1 \qquad (21\text{–}16)$$

Fast neutrons whose energy is 1 mev or greater are required to produce fission of U^{238}. The neutrons produced by a process like (21–16) are fast neutrons with energies of the order of 1 mev. The probability of these fast neutrons producing fission of U^{235} is small. Hence in order to study a possible chain reaction in U^{235} it is necessary to slow down the fast neutrons to thermal neutrons by some method without losing the neutrons while they are being slowed down. A substance which has the property of slowing down neutrons is called a *moderator*. It was at first thought that ordinary hydrogen would be a good moderator. Since a neutron and an H^1 nucleus have nearly the same mass, an elastic collision results in the neutron losing half its energy on the average for each collision. To reduce the energy of a neutron from 1 mev to 0.038 ev requires log (1 mev/0.038 ev)/log 2 = 25 collisions on the average. However, hydrogen is a gas and takes up too much space. Water (H_2O) might be used but it contains oxygen and this complicates matters. Furthermore, the neutron can be captured by H^1 in an inelastic collision ~~according to~~

$$_1H^1 + {}_0n^1 \rightarrow {}_1H^2 \qquad (21\text{–}17)$$

A good moderator is one that slows down neutrons without capturing them. The inelastic collision described by (21–17) means that some of the neutrons will be permanently lost. For this reason H^1 is not a good moderator. Hydrogen-2, helium, beryllium, and carbon were considered as possible moderators.

21.7. The Uranium Pile.—When a neutron passes through a lump of natural uranium, there are four different processes which may occur with respect to the U^{238} isotope: (1) elastic scattering of the neutron, (2) inelastic scatttering, (3) capture without subsequent fission, or (4) capture with fission. The third process is known as *resonance* capture.

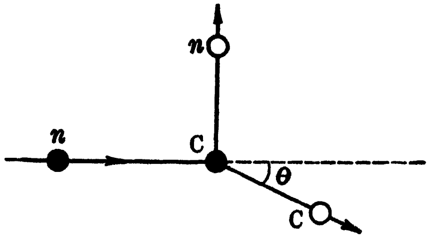


FIG. 21.2.—Elastic scattering of a neutron by a carbon nucleus.

It occurs only when the neutron has an energy in a small range about some energy intermediate between that of a thermal and that of a fast neutron. However, when the neutron is in this energy range the resonance cross section and the probability of capture are very great. The *multiplication factor k* is the number of neutrons which, being produced by a fission process such as in (21-16), themselves produce fission. When $k > 1$, a chain reaction is possible. The high probability of resonance capture by a U^{238} nucleus makes the mulitplication constant $k < 1$ and so a chain reaction is impossible in natural uranium.

At first it was assumed that the uranium and the moderator would be mixed in a homogeneous mass. However, about 1941 Fermi, then in the United States, and Szilard suggested that the moderator and uranium be kept separate and arranged in a certain manner. The arrangement is known as a "pile." The first pile was successfully operated at the University of Chicago on December 2, 1942. Blocks of graphite, some of which contained lumps of natural uranium, were so arranged that the uranium lumps were in a cubic lattice somewhat like the atoms of sodium and chlorine in a rocksalt crystal. If fission occurs in a lump, some of the neutrons get out of the lump before suffering resonance capture by U^{238} and are then slowed down by the moderator to thermal energies. These thermal neutrons in turn enter another lump and therein produce another fission process like (21-16). According to Cork, the "multiplication factor for the pile" may be defined as the ratio of the neutron density with the uranium lumps in place to the corresponding density with the lumps removed. If this factor is greater than one, a chain reaction results. As the blocks were assembled, the neutron density was measured as each few blocks were added. This was necessary in order that the chain reaction would not get out of hand and perhaps proceed with explosive violence. The size of the assemblage of blocks when the multiplication factor becomes greater than one may be called the *critical size* of the pile.

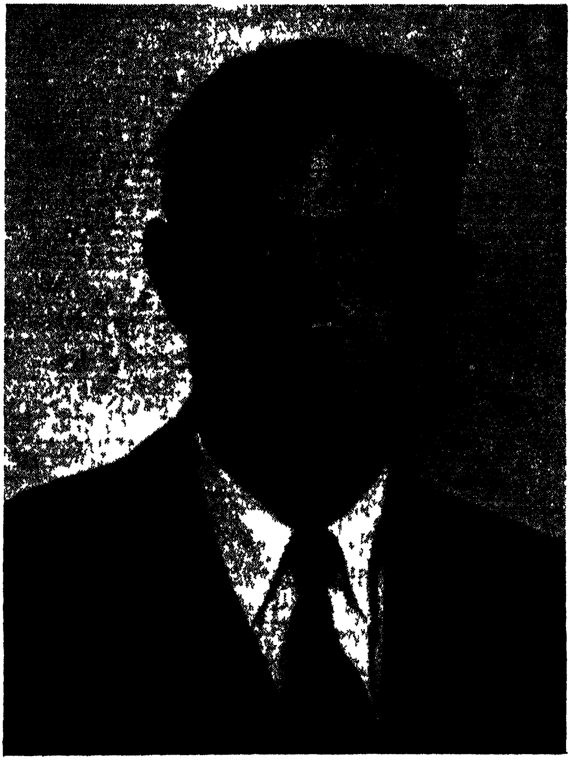


FIG. 21.3.—Enrico Fermi.

For reasons of safety if for no other reasons it was necessary to have some way of arresting the chain reaction when the critical size had been attained. Both boron and cadmium are good absorbers of thermal neutrons. In order to arrest the action of the Chicago pile, rods of cadmium or boron steel were inserted into the pile.

The density of neutrons at a location in the pile was measured by means of a *neutron counter.* Such a device is a Geiger counter filled with boron trifluoride, BF_3, which is a gas at room temperature, or a Geiger counter lined with boron. The nuclear reaction which occurs is described by

$$_5B^{10} + {}_0n^1 \rightarrow {}_3Li^7 + {}_2He^4 \tag{21-18}$$

The fraction of B^{10} in natural boron is 18.4 per cent and so natural boron may be used. The α-particles ($_2He^4$) which are produced in (21–18) ionize the gas in the Geiger counter and thus generate a pulse in the counter. These counters operate so efficiently because boron has a large cross section for the process shown in (21–18).

21.8. Plutonium.—We have spoken of resonance capture of neutrons of intermediate energy by U^{238}. This capture is represented by (21–2). However, resonance capture is not followed by fission. The U^{239} which is produced is a radioactive isotope of uranium with a half-life of 23 min. It breaks up with the emission of β-rays according to

$$_{92}U^{239} \rightarrow {}_{93}Np^{239} + {}_{-1}e^0 \tag{21-19}$$

The new element Np is called neptunium. This is a true transuranic element. But neptunium is also radioactive with a half-life of 2.33 days. It breaks up with the emission of β-rays according to

$$_{93}Np^{239} \rightarrow {}_{94}Pu^{239} + {}_{-1}e^0 \tag{21-20}$$

The new element Pu is called plutonium and is another transuranic element. Pu^{239} is radioactive with a half-life of some 10,000 years or so. It breaks up with the emission of α-rays according to

$$_{94}Pu^{239} \rightarrow {}_{92}U^{235} + {}_2He^4 \tag{21-21}$$

However, there are two important facts about plutonium. Its Pu^{239} isotope (there is also a Pu^{238} isotope) is fissionable by thermal neutrons according to a process similar to

$$_{94}Pu^{239} + {}_0n^1 \rightarrow {}_{57}La + {}_{37}Rb + x{}_0n^1 \tag{21-22}$$

where Rb represents rubidium. Also plutonium is a different chemical element from uranium and so can be separated from uranium by chemical means.

By changing the sizes of the carbon blocks and the uranium lumps in a pile, conditions can be obtained such that neutrons of intermediate energy enter each lump and the probability of resonance capture by U^{238} nuclei highly enhanced. As the pile operates, more and more plutonium appears in the uranium lumps of the pile. However, fission of U^{235} is also occurring in each lump and so the grains contain not only Pu^{239} but also fission products like La, Ba, Rb, and Kr, all of which are radioactive. After a certain interval, the uranium lumps are removed from the pile and subjected to chemical methods for separating the plutonium and the radioactive fission products from the uranium and from each other.

21.9. Atomic Power.—The large energy carried by the fission fragments in a pile finally appears as heat and provision must be made for cooling the pile. Air or water cooling may be used. If, however, we take another point of view, we can look upon the pile as a source of heat. This heat can then in turn be changed into mechanical energy by some sort of heat engine. A few days after December 2, 1942, the Chicago pile was run at a rate of 200 watts. It is possible to build piles whose power output would be many thousands of kilowatts. The city of St. Louis has a population of about 800,000 and a peak power consumption of about 800,000 kw. This amount of power is equivalent to 5×10^{21} mev/sec. If we assume 100 per cent efficiency, this would require the "burning" of $5 \times 10^{21}/200$ nuclei of U^{235} each second, or 2.5×10^{19} nuclei per second. A nucleus of U^{235} weighs approximately $235 \times 1.66 \times 10^{-24}$ gm $= 3.76 \times 10^{-22}$ gm. The rate of burning of the U^{235} is therefore 0.0089 gm/sec = 770 gm/day. This is nearly 2 lb of U^{235} per day. This is to be compared with the burning of 2800 tons of coal or 1800 tons of gasoline per day to produce a power of 800,000 kw. We see then what a tremendous source of power a uranium pile may be.

The critical size of a pile may be reduced by surrounding the pile with some material which reflects the neutrons which would otherwise escape back into the pile and also by using lumps of uranium which contain more than the natural amount of U^{235}. Such lumps are said to be *enriched* and we speak of an enriched pile. This enrichment requires a partial separation of U^{235} from natural uranium and is a lengthy and tedious process. U^{235} cannot be separated from U^{238} by chemical means. The separation must be effected by a physical means. Separation can be obtained for small quantities by means of a mass spectroscope. Also the rate of diffusion of a substance depends on the mass of the molecules of the substance, but the substance must be in the form of a gas. The only known gaseous compound of uranium is uranium hexafluoride, UF_6. The molecular weights of this substance for the two isotopes, U^{235} and U^{238}, are 349 and 352, respectively. The two molecular weights are so close that many consecutive diffusion experiments must be carried out. Still, it was possible to obtain uranium hexafluoride enriched in the U^{235} isotope.

The fission products of a pile are radioactive and this radioactivity may be very lethal. A pile must therefore be operated by remote control and careful provision must be made for shielding the operators from any lethal rays.

A pile is not only a source of power but also a source of neutrons. A very strong beam of neutrons can be obtained from a pile. This beam can be used for bringing about all sorts of nuclear reactions in the targets that they are caused to hit. Apertures in cadmium plates may be used to give a well-collimated beam of neutrons.

21.10. A Useful Notation.—In a nuclear reaction like that described in (21–1), the beryllium is the target material, the α-particle is the projectile which produces the reaction, the neutron is the ejected particle, and the carbon is the residue usually remaining in the material of the target. The reaction (21–1) is therefore abbreviated to $Be^9(\alpha,n)C^{12}$. The first symbol on the left refers to the target material, the first symbol in the parenthesis refers to the bombarding particle, the second symbol in the parenthesis refers to the ejected particle, and the last symbol, which is outside the parenthesis, refers to the residue nucleus. The bombarding and ejected particles are α-particles, deuterons, protons, electrons (sometimes the ejected particle is a positron), and gamma rays. These are represented by the symbols α, d, p, $-e$, $+e$, and γ, respectively. The nuclear reaction (21–18) is abbreviated to $B^{10}(n,\alpha)Li^7$. The reaction (21–1) is called an (α,n) reaction and the reaction (21–18) an (n,α) reaction.

21.11. *K* Electron Capture.—In 1936 Yukawa and Sakata of Japan predicted that under certain conditions a positron-emitting nucleus may achieve the same end result by capturing an electron from the *K* shell surrounding the nucleus. When a positron is emitted, the charge on the nucleus and, consequently, the atomic number are each decreased by one, the mass number remaining unchanged. If the nucleus captures an electron from the *K* shell of the atom of which the nucleus is a part, the atomic number is again decreased by one, the mass number remaining the same.

In 1938 Alvarez of the University of California found a case of *K* electron capture. When titanium is bombarded by deuterons in the process $Ti^{46}(d,n)V^{47}$ the residue vanadium has a half-life of 600 days, the residue vanadium isotope, ${}_{23}V^{47}$, decaying to ${}_{22}Ti^{47}$ according to

$$ {}_{23}V^{47} + {}_{-1}e^{0} \rightarrow {}_{22}Ti^{47} \tag{21-22}$$

When this occurs the daughter titanium finds itself with only one instead of two electrons in its *K* shell. The vacancy is filled by an electron falling from an outer shell into the *K* shell and the *K* x-ray spectrum of titanium is emitted. The absorption of these rays was measured and was found to be that of the characteristic *K* x-rays from titanium.

The material for this chapter has largely been gathered from L. A. Turner's article on *Nuclear Fission* in the *Reviews of Modern Physics* for 1940, G. T. Seaborg's *Table of Isotopes* in the *Reviews* for 1944, and J. M. Cork's *Radioactivity and Nuclear Physics* published in 1947. The reader might well refer to any of these sources for further information.

CHAPTER XXI

PROBLEMS

1. The latent heat of vaporization of water at 100° C is 540 cal/gm. Determine the energy in electron-volts per molecule of water (H_2O) evaporated.

2. Determine the number of collisions on the average to reduce the kinetic energy of a neutron from 0.1 mev to 1 ev by elastic collisions with H^1 nuclei.

3. Determine the area of the circular Bohr orbit for one *K* electron revolving about a lead ($Z = 82$) nucleus stripped of all the other electrons. Determine the mass absorption coefficient of x-rays of a wave-length just shorter than the critical absorption wave-length in lead and find the corresponding absorption cross section.

4. If by some process two protons and two neutrons could be assembled to produce one nucleus of He^4, there would be a loss of rest mass. Find the release of energy (*a*) per atom, and (*b*) per kg of helium produced.

5. Find the mass of water evaporated at the boiling point by the production of 1 lb of helium as in Problem 4.

6. The peak electrical power output in the United States is given as about 50,000,000 kw. Assuming that this is maintained throughout 24 hours and that the efficiency is 100 per cent, calculate the amount of bituminous coal burned per day, given that the heat of combustion of the coal is 12,000 B.T.U./lb. Also calculate the amount of U^{235} "burned" per day to provide the same power.

CHAPTER XXII

COSMIC RAYS

22.1. Early Experiments.—In 1785 Coulomb of France showed that, when proper precautions are taken either to prevent or to correct for the leak across insulators, an electroscope slowly discharges through air. Over a hundred years later it was shown that much of this conductivity of the air is due to ions produced in the air by the α-, β-, and γ-rays of radioactive substances. Soon after 1900 it was shown that, when correction was made for the radioactivity of the walls of an ionization chamber and when all known radioactive sources were removed, the air in an ionization chamber is still slightly conductive. At that time it was thought that the ions causing this conductivity were produced by penetrating γ-rays coming from outside the chamber. It was supposed that these penetrating rays came from radioactive substances in the rocks and soil at the surface of the earth. If this is so, the rate of discharge of an electroscope should diminish with height above the earth. This inference was tested by balloon ascensions in Germany. In 1909–1911, Gockel, ascending to a maximum height of 14,000 feet and using an open ionization chamber, found that the rate of discharge diminished much more slowly with height than could be explained on the theory that the rays producing the ionization were of terrestrial or atmospheric origin. In 1911 Hess and somewhat later Kolhoerster using sealed ionization chambers made further balloon observations and found that at high altitudes the penetrating radiation actually increases with height. Hess then stated his conclusion that the origin of the penetrating rays was outside the earth and its atmosphere. This was the first suggestion of the cosmic origin of the rays.

22.2. Beginning of Recent Investigations.—With the coming of the World War I in 1914 all work on these penetrating rays was abandoned and it was not until 1925 that further investigations on these rays were begun by Millikan and his associates of the California Institute of Technology and by Kolhoerster in Germany. With the opening of this new attack on the problem there began one of the most romantic, interesting, and exciting periods in the history of physics. Physicists took their electroscopes and ionization chambers to the tops of high mountains; they sank them in deep lakes; they carried them to odd corners of the earth; they sailed on ships and ascended in balloons; and some of the most adventurous lost their lives in the quest for knowledge of the rays which soon after 1925 became known as cosmic rays. Furthermore, physicists of many nations vied with each other in friendly competition to unravel the mystery of the rays which come to our earth from outside the atmosphere. The nations represented in these investigations were: United States, British Commonwealth, Belgium, France, Germany, Holland, Italy, Mexico, Japan, Russia, Switzerland, and others. The atomic bomb was the result of the co-operation of many scientists and others working in secret in the United States. Cosmic rays, on the other hand, have been studied intensely by physicists in all parts of the one world in which we must all live.

Before 1925 the pre-war observations of Hess and Kolhoerster which indicated that at high altitudes the intensity of the penetrating rays increases with height were held in considerable doubt, and the deduction of Hess that the origin of the rays is outside the earth was strongly criticized. Millikan's first investigations were carried out by means of sounding balloons. The balloons were sent up in pairs, each pair carrying a recording electroscope, thermometer and barometer. The records were made on a moving photographic film driven by a clock spring. The pairs attained a height of about 10 miles when, after the bursting of one balloon, the other balloon of a pair slowly lowered the recording apparatus to the surface of the earth. Search for the apparatus was stimulated by offering a reward and as a result Millikan collected valuable information. It was found that the intensity of the penetrating rays increases with height but that the increase is less rapid than the experiments of Hess and Kolhoerster at lower altitudes had indicated. The rays were found to be far more penetrating than had previously been suspected and Millikan became convinced that the origin of the rays was far outside the atmosphere of the earth, and the name cosmic rays began to come into use.

It is very necessary in experiments on cosmic rays to avoid ionization in the electroscope or ionization chamber by radioactive substances at the surface of the earth. This can be done by surrounding the chamber with a sufficient thickness of lead, water or air. Since lead is slightly radioactive, Millikan and also Regener chose "snow fed" lakes. The water of such lakes is almost entirely free of radioactive contaminating material. The recording electroscope and ionization chamber are lowered into a lake. The rate of discharge of the electroscope is then plotted against the depth and in this way the absorption coefficient of the rays can be found. The results indicated that the rays were not homogeneous—that is, they do not all have the same absorption coefficient. Four values of μ, the linear absorption coefficient in water, were indicated—$\mu = 0.0080$, 0.0020, 0.0010, and 0.0003 cm^{-1}. These are all smaller than the smallest μ known for γ-rays. If the cosmic rays are of the same nature as γ-rays, it was inferred from a certain relation between wave-length and absorption coefficient that $\mu = 0.0003$ indicated a wave-length of 5×10^{-5} A or a photon of energy about 250 mev. However, the relation between wave-length and absorption coefficient was uncertain for very short wave-lengths and this value is to be taken only as indicating the immense energy of the photons.

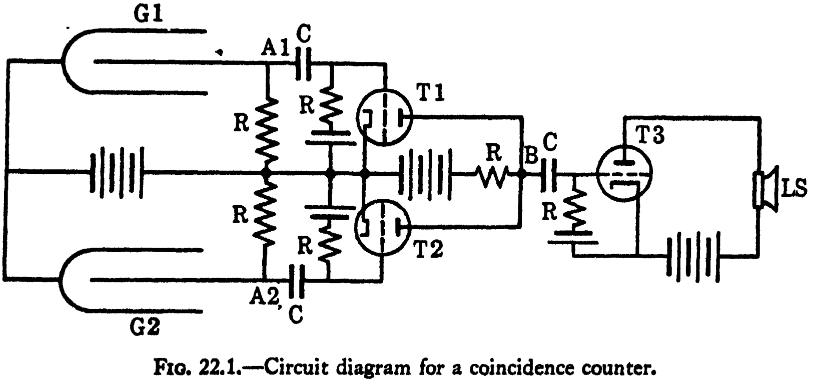


Fig. 22.1.—Circuit diagram for a coincidence counter.

22.3. Coincidence Counter Circuit.—It is desirable for many purposes in cosmic ray work to record ionizing events which occur simultaneously in two Geiger counters. If the same ionizing particle traveling with a speed approaching that of light passes through two Geiger counters 30 cm apart, the interval between the ionizing events in the two counters is 10^{-9} sec. This interval is so short that the two events may be considered as coincident.

The circuit diagram for coincidence counting is shown is Fig. 22.1. Geiger counters are shown as *G*1 and *G*2; triodes are shown as *T*1, *T*2, and *T*3. Capacitances and resistances are shown as *C* and *R*, respectively, although all the C's and all the R's may not be equal. If an ionizing particle passes through $G1$ and not through $G2$ there is a surge of current from the wire to the cylinder. This surge passes through the resistance in the circuit of $G1$ and suddenly lowers the potential of the point $A1$ in the Fig. 22.1. It should be remembered that the wire is always the anode. A negative pulse of potential appears at $A1$ and this is transmitted through C to the grid of $T1$. The potential of this grid being suddenly made more negative, $T1$ becomes less conducting and there is a small voltage drop in R of the plate circuit of $T1$. Consequently a positive pulse of potential appears at B and this is transmitted through C to the grid of $T3$. The bias of the grid of $T3$ is more negative than cutoff and is represented by $\overline{OF}$ in Fig. 22.2. The graph shown in Fig. 22.2 is the transfer characteristic of $T3$. The positive pulse is represented by $\overline{FG}$, Fig. 22.2. However, $\overline{OG}$ is more negative than cutoff and $T3$ does not conduct. Consequently, the recording device LS does not register.

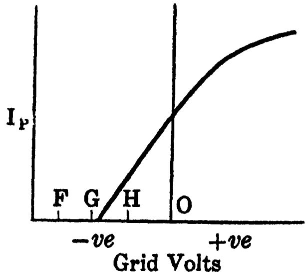


Fig. 22.2.

Now suppose that the same ionizing particle passes through $G1$ and $G2$. In this case, negative pulses of potential appear at both $A1$ and $A2$, Fig. 22.1, and are transmitted through the condensers C to the respective grids of $T1$ and $T2$. A double-size positive pulse of potential appears at B and this is transmitted through C to the grid of T_3. This double positive pulse is represented by $\overline{FH}$, Fig. 22.2, the grid of which is now more positive than cutoff and the recording device LS registers. One or two amplifying stages may be needed between $T3$ and LS in Fig. 22.1. If LS is a loud speaker a click will be heard whenever an ionizing particle passes through both $G1$ and $G2$.

22.4. Photons or Charged Particles.—The great penetrating power of the cosmic rays was at first believed to indicate that the rays must be photons of the same nature as γ-rays. However, in 1929, experiments by Bothe and Kolhoerster of Germany gave evidence that the cosmic rays consist largely of electrically charged particles. Their experimental arrangement is shown in Fig. 22.3. Two Geiger counters, GM, were placed one above the other and were surrounded by walls of lead and iron each about two inches thick. The purpose of the lead is to shield the tubes from rays of radioactive origin coming from terrestrial sources, while that of iron is to shield the tubes from the rays coming from the slightly radioactive lead. This use of a double wall for shielding instruments for measuring cosmic rays became standard practice for a few years. A slab of gold of thickness 4.1 cm could be placed at will between the counters. The Geiger counters are connected so as to record coincidences.

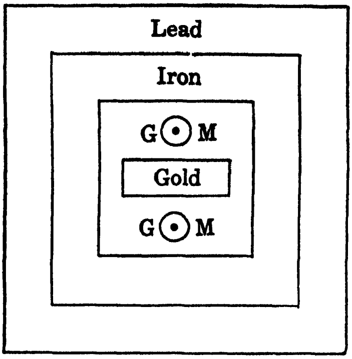


FIG. 22.3.—*Bothe and Kolhoerster's experiment.*

A counter tube only discharges when ions are produced in the tube by some agent. It was argued by Bothe and Kolhoerster that, if the agent is a photon, the photon is either absorbed in one chamber or the other. A photon proceeds through matter such as the gas in the counter without producing ions until it produces a recoil electron by means of the Compton effect or is entirely absorbed by an atom when by means of the photoelectric effect an electron is ejected. In either case the electron produces ionization along its path, but, so the argument runs, the electron will travel in any direction and will not usually pass through the second counter. On the other hand a high energy charged particle traveling in the downward direction will first pass through the upper counter tube producing a discharge and then through the lower tube producing a discharge. A charged particle produces ions along its path. High energy particles travel with a speed only slightly less than that of light. Bothe and Kolhoerster found that the gold block reduced the number of coincidences per hour to 0.78 of the number without gold. But this fraction is nearly equal to the reduction in cosmic rays observed in ionization chambers surrounded by an equivalent screen. Hence

Bothe and Kolhoerster inferred that the cosmic rays consist mostly of charged particles, since photons would not produce coincidences.

Following the publication of the above result and inference there arose a storm of argument. For a while physicists were divided into two camps—those who held to the photon theory and those who held to the charged particle theory of cosmic rays. Those who held to the photon theory admitted the experimental result found by Bothe and Kolhoerster and even the inference that the coincidences were due to charged particles passing through both tubes but they argued that these particles were not the original cosmic rays. The controversy then shifted to the question of primary and secondary rays. The photon supporters said that the rays which came to the earth from outside were photons and that these primary rays upon arriving in the atmosphere produced secondary rays which were charged particles and traveled in a direction very close to that of the primary ray. In support of this idea, these physicists referred their opponents to a closer study of the processes by means of which the photon loses energy—the process of photoelectric absorption and that of the Compton effect. In the first process the photon disappears and gives up all its energy, while in the second the photon gives up part of its energy to the electron. Let us now consider a head-on collision between a photon of energy $h\nu$ and an electron at rest. After the collision the electron recoils in the direction in which the photon was originally traveling and the photon bounces back with an energy $h\nu'$. Conservation of energy gives

$$m_0c^2 + h\nu = h\nu' + m_0c^2/\sqrt{1 - \beta^2} \tag{22–1}$$

where $m_0c^2/\sqrt{1 - \beta^2}$ is the mass-plus-kinetic energy of the recoil electron and βc is its velocity. Conservation of momentum gives

$$h\nu/c = -h\nu'/c + m_0\beta c/\sqrt{1 - \beta^2} \tag{22–2}$$

Eliminating ν' from (22–5) and (22–6), we have

$$(1 + \beta)/\sqrt{1 - \beta^2} = 1 + 2\alpha \tag{22–3}$$

where $\alpha = h\nu/m_0c^2$. From this, $\beta = 2(\alpha + \alpha^2)/(1 + 2\alpha + 2\alpha^2)$ and the kinetic energy of recoil is

$$m_0c^2\left(\frac{1}{\sqrt{1 - \beta^2}} - 1\right) = m_0c^2 \cdot \frac{2\alpha^2}{1 + 2\alpha} \tag{22–4}$$

But $\alpha = h\nu/m_0c^2$, so that the

$$\text{Recoil K. E.} = \frac{2\alpha}{1 + 2\alpha} \cdot h\nu \qquad (22\text{–}5)$$

But the energy of the photon is at least of the order 100 mev while $m_0c^2 = 0.5$ mev, so that α is of the order 200. In this case all but 0.25 per cent of the energy of the photon is given to the recoil electron.

Although the process just described is very rare, yet it does occur sometimes and we see that a photon entering the atmosphere may be replaced by a high energy electron traveling in the same direction. However, the high energy electron may make the kind of impact with an atom or nucleus that the cathode ray electron in an x-ray tube makes with an atom or nucleus of the target and so produce in its place a high energy photon. Also the electron will knock other electrons from some of the atoms it passes and these in turn may produce photons. Hence a photon or electron entering the atmosphere is finally replaced by a troupe of electrons and photons all traveling in approximately the direction of the primary photon or electron. This was the picture generally held in 1931.

22.5. Latitude Effect.—As shown in the previous section, the result found by Bothe and Kolhoerster with their coincidence counter did not necessarily mean that the primary cosmic rays entering the atmosphere were charged particles. However, the earth's magnetic field extends far beyond the atmosphere, and although the strength of the field is small, yet it extends over such large distances that the paths of charged particles should be bent by the field. The phenomenon of the aurora is explained as being due to electrons coming to the earth from the sun which emits electrons due to the thermionic effect. The paths of these electrons are bent by the earth's magnetic field so that there is a concentration of electrons in the neighborhood of the north and south magnetic poles. These electrons strike the upper rarefied air and produce the auroral displays which are seen in the northern latitudes of the northern hemisphere and in the southern latitudes of the southern hemisphere. Bothe and Kolhoerster pointed out that if the primary cosmic rays are charged particles the rays should be more concentrated near the poles than near the equator and they quoted a result found by Clay of Holland on a voyage in 1928, that a cosmic ray electroscope discharged more rapidly in Holland, latitude 52° N, than in Java, latitude 8° S.

However, no great weight was given to Clay's result since other experimenters found no change in the intensity of cosmic rays between California, 34° N, and Bolivia, 16° S, or between Melbourne, Australia, 39° S, and Antarctica, 63° S. Nevertheless, in 1930, Clay repeated his findings in Holland and Java. The question of the existence of the latitude effect lay open until 1932 when A. H. Compton, then of Chicago, and many others undertook a world-wide survey of the intensity of cosmic rays. Compton made readings on a voyage around the Pacific Ocean and established a very definite latitude effect. The effect is shown by the curve of Fig. 22.4. It will be noted that the curve consists of two flat portions *AB* and *CD*. Variation of latitude in either of these ranges will have no effect and so explains the lack of any effect found by several observers. Clay, however, happened to take one reading on the portion *CD* and the other on the portion *AB*. Fig. 22.4 is for measurements made at sea-level. The ratio of the cosmic ray intensity near the poles to that near the equator is 1.16 at sea-level. The best agreement between observers at different points on the earth's surface is obtained when the results are plotted against the geomagnetic rather than the geographic latitude. This result might have been expected since the magnetic pole of the northern hemisphere does not correspond with the geographic north pole. The geomagnetic latitude is used in plotting the curves of Figs. 22.4 and 22.5.

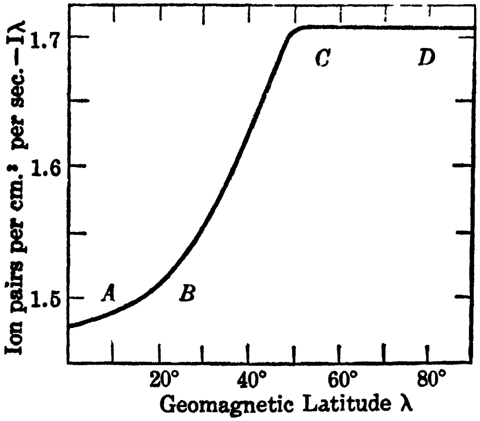


FIG. 22.4.—Variation of cosmic-ray intensity with latitude.

The latitude effect is more pronounced at altitudes above sea-level, as is shown by Fig. 22.5. The ratios I_{90°/I_{0° are 1.16, 1.22 and 1.33 for altitudes 0, 2000 and 4360 meters, respectively. The ordinates of Fig. 22.5 represent the number of ion pairs produced in 1 cm^3 of air at normal pressure and temperature in one second. When an electron is ejected from an atom of one of the molecules of air a pair of ions, one positive and the other negative, is produced. The rate of discharge of an electroscope or the ionization current in an ionization chamber measures the number of pairs of ions which are produced per second.

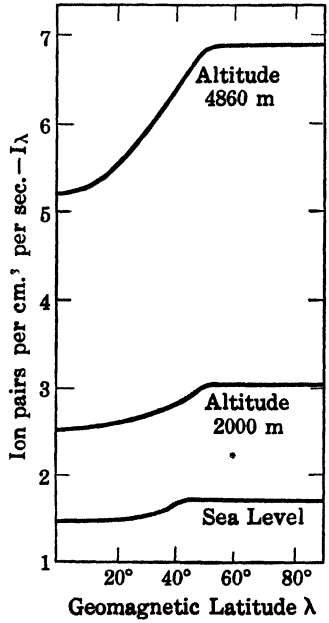


FIG. 22.5.—Variation of cosmic-ray intensity with latitude.

It has become the practice to measure the intensity of cosmic rays in the way just described. *In the United States at sea level the cosmic rays produce from 1.65 to 1.70 ion pairs per* cm^3 *per sec.* From Fig. 22.5 we see that not only does the ratio I_{90°/I_{0° increase with the altitude but the absolute magnitudes also increase. At an altitude of 4360 over the United States the cosmic rays produce from 6.5 to 7.0 ion pairs per cm^3 per sec.

From recent high altitude measurements in the stratosphere by means of balloons, the ratio I_{52°/I_{0° has been found to be about 40, while the value of I_{52° is about 500 ion pairs per cm^3 per sec. These measurements were obtained at a height such that the air pressure was only one-tenth of that at the surface of the earth. Compton estimates that at the top of the atmosphere the ratio I_{90°/I_{0° is not less than 100 and that I_{90° is not less than 1200 ion pairs per cm^3 per sec.

22.6. Theory of the Latitude Effect.—In 1933 Lemaitre of Belgium and Vallarta of Mexico announced the results of a theoretical investigation in which they assumed that charged particles from remote space are approaching the earth equally from all directions. They supposed a magnet of magnetic moment M to be placed at the center of the earth and to point in the magnetic north-south direction. As the particles approach the earth their lines of travel begin to be affected by the magnetic field of the earth. At a point some 15,000 miles above a point on the surface of the earth the earth's magnetic field is still 1 per cent of its value at the point on the surface. The bending of the rays by the field in a centimeter of path is very small but there are plenty of centimeters so that the bending by the time the rays arrive at the earth's surface may be considerable. The energy of a charged particle is not changed as it passes through the magnetic field, for the effect of the field is to supply a force which always acts at right angles to the line of travel of the particle and so the particle is neither slowed down nor speeded up by the field. Let us introduce the quantity x defined by

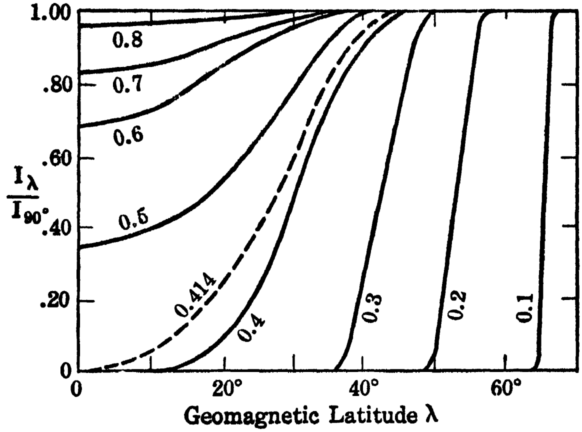


FIG. 22.6.—Curves for the theoretical variation of cosmic-ray intensity with latitude.

$$x = r(mv/ZeM)^{1/2} \qquad (22\text{–}6)$$

where m is the relativity mass of a particle traveling with velocity v and carrying a charge Ze and r is the radius of the earth. Lemaitre and Vallarta give the curves of Fig. 22.6 for the ratio of the number of particles I_λ arriving at a point on the earth's surface at a latitude λ to the number $I_{90°}$ arriving at the pole. The number on each curve refers to the value of x for that curve. We note that, for particles of a given value of x, the curve is horizontal from 90° to λ_1, the value of the latitude at the knee of the curve. The cosmic particles of a given value of x arrive at a point on the earth's surface from all directions above the horizon if the geomagnetic latitude of the point is in the range λ_1 to 90°. Also we note that, for particles of a given value of x which is not greater than $x = 0.414$, the curve drops to zero at a geomagnetic latitude λ_2. No particles of this value of x arrive at points between the geomagnetic equator and the latitude λ_2. Particles arrive from some directions above the horizon but not from others at points whose geomagnetic latitudes are between λ_2 and λ_1. When the particles have a value of x which is not less than 0.414 (see Fig. 22.6), the particles arrive at all latitudes, even arriving from some directions at the geomagnetic equator.

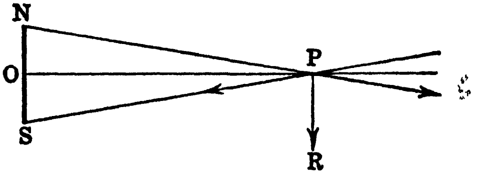


Fig. 22.7.—Magnetic field in the neighborhood of a magnet.
Note:—A + magnetic pole is assumed at the N pole in the figure. Actually a − pole. occurs near the N geographic pole of the earth.

Let NS in Fig. 22.7 represent a magnet with poles of strengths μ and $-\mu$ at N and S respectively. The field at P due to μ at N is μ/NP^2 along PQ and that due to $-\mu$ at S is μ/PS^2 along PS. The resultant of these is along PR and is equal to $2(\mu/NP^2) \cos \angle QPR$. But $\cos \angle QPR = NO/NP$ and $\mu \times 2NO = M$. Hence $H = M/NP^3$. But, if P is at a great distance from O, $NP \sim OP$. So, placing $OP = r$, we have

$$H = M/r^3 \tag{22-7}$$

The strength of the earth's magnetic field at the magnetic equator is about 0.32 gauss, while the radius of the earth is 6.37×10^8 cm, so that $M = 8.2 \times 10^{25}$ emu-cm. If the kinetic energy and the rest-mass energy of each of the particles are W and W_0 as measured in mev, the relation connecting x, W, and W_0 is, from (22-6)

$$(W^2 + 2WW_0)^{1/2}/Z = 6.05 \times 10^4 x^2 \tag{22-8}$$

where Ze is the charge on each particle.

Now it is to be expected that the particles approaching the earth from remote space are traveling with various energies, so that the sea-level curve of Fig. 22.4 is the weighted sum of curves such as shown in Fig. 22.6. The experimental curve may well consist of curves whose x values extend from 0.2 to 0.8. If the particles are electrons ($Z_0 = 0.51$ mev), we find from (22–8) that W extends from 2,420 to 38,750 mev. If the particles are protons ($Z_0 = 1.008 \times 934 = 940$ mev), W extends from 526 to 38, 280mev. It is worth noting that for high values of x, protons and electrons have the same energy. The paths of high energy protons are bent by the same amount in a magnetic field as those of high energy electrons. In 1933 it therefore appeared that particles of kinetic energy 3.8×10^{10} ev are present in cosmic rays.

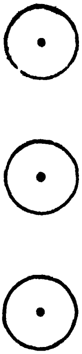


FIG. 22.8.—Triple coincidence counters.

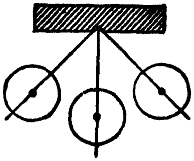
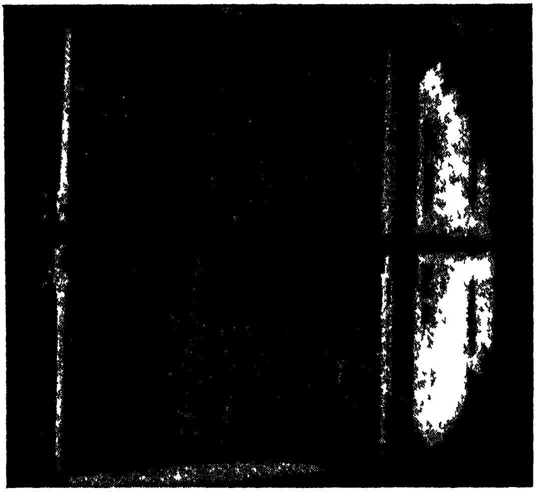


FIG. 22.9.—Rossi's experiment showing shower production.

22.7. Cosmic Ray Showers.—Interesting experiments have been performed with Geiger counters arranged to record triple coincidences. The circuit for such counters is based on the same principle as that underlying Fig. 22.1. The bias of the grid of the triode operating LS is such that LS does not register unless a triple-size positive pulse of potential is applied to the grid of the triode. If three counters are arranged in a vertical line as in Fig. 22.8, triple coincidences are recorded when an ionizing particle passes through all three counters. If there is no heavy material above the top counter the number of triple coincidences drops suddenly when the lowest counter is moved out of line with the upper two. In 1930 Rossi of Italy placed three counters which were out of line under a sheet of lead, as shown in Fig. 22.9. Rossi's experiment implied that at least three ionizing particles were simultaneously emanating from a point in the lead. A cosmic ray shower may be defined as that radiation emerging from a non-radioactive block of matter exposed to cosmic rays which will simultaneously discharge three Geiger counters placed out of line. The implication of Rossi's experiment induced physicists to look for evidence of showers by means of a Wilson cloud chamber, across the middle of which a lead plate had been placed. Nothing of interest can be shown in a cloud chamber picture unless an ionizing expansion is made. When the expansions occur at random instants, most of the cloud chamber pictures show no tracks. In 1933 Blackett and Occhialini, working at Cambridge, England, devised an ingenious arrangement in which the cloud chamber is put between two Geiger counters placed so that one is vertically above the other and connected so as to record coincidences. When a coincidence occurs, the ionizing particle must have passed through the cloud chamber between the counters. The counters are connected to a device which sets off the expansion mechanism of the cloud chamber and a picture is taken at those instants when an event of interest has occurred in the cloud chamber. The picture shown in Fig. 22.10 was obtained by Stevenson and Street of Harvard University. Evidently many



Courtesy of E. C. Stevenson and J. C. Street and The Physical Review

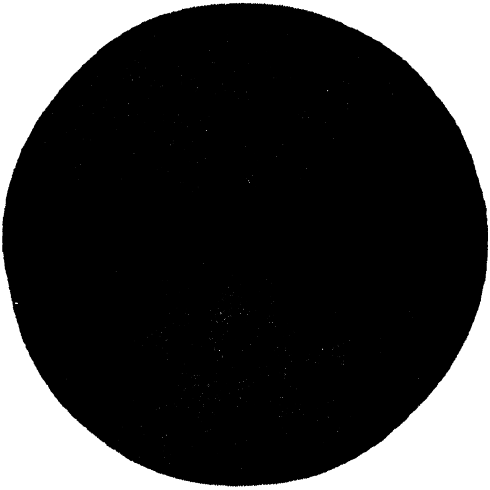
FIG. 22.10.—A shower of 22 rays produced by an electron. The slightest displacement of the ray above the lead is due to an irregularity in the expansion, a fact which was verified by a steroscopic photograph.

ionizing particles emanate from a small region in the lead. Fig. 22.10 shows a shower produced by an ionizing particle.

In order to obtain a cloud chamber picture of a shower which is produced by a non-ionizing ray entering the cloud chamber it is necessary to have three Geiger counters placed out of line under the cloud chamber. When a triple coincidence occurs the expansion mechanism of the cloud chamber is set off. Anderson of the California Institute of Technology found shower tracks emanating from the under side of the lead sheet, but no cloud track above the sheet. He inferred that the shower is produced by a non-ionizing ray, which may be either a photon or a neutron. This non-ionizing ray travels through the atomosphere and other matter without knocking electrons from the atoms until it makes a catastrophic collision with a nucleus. The energy of the ray is added to the nucleus, which then explodes into protons, electrons, and α-particles.

Showers are recognized by the small number of tracks—usually less than twenty—and the fact that the particles of the shower travel predominantly in the direction of the agent which produced the shower. The average angle of divergence is about 20°.

22.8. The Positron.—In 1932 Anderson began examining the bending of the tracks of ionizing cosmic rays by a magnetic field. He was fortunate enough to obtain a large electromagnet which not only gave a strong magnetic field but also a field of considerable extent. On August 2, 1932, Anderson obtained the Wilson cloud photograph shown in Fig. 22.11. The plane of the figure is vertical. A horizontal magnetic field was applied to the space occupied by the cloud chamber, and a sheet of lead of several millimeters thickness was placed across the chamber as shown. The curvature of the track is less below the lead sheet than above. Tracks due to protons and α-particles are thicker than those due to high energy electrons. The appearance of the track in Fig. 22.11 indicated that it was the track of an electron. The curvature of the track was such that, if the particle was traveling downward, it was charged negatively, whereas, if it was traveling upward, it was charged positively. The answer depends upon which way the particle was traveling. Now, the curvature of the path of a charged particle increases when the velocity of the particle becomes less. Hence the velocity of the particle above the lead was less than that below the lead. We can only imagine that the particle was slowed down as it passed through the lead and we therefore infer that the charged particle was moving upward. Anderson immediately recognized the particle as having at least approximately the same mass as the electron but as having a positive charge numerically equal to that of the electron. The name *positron* was given to this new particle. The name negatron is sometimes given to the ordinary electron in order to distinguish it from a positron. However, we shall use the terms electron and positron.



Courtesy of Carl D. Anderson and The Physical Review

FIG. 22.11.—Anderson's famous first picture of a positron cloud track. The 63 mev positron passed upwards through a 6 mm. lead plate, emerging on the upper side with an energy of 23 mev.

22.9. Electron-Positron Pairs.—In 1933 Chadwick, Blackett and Occhialini allowed the α-rays from polonium to fall upon a sheet of Be as shown in Fig. 22.12. The neutrons and photons which were produced passed through the wall of a cloud chamber and then through a sheet of lead into the main part of the chamber. A magnetic field was applied to the chamber. Pairs of tracks were observed, one track of a pair being curved upward and the other downward. The appearance of the tracks indicated that each track was due to an electron, while the curvatures of the tracks indicated that one track was produced by a positron and the other by an electron.

Blackett and Occhialini have suggested the hypothesis that under certain conditions a photon transmutes itself into an electron and positron. For this to occur

$$h\nu \geqq m_0c^2 + m_0c^2 \tag{22-9}$$

where $h\nu$ is the energy of the photon and m_0 is the rest mass of either the electron or positron. The process (22-9) cannot occur alone. If the photon disappears and an electron-positron pair appears in its place, the positron traveling with velocity $\beta_1 c$ and the electron with velocity $\beta_2 c$, conservation of energy requires that

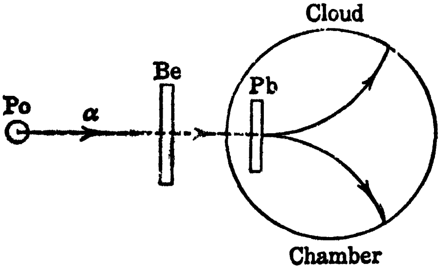


Fig. 22.12.—Production of electron (+ and −) pairs.

$$h\nu = \frac{m_0c^2}{\sqrt{1-\beta_1{}^2}} + \frac{m_0c^2}{\sqrt{1-\beta_2{}^2}} \tag{22-10}$$

Let θ_1 and θ_2 be the angles which the directions of travel of the positron and electron make respectively with the original direction of travel of the photon. Conservation of momentum requires that

$$\frac{h\nu}{c} = \frac{m_0c\beta_1 \cos\theta_1}{\sqrt{1-\beta_1{}^2}} + \frac{m_0c\beta_2 \cos\theta_2}{\sqrt{1-\beta_2{}^2}} \tag{22-11}$$

or, after multiplying throughout by c,

$$h\nu = \frac{m_0c^2\beta_1 \cos\theta_1}{\sqrt{1-\beta_1{}^2}} + \frac{m_0c^2\beta_2 \cos\theta_2}{\sqrt{1-\beta_2{}^2}} \tag{22-12}$$

Eqs. (22-10) and (22-12) are incompatible because $\beta_1 \cos\theta_1 < 1$ and $\beta_2 \cos\theta_2 < 1$. The transmutation can occur in the strong electric field near a nucleus, which takes part of the momentum. It is then possible for momentum to be conserved, although practically all the energy goes into the electron-positron pair which is produced. The energy m_0c^2 is 0.51 mev, so that the photon (γ-ray) must have an energy greater than 1.02 mev to produce an electron pair, as an electron-positron pair is now conveniently called. Photons of less than this energy have never been found to produce either electron pairs or positrons. The showers described in the previous section include positron as well as electron tracks.

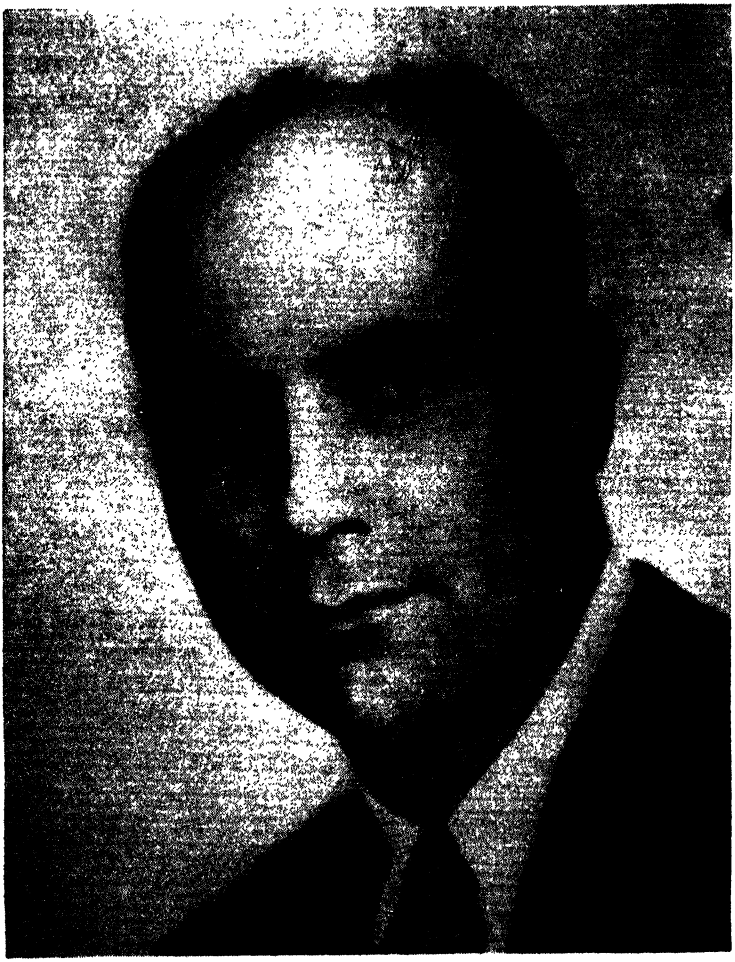


FIG. 22.13.—Carl David Anderson.

The production of electron pairs as photons pass through matter adds another cause for absorption of high-energy γ-rays in matter. In Sec. 22.2 it was stated that at the time when the absorption of cosmic rays in the water of snow-fed lakes was measured, the then assumed relation between absorption and wave-length indicated a wave-length of 5×10^{-5}A for the most energetic photons. However, the production of electron pairs increases the absorption coefficient. Not only do high-energy photons produce electron pairs but they may also shatter the nuclei, thus giving another cause for absorption. The relation between absorption and wave-length is thus very complicated for high energy photons. The departure from the ordinary law of absorption in the x-ray and soft γ-ray region begins at $\lambda = 0.048$A, which corresponds to a photon of 1.02 mev. Hence the estimated wave-length in Sec. 22.2 is much too large and the estimated energy much too low.

The reverse process of an electron and positron combining to produce a photon cannot occur unless a third body such as a nucleus is present. However, an electron and a positron can combine to produce two photons. If the electron and positron have small velocities, the energy of each of the two photons is m_0c^2 or 0.51 mev. The photons travel in opposite directions. In this case momentum can be conserved.

The hypothesis that a single photon can of itself produce an electron-positron pair is seen to be untenable for the following reason: According to the Doppler principle the frequency or energy of a photon emitted by an atom or some other source depends upon the relative velocity of the source to an observer. A photon of energy 1.02 mev as viewed from the source would have an energy of $0.9 \times 1.02 = 0.92$ mev when viewed by an observer moving with a velocity of 0.1c away from the source and in a direction opposite to that of the emitted photon. According to the observer the 0.92 mev photon has insufficient energy to produce an electron-positron pair. By means of the Doppler effect an imaginary observer can reduce the energy of any photon to below 1.02 mev. It is beyond all reason to believe that the production of an electron-positron pair depends upon the velocity of the observer relative to the source which emitted the photon.

22.10. East-West Asymmetry.—The theory of Lemaitre and Vallarta requires that, if more of the particles approaching the earth from remote space are charged positively than negatively, there should be more particles coming in from the west than from the east at points near the equator. To test this two coincidence counter tubes were set up as shown in Fig. 22.14. The ratio of the number of coincidence counts per minute when θ was on the west side of the vertical to the number of counts when θ was on the east side of the vertical was found. The results are shown in Table 22.1. The experiment was made in Mexico City, which is fairly close to the equator and has an elevation of 2310 meters. Further experiments have verified this east-west asymmetry and indicate that the cosmic rays consist of a greater number of positively charged particles than of negatively charged particles. Johnson has found that the east-west asymmetry vanishes at geomagnetic latitudes greater than 45°. This again is in agreement with the theory.

TABLE 22.1

θ	$\dfrac{\text{West Counts}}{\text{East Counts}}$
30°	1.08
45°	1.10
65°	1.05

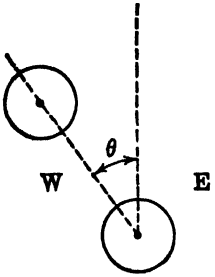


Fig. 22.14.—East-west asymmetry.

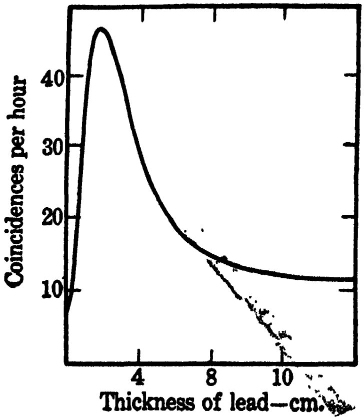


Fig. 22.15.—Rossi curve for small showers at an altitude of 2 miles.

22.11. Mechanism of Shower Production.—Rossi, using the arrangement shown in Fig. 22.9 with the counters connected to record triple coincidences, placed sheets of lead above the array of counters. By varying the total thickness of the lead he was able to find the relation between the thickness of the lead and the number of triple coincidences recorded per hour on the average. He obtained the graph shown in Fig. 22.15. The first few millimeters of lead increase the number of showers per hour. The rate of shower production increases until the thickness of lead is about 1.5 to 2.0 cm. In Fig. 22.15, the maximum rate of shower production is about 5 times the rate without a lead sheet. After reaching a maximum, the rate of shower production decreases with increasing thickness of lead. When the thickness of the lead is more than 5 or 6 cm the curve becomes almost horizontal, showing a very small decrease for thicknesses up to 15 cm. It is noteworthy that the curve does not drop to zero for large thicknesses of lead.

The shape of the Rossi curve is explained by the following three processes:

(1) When electrons hit the target of an x-ray tube they produce a line spectrum of x-rays which depends on the atomic number of the target material and a continuous spectrum whose minimum wavelength is independent of the atomic number of the target material. About 1900, Stokes and Thomson suggested that x-rays are produced by the slowing down of electrons. In 1946 it was believed that when an electron passes near a nucleus it is deflected and retarded. This retardation causes the electron to radiate energy according to the electromagnetic theory given in Sec. 6.4. The continuous spectrum is explained somewhat in this way. The line spectrum is explained by a process described in Sec. 18.4. The electromagnetic radiation produced by deflection and retardation of an electron is called *Bremstrahlung*. This might be translated "braking radiation." We suppose then that a positive or negative cosmic ray electron, on passing through the strong field of a nucleus, is retarded and produces *Bremstrahlung*. However, we must remember the quantum theory. The *Bremstrahlung* appears as a photon. As an aside at this place we should note that Sir William Bragg once said that on Mondays, Wenesdays, and Fridays we use the quantum theory and that on Tuesdays, Thursdays, and Saturdays we use the electromagnetic wave theory. And so the original cosmic ray electron becomes partly or wholly replaced by a photon.

(2) The *Bremstrahlung* process is followed by the inverse phenomenon: A photon passes through the strong field near a nucleus and gives rise to an electron pair.

(3) A slow positron passes near an electron and fuses with it, giving rise to two photons. The thickness of the sheet above the counters which gives the maximum rate of shower production increases with increase of the atomic number of the material of the sheet. This is consistent with the theory of *Bremstrahlung*. The intensity of the contin-

uous spectrum emitted by an x-ray tube increases with the atomic number of the target. Also it should be noted that in experiments with thin targets the continuous x-ray spectrum is more intense in the forward than in the backward direction relative to the direction of flight of the cathode rays in the x-ray tube. As a result of the foregoing three processes the particles of the shower proceed predominantly in the direction of the original particle or photon which produced the shower.

To explain the maximum in Rossi's curve we note that for small thicknesses of the lead screen more and more primary particles pass close to a nucleus as the thickness is increased. Consequently, increasing thickness gives more showers. But the primary particles also ionize the atoms of the screen and their energy is absorbed. When the thickness of the screen is greater than a certain critical value absorption overcomes shower production and the number of showers decreases with increasing thickness.

22.12. Cosmic Ray Bursts.—We have stated that at sea level in the United States the rate of production of ion-pairs in air at standard pressure and temperature by cosmic rays varies from 1.65 to 1.70 ion-pairs per second. The ionization does not occur continuously but occurs intermittently as ionizing particles pass through the ionization chamber. If the ionization chamber is connected to an electrometer the deflection of the electrometer grows unsteadily with the time. Some ionizing particles traverse a longer path in the chamber than others. Because the number of ionizing particles entering an ionization chamber varies from one interval of time to the next, and because the number of ion-pairs produced by each particle varies, the rate of deflection of the electrometer taken over short intervals of time fluctuates. We therefore speak of the statistical fluctuation in the rate of deflection. The rate of production of ion-pairs just given is the average over a long period of time. The increasing deflection of the electrometer can be recorded photographically on a moving film. In 1927 Hoffmann of Germany while studying such a record found evidence of a sudden jump in the deflection of the electrometer. This sudden jump was several times greater than the mean statistical fluctuation. A Geiger counter multiplies by several hundred or thousand times the number of ion-pairs produced by the original ionizing event, whereas the multiplication factor for an ionization chamber is unity. The jump in the deflection indicated a sudden production of the order of 10^7 ion-pairs in the ionization chamber. This phenomenon is called a burst. An α-particle of

10 mev energy produces 10 mev/35 ev = 385,000 ion-pairs if its energy is spent entirely in the ionization chamber. Consequently a burst is not produced by an α-particle. Since 1933 much has been learned about bursts with the aid of high-pressure ionization chambers. If an ionizing particle traverses an ionization chamber filled with argon at a pressure of 30 atmospheres the number of ion-pairs produced per cm is roughly 30 times greater than if the argon is at atmospheric pressure. If, in addition, the ionization chamber is a large sphere, say, 3 or 4 ft in diameter, the total number of ion-pairs produced in the ionization chamber can be very large. Such ionization chambers were constructed at Chicago by A. H. Compton and his co-workers. Bursts producing up to 10^{10} ion-pairs have recently been observed. This means that the energy of the original particle must have been 3.5×10^{11} ev.

Bursts were originally distinguished from showers by the fact that a burst is registered by an ionization chamber and electrometer without amplification whereas a shower is registered by 3 or more Geiger counters placed out of line. It was inferred that bursts were large showers. To test this inference, Wilson cloud pictures were taken. In 1936 Anderson and Neddermeyer obtained such a picture showing a shower of more than 300 tracks. The tracks had the appearance of electron tracks and Anderson and Neddermeyer estimated a total energy of over 1.5×10^{10} ev. A Rossi curve for bursts shows a maximum for a lead screen between 3 and 4 cm thick.

However, although 99 per cent of the cloud tracks produced by a burst are electron tracks, about 1 per cent of the tracks are proton and α-particle tracks with the proton tracks preponderating. The ionization produced by one of these heavy particles is so dense that, if a few such particles pass into an ionization chamber, the electrometer may register a burst. Consequently, not all bursts as registered by an electrometer are caused by large showers. These proton and α-particle tracks are believed to be the result of a nuclear explosion. However, these are best studied by a new technique which we shall describe in the next section.

22.13. Photographic Emulsion Technique.—This is an extension of the principle of the Wilson cloud chamber. When an ionizing particle passes through the gas of a Wilson cloud chamber, a trail of positive and negative ions is produced. The ions immediately begin to recombine and in a second or so the trail disappears. If an expansion occurs a few hundredths of a second after the trail is produced a cloud track is obtained. This cloud track must be photographed before it has time to disperse. A photographic emulsion consists of a colloidal mixture of silver bromide grains and gelatin. This emulsion is supported by a glass plate or by a flexible film of cellulose acetate. The diameter of a grain of silver bromide is of the order of the wave-length of visible light or about 5×10^{-5} cm. When light falls upon a photographic emulsion, a grain is either *infected* or not. If it is infected, the grain is reduced to metallic silver by subsequent development. A photographic emulsion need not be developed immediately after the picture is taken. The latent image, as it is called, may remain undeveloped for months, the infected grains remaining infected during this period. If now an ionizing particle passes through a photographic emulsion a trail of infected silver bromide grains is produced. This trail can be developed as a track of metallic silver grains several weeks or months later. The development corresponds to the expansion in a Wilson cloud chamber. The photographic emulsion technique has a great advantage over the cloud track method when a rare phenomenon is being studied. The plate with its emulsion can be exposed to cosmic rays for months. The nuclear explosions noted in the previous section can be studied in this way. The developed emulsion is examined under the microscope for interesting tracks. If anything of interest is observed a photomicrograph is taken of the developed emulsion. A diagram of such a microphotograph is shown in Fig. 22.16. In one case five tracks were found to diverge from a point in the emulsion. Such a phenomenon is described as a cosmic ray "star."

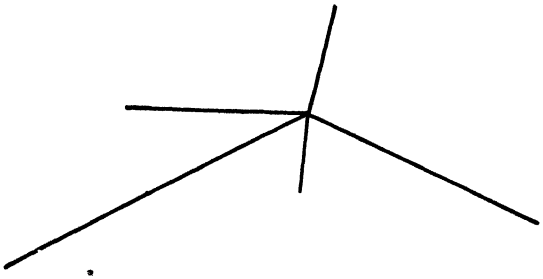


FIG. 22.16.—Diagrammatic representation of a fivefold "star" obtained in a photographic emulsion. Magnification about 500 diameters.

The rate of star production increases with rise in the atmosphere.

In 1941 Shapiro of Chicago found that a photographic plate exposed on Mount Evans, Colorado (14,200 ft) for eight months showed more than 100 stars per cm^2 of the emulsion. Several hours exposure in the stratosphere furnished a comparable yield of stars. Shapiro investigated many stars and found that the ionizing particles on an average travel equally in all directions—there is no preponderance in the direction of the agent which produced the star. This supports the idea that a star is the result of a nuclear explosion with the fragments as likely to travel in one direction as another.

22.14. Nuclear Drop Theory.—Bohr has likened the nucleus to a drop of liquid. An ordinary liquid drop is held together by cohesive forces which come into play when two molecules are within a certain small distance of each other. An ordinary drop has a temperature and the molecules of the drop have a certain average energy of agitation. At temperatures near the freezing point the rate of evaporation of the liquid drop is very small. However, if the drop is suddenly heated by some means a sizeable fraction of the molecules become sufficiently agitated to evaporate from the drop. Applying this general idea to the nucleus, Bohr ventures the hypothesis that the protons and neutrons of the nucleus are held together by cohesive forces. If now a primary cosmic ray electron passes through the nucleus, the *nucleons* (a nucleon is the name applied to a proton or a neutron since either may be a particle in the nucleus) may become agitated and the "temperature" of the nucleus rise sufficiently for the nuclear drop to divest itself of droplets in the form of protons and neutrons. The electron is not captured by the nucleus in this process. It loses energy to the nucleons in the nucleus and some of these nucleons evaporate. According to Bohr the evaporating nucleons have considerable kinetic energy. The protons produce dense ionization along their paths in a gas or a photographic emulsion.

22.15. Cosmic Ray Neutrons.—Bohr's nuclear drop theory implies that neutrons as well as protons should evaporate from a nuclear drop. Furthermore, the number of neutrons should be equal to or somewhat greater than the number of protons. A two-fold "star" such as that shown in Fig. 22.16 should be accompanied by perhaps two neutrons. However, neutrons make no record in a photographic emulsion, a Wilson cloud chamber, or an ordinary ionization chamber. Korff of New York University has made a neutron counter by filling the counter tube with boron trifluoride, BF_3, or by lining the inside surface of the metal cylinder with boron. The absorption cross section for a thermal neutron by a boron nucleus is comparatively large and the nuclear reaction

$$_5B^{10} + _0n^1 \rightarrow _3Li^7 + _2He^4 \qquad (22\text{-}13)$$

occurs fairly frequently. The α-particle, $_2He^4$, produces a dense trail of ions in the gas of the counter. If the counter is operated at a voltage somewhat below the threshold voltage for Geiger counter action (see Sec. 20.9), the counter operates as a *proportional counter.* In such a counter the size of the positive pulse of potential arriving on the grid of the triode *T*2 in Fig. 20.8 is proportional to the number of ions originally produced in the counter. If the bias of the grid of *T*2 is sufficiently negative, α-particles in the counter will be registered but not cosmic-ray electrons because these latter produce too few ions to drive the potential of the grid of *T*2 above cutoff.

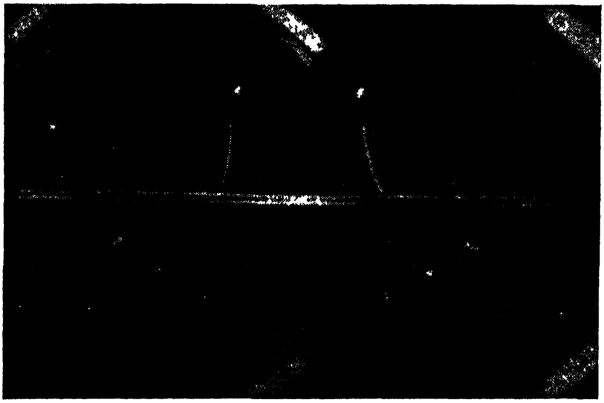


FIG. 22.17.—A stereoscopic cloud-tract picture. The thick, bent track above the lead sheet is a mesotron track in argon obtained at Pike's Peak (elevation 4300 meters). The magnetic field was 7900 gauss.

Korff found evidence for a neutron component of cosmic rays at sea level. The number of neutrons like the number of "stars" increases with height in the atmosphere.

22.16. The Mesotron.—After a large number of cosmic-ray cloud-track pictures had been taken by various experimenters, it was noticed that a few pictures showed tracks which were too dense for electron tracks but not dense enough for proton tracks. These anomalous tracks were so few in number that no conclusion concerning the nature of the particle producing them could be made. In 1936 Anderson and Neddermeyer obtained the cloud track picture shown in Fig. 22.17. The picture was taken with the cloud chamber in a magnetic field. There is a thick track above the lead screen placed across the chamber. The thin tracks below the lead screen are caused by electrons. Assuming that the particle above the screen is a proton, its range in the argon gas of the cloud chamber gives 1.5×10^6 ev for its energy, whereas its curvature in the magnetic field of 7900 gauss gives 1.6×10^5 ev for its energy. This difference of a factor of nine is outside the error of measurement. Relations between the energy of an ionizing particle and its range in air are given in (20–7) and Table 20.1. The numerical constants are 1.24, and 1.57, for protons and deuterons, respectively. These two constants are nearly proportional to $m^{1/3}$, where m is the mass of the ionizing particle. If this rule holds in the present case, we see from the form of (20–7) that we must assume a mass of $1/9^{2/3}$ of that of the proton for the anomalous particle. This would give it a mass of $1845/9^{2/3} = 430$ times that of the electron. This is a very rough estimate indeed. The evidence for the existence of this particle was very meager in 1936.

In 1933 Rossi of Italy and somewhat later Auger of France made experiments with the arrangement shown in Fig. 22.18. Two Geiger counters $G1$ and $G2$ with their axes parallel were arranged with $G1$ vertically above $G2$. The distance between the two counters was kept constant while slabs of lead of different thickness were placed between the two counters, which were connected to record coincidences. The curve of Fig. 22.19 was obtained at sea level. The number of coincidences per hour decreases very rapidly as the thickness t increases up to $t = 8$ cm. For $t > 8$ cm, the decrease is much less rapid. Auger therefore postulated two groups of particles: a "soft" group or component which is completely absorbed in about 10 cm of lead and a "hard" component which can penetrate more than 100 cm of lead. The ratio of the abundance of soft to hard particles at sea level is 0.3; at a height

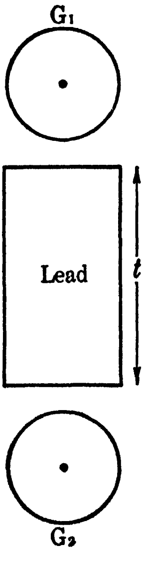


FIG. 22.18.—Rossi-Auger arrangement of coincidence counters.

of 2.2 mi, 1.0; at 9 mi, perhaps more than 4. Below rock of thickness equivalent to 80 ft of water the ratio approaches 0.1 but does not become smaller than this for greater depths. It would seem that under ground the soft particles are produced by the hard particles.

In 1946 it was believed that the cosmic rays entering the atmosphere from outside consist almost entirely of protons, positrons, and electrons. The absorption of positrons and electrons in lead has been studied from the theoretical point of view and it has been found that, owing to the production of *Bremstrahlung*, no electron of either sign can penetrate more than 15 cm of lead, no matter how great its energy. Further study is required for the case of protons. Yet the hard component does penetrate more than 100 cm of lead. The situation requires considerable thought.

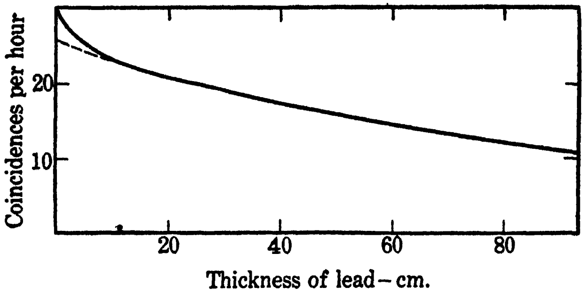


FIG. 22.19.—Hard and soft components in cosmic rays at sea level.

In the quantum theory certain molecules such as H_2 are described as being held together by *exchange forces*. An H_2 molecule consists of two H^1 nuclei and two electrons. At a certain instant of time both electrons are associated with one nucleus making it a negative ion. The other nucleus is stripped so that it is a positive ion. The negative and positive ions attract each other and so the molecule holds together. At a later instant of time the first nucleus is stripped and the second with its two electrons constitutes a negative ion. Again we have attraction and the molecule holds together. The electrons are exchanged or shared between the two nuclei. Such forces are called exchange forces. In 1934 Fermi applied this principle to the forces holding the protons and neutrons together in a nucleus. He thought of an electron being exchanged between a neutron and a proton so that a nucleon (see Sec.

22.14) would at one instant be a neutron and at another a proton. However, the forces so calculated were too weak to hold the nucleus together. In 1935 Yukawa of Japan made a theoretical investigation of this problem and found that if a particle with a charge equal in numerical value to that of an electron but with a mass about 200 times that of an electron were the exchange particle, the calculated exchange forces would be sufficient to hold the nucleus together. These exchange particles later became known as *mesotrons* or *mesons*. In 1937 cosmic-ray cloud-track pictures giving further evidence for the existence of charged particles of mass between that of an electron and that of a proton were obtained by Anderson and Neddermeyer at Pasadena and also by Street and Stevenson at Harvard. In 1938 Yukawa suggested that these particles might be the abnormal particles occasionally observed in cloud-chamber pictures. However, Yukawa found it necessary to endow these strange particles with the kind of spontaneous decay that is exhibited by radioactive nuclei. The mesotron, which may be either positively or negatively charged, has a half life of about 2 microseconds. However, the mesotron as observed in cosmic rays is moving with a high speed relative to the observer. The half life of 2 μ sec is that measured on a clock traveling with the mesotron; the half life as measured by an observer on the earth is increased by the factor $1/\sqrt{1 - v^2/c^2}$, where v is the velocity of the mesotron relative to the earth.

A mesotron decays with the production of a β-particle or electron. When a radioactive nucleus such as $_{88}\text{Ra}^{226}$ decays, an α-particle is ejected and the nucleus of $_{86}\text{Rn}^{222}$ recoils. Consequently, momentum can be conserved. When a mesotron decays with the ejection of an ordinary electron, there seems to be nothing to recoil and so momentum cannot be conserved. But if there is another particle, momentum can be conserved. Mesotrons decay far from any nucleus and so we cannot bring a nucleus into play as we do when a photon produces an electron pair. Fortunately another particle had been postulated by Pauli, then of Germany, for another purpose. This particle is called the *neutrino*, the little neutron. It carries no charge and has a mass many times smaller than that of an electron. It seems to be unobservable and is sometimes called a phantom particle. If a mesotron decays into an electron and a neutrino, momentum can be conserved, but momentum can also be conserved if a mesotron decays into an electron and a photon.

Physicists do not care to invent unobservable phantom particles.

The only justification for doing so is that these phantom particles serve more than one purpose and that they enable the physicist to refrain from violating a fundamental principle like that of the conservation of momentum. But the neutrino serves more than one purpose. When an α-particle is ejected by a Ra nucleus, the α-particle always has the same initial energy of 4.79 mev, but when a β-particle is ejected by a $_{83}RaE^{210}$ nucleus, the β-particle has an energy anywhere between zero and a certain limiting value of 1.7 mev. This is very strange. Pauli therefore assumed that two particles—a β-particle and a neutrino—are ejected by the RaE nucleus. In this way he was able to explain what is called beta decay.

Still, although the mesotron is not a phantom particle, the physicists wished to make a test which would indicate some property of the mesotron not possessed by any other particle found in cosmic rays. Obviously the test to make is one concerning the half life of a mesotron. A mesotron produces ionization as it passes atoms and so we have absorption of its energy as it passes through air or lead. A layer of air 0.6 mi thick and a sheet of lead 10 cm thick are equivalent so far as the production of ionization is concerned. However, the observed "absorption" is greater in the air than in the lead. This is caused by the fact that more mesotrons decay when they pass through the air than when they pass through lead. A mesotron traveling with a speed $0.5\ c$ takes $6.4\ \mu$ sec to pass through 0.6 mi of air. From various experiments of this type a half life of between 2 and $3\ \mu$ sec has been estimated.

The mesotrons must be secondary particles because of their half life. They seem to be produced high in the atmosphere. Mesotrons constitute the hard component of cosmic rays found by Auger, Rossi, and others. The electrons produced by the decay of mesotrons constitute the soft component at sea level. They could not possibly have penetrated the atmosphere themselves, whereas a mesotron can penetrate the atmosphere. In 1946 it was believed that mesotrons neither produce showers nor give rise to *Bremstrahlung*. Their energy is dissipated by ionization and their numbers by decay as they pass downward through the air. As of 1946 the production of mesotrons is the result of an unknown process. Also as of 1946 it is the general belief that their rest-mass is variable, perhaps between 100 and 250, although some physicists hold the view that the mass is fixed and has a value about 200 times that of an electron.

22.17. Auger Showers.—Provided it does not decay, a mesotron loses energy only by ionization. The chance of a passing charged particle ionizing an atom is a decreasing function of its speed. We may look at the matter as follows: Consider a negatively charged particle which is passing near to an atom. One of the orbital electrons will be closest to the particle. There is a repulsive force between this electron and the particle. In mechanics we have learned that the impulse of a force F is Ft where t is the time during which the force acts. The impulse which is given to the electron depends on the time the particle stays near the electron. But this time is inversely proportional to the velocity v of the particle. Therefore, a slow particle should in passing the atom on the average give more impulse and therefore more momentum to the orbital electron. Consequently more energy is given to the electron, and, if this energy is greater than the binding energy of the electron, the atom becomes ionized. Ionization is thus more likely with slow than with fast particles. From (13–39) the speed of a particle of rest-mass m_0 and kinetic energy W is

$$v = c\sqrt{1 - 1/(1 + W/m_0c^2)^2} \tag{22–14}$$

It is seen that for high energy particles the speed approaches c and so all such particles stay the same time near an orbital electron and the chance of ionization of an atom becomes independent of the speed of the particle. A fast mesotron produces about 2×10^6 ion-pairs per foot of water. This means an energy loss of about 7×10^7 ev/ft. The whole atmosphere is equivalent to 34 ft of water (the height of the water barometer). The mesotrons which get through the atmosphere from the top layers have an energy of $34 \times 7 \times 10^7 = 2.4 \times 10^9$ ev. However, according to Auger, mesotrons have been detected at depths equivalent to 1 mi of water-equivalent. These mesotrons must have an energy of $5280 \times 7 \times 10^7 = 3.7 \times 10^{11}$ ev. Furthermore, there seems to be no limit. Particles with an energy of one million mev are possible.

In 1938 Auger and Maze began looking for showers each of which was distributed over a large horizontal area. They used the coincidence counter technique. They used two Geiger counters with their axes parallel to each other but in the same horizontal plane. They placed no shower-producing material such as lead above the counters. The counters were placed some 3 or 4 meters apart and coincidences were found. They were still found at a separation of 20 meters at sea level. Auger and Maze took their apparatus to the Swiss glaciers at an altitude 2.2 mi. The counters separated by 990 ft give 1 significant coincidence per hour. The word "significant" is used because there are random coincidences. There is a small probability that two particles, having no connection with each other, will pass through two counters within the resolution period during which two particles are recorded as one by a single Geiger counter acting alone. The number of random coincidences per hour can be calculated. The number of coincidences found by Auger and Maze was one more than the random value. With a separation of 990 it was necessary to correct for the time required for a signal to travel along this length of wire. Such evidence showed that a shower had been observed with a horizontal extension of 25 acres. Such showers are known as *extensive* or *Auger showers*. These Auger showers originate very high in the atmosphere. There is evidence that in such a shower there are 25 particles per square yard. In 25 acres there would be 1.21×10^5 particles arriving at a level 2.2 mi high in the atmosphere. Auger estimates the energy of the particle which initiated the shower at 10^{15} ev. As Auger says (translation by Shapiro), the sky is the limit!

The desire of physicists to learn more about cosmic rays lures them to the tops of the highest mountains—the Andes in South America, the Caucasus in Russia, and the Himalayas in India. However, these expeditions are not always feasible and it has become necessary to use airplanes or balloons. But airplanes are limited to a height of 9 mi and so physicists have used unmanned balloons. These carry either recording instruments or radio broadcasting devices which transmit the records to receiving devices on the ground. More recently rockets have been used. The exploration of the upper atmosphere is a challenge to the imagination and ingenuity of the physicists of 1946.

22.18. The Betatron.—The cyclotron cannot be used for the acceleration of electrons because, according to relativity theory and to fact, the mass of an electron increases with its velocity. As a result the equations of Sec. 20.6 do not hold. It is therefore necessary to use some principle of physics different from the principles used in Sec. 20.6.

A device for obtaining electrons of high energies was developed by Kerst at the University of Illinois in 1940. This device is known as a betatron. The action of the betatron depends upon well-known principles of physics. Let us consider a circle with its plane perpendicular to a magnetic field. If the number of lines of force—that is, the magnetic flux—of the field passing through the circle is changing, Faraday's law of electromagnetic induction (see Secs. 4.3 and 5.7) requires that the induced e.m.f. around the circle be given by

$$E = -dN/dt \qquad (22\text{–}15)$$

where N is the flux. Let us further consider the case where the flux density or magnetic field strength is greater at the center than at the circumference of the circle, so that at an instant t the average magnetic field strength H_{ave} at points within the circle is greater than the magnetic field strength H_{orb} on the circular orbit. The flux through the circle at the instant t is then $\pi r^2 H_{ave}$, so that from (22–15)

$$E = \pi r^2 d(H_{ave})/dt \qquad (22\text{–}16)$$

where the minus sign has been omitted. The electric field R on the circle is then $E/2\pi r = (r/2)d(H_{ave})/dt$. If an electron from a hot filament is constrained by some means to move in the circle, the force acting on the electron in a direction tangential to the circle is $(er/2) \cdot d(H_{ave})/dt$. From Newton's second law of motion the rate of change of momentum of the electron is then given by

$$d(mv)/dt = (er/2) \cdot d(H_{ave})/dt \qquad (22\text{–}17)$$

But the electron is moving in a magnetic field H_{orb}, which is at right angles to the motion of the electron. Consequently, it is acted upon by a force evH_{orb} at right angles both to the motion of the electron and the magnetic field. If this force is directed toward the center of the circle and if the centrifugal force mv^2/r equals the force toward the center, so that

$$mv^2/r = evH_{orb} \qquad (22\text{–}18)$$

the electron will be constrained to move in the circle of radius r. Canceling v from (22–18) and differentiating with respect to the time, we obtain

$$d(mv)/dt = e\, d(rH_{orb})/dt \qquad (22\text{–}19)$$

Equating the right sides of (22–17) and (22–19) and canceling e, we have

$$(r/2)\, d(H_{ave})/dt = d(rH_{orb})/dt = r\, d(H_{orb})/dt + H_{orb}\, dr/dt \qquad (22\text{–}20)$$

If the circular orbit is stable, so that $r =$ constant and $dr/dt = 0$, (22–20) becomes

$$d(H_{orb})/dt = (\tfrac{1}{2})d(H_{ave})/dt \qquad (22\text{–}21)$$

The condition for a stable circular orbit of constant radius r is therefore that ratio of the time rate of increase of H_{orb} to the time rate of increase of H_{ave} is 1 : 2. This is known as the "1 : 2" condition. If $H_{orb} = 0$ when $H_{ave} = 0$, the condition is sometimes given as the ratio $H_{orb} : H_{ave} = 1 : 2$.

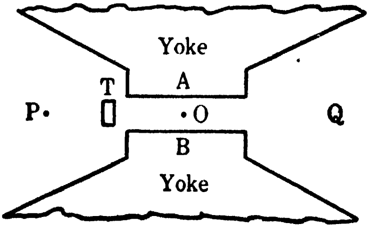


FIG. 22.20.—Diagram of Kerst's betatron.

The magnetic field is produced by the electromagnet shown diagrammatically in Fig. 22.20. The stronger average magnetic field within the circular orbit is brought about by tapering the poles of the electromagnet as shown in the figure. The electron is injected into an evacuated space at P, then finds a stable orbit and finally hits the target T. The exciting current for the electromagnet has a frequency in the range of 60 to 600 cycles per second. The momentum of the electron increases during one quarter of a cycle (not cycle per second). For most of this interval the 1 : 2 condition holds. However, the electron after a trip around its orbit must not hit the filament and so for the first few trips the electron must spiral inward or outward. While it is spiraling, the 1 : 2 condition cannot hold. After the electron has made many trips around the stable circular orbit it must spiral out of this stable orbit so as to hit the target. Consequently the 1 : 2 condition must break down during short intervals respectively at the beginning and end of a quarter cycle. To accomplish this during the later interval the pole pieces and their windings are so designed that some parts of the iron reach magnetic saturation before other parts. One way of upsetting the 2 : 1 condition during earlier interval is to add to the turns carrying the alternating current a few back turns about the pole tips carrying a steady direct current. From (22–20), we see that the sign of dr/dt depends on the relative values of $d(H_{orb})/dt$ and $d(H_{ave})/dt$ and on the sign of H_{orb}. The electron spirals outward or inward according as dr/dt is positive or negative. The hot filament, stable orbit, and the target are contained in a glass envelope (not shown in Fig. 20.20) whose shape is that of the surface of a doughnut. The axis through the center of the "hole" of the doughnut coincides with the axis of the poles of the magnet.

If H_{orb} increases from nearly zero to H_{max} while the 1 : 2 condition holds, (22–19) shows that the momentum of the electron increases by erH_{max}. For $r = 10$ cm and $H_{max} = 1000$ emu, the increase of momentum is 1.6×10^{-16} gm cm sec^{-1}. Since $m = m_0/\sqrt{1 - v^2/c^2}$, this corresponds to a velocity of 0.9857 times that of light and an energy of 2.52 mev. This is approximately the energy obtained by Kerst in 1941. For most of the quarter cycle the electron is traveling with a speed nearly equal to that of light, so that for a frequency of 600 cycles per second and $r = 10$ cm the electron makes about 200,000 trips around its stable orbit before hitting the target. It was reported in 1946 that 100 mev electrons had been obtained by means of a betatron. On hitting the tungsten target T of Fig. 22.20, such electrons produce x-rays of minimum wave-length $12{,}345/10^8 = 1.23 \times 10^{-4}$ A. These x-rays are in the region of gamma rays and are very penetrating.

CHAPTER XXII

PROBLEMS

1. If electrons of kinetic energy 5.4×10^3 mev are approaching the earth from all directions in space, at what latitude will these electrons fail to reach the surface of the earth?

2. Find the kinetic energies of protons for which x in Fig. 22.6 equals 0.4. These protons arrive from all directions at magnetic latitudes between λ_1 and 90°. What is the value of λ_1? Find the value of x for α-particles of kinetic energy 1.5×10^9 ev.

3. Determine the number of ion-pairs produced in air by an α-particle whose kinetic energy is 21 mev. Calculate the range of this particle in air at a pressure of 4 atmospheres and a temperature of 30° C.

4. The radius of curvature of the path of a charged particle in a magnetic field decreases as the particle slows down. Let r be the radius of curvature at the beginning of the path of a particle of mass m carrying a charge $\pm 4.8 \times 10^{-10}$ esu in a magnetic field of strength H. Assuming that the constant in the range-energy relation (20–7) is proportional to $m^{1/3}$, derive the relation $m^2 = KH^3r^3T/(Rp)$, where R is the range in air at a pressure p and a temperature T and K is a numerical constant. Determine the value of K when m is measured in atomic mass units ($O^{16} = 16$), R and r in cm, H in gauss (emu), p in atmospheres (76 cm of mercury = 1 atmosphere), and T in degrees absolute. Neglect relativity corrections.

5. Using the relation found in Problem 4, determine the mass in atomic mass units of a particle producing a path of initial radius of curvature 9 cm in a magnetic field of 6000 gauss and having a range of 2 cm in air at a pressure of 3 atmospheres at a temperature of 20° C. Give the ratio of the mass of the particle to the rest-mass of the electron.

CHAPTER XXIII

ASTROPHYSICS

23.1. Cavendish Experiment. Newton's law of gravitation may be written

$$F = \frac{GMm}{r^2} \tag{23-1}$$

where F is the force between two point masses M and m which are at a distance apart equal to r and G is a constant of proportionality known as the gravitational constant. When a planet revolves about a sun (or a moon about a planet) in a circular orbit the force of attraction just balances the centrifugal force. However, it is not quite true that the planet revolves about the sun but rather both the sun and the planet revolve about their common center of gravity. If a is the distance of the center of the sun of mass M from the common center of gravity and V is the velocity of the sun in its orbit, which we shall for simplicity suppose to be circular, and b, m and v are the corresponding quantities for the planet in its orbit about the common center of gravity, we have for the sun and planet respectively

$$\frac{MV^2}{a} = \frac{GMm}{d^2} \quad \text{and} \quad \frac{mv^2}{b} = \frac{GMm}{d^2} \tag{23-2}$$

where $d = a + b$, the distance of the center of the planet from the center of the sun. However, the period of the sun in its orbit is the same as that of the planet in its orbit. If this common period is P, then $V = 2\pi a/P$ and $v = 2\pi b/P$, so that $Ma = mb$ from (23–2). This must be so because a and b are measured from the common center of gravity. Since $Ma = mb$ and $d = a + b$, we have $a = md/(M + m)$ and $b = Md/(M + m)$. Either of equations (23–2) therefore leads to

$$\frac{4\pi^2 Mmd}{(M + m)P^2} = \frac{GMm}{d^2}$$

whence, on canceling Mm,

$$G(M + m) = \frac{4\pi^2 d^3}{P^2} \tag{23-3}$$

The quantities d and P can be found experimentally. Hence, the product of the gravitational constant and the sum of the masses of the two heavenly bodies which are rotating about each other can be found. Further, if the distances a and b can also be measured, the values of GM and Gm can be obtained, since $M = b(M + m)/d$ and $m = a(M + m)/d$. However, although the products $G(M + m)$, GM and Gm for the sun, planets and moons of the solar system can be measured, yet G and the masses $(M + m)$, M and m cannot themselves be determined. It is thus impossible to find the value of the gravitational constant G from astronomical observations alone, and it is necessary to devise a terrestrial experiment in order to determine G.

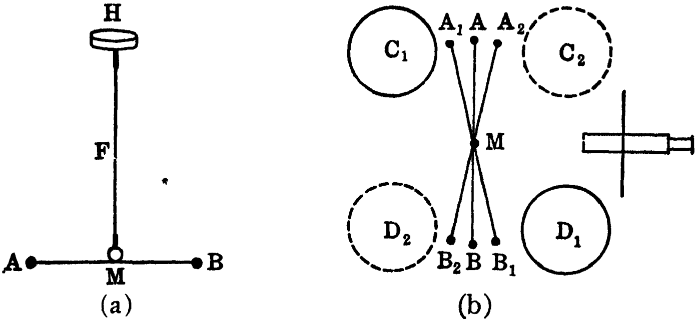


FIG. 23.1.—Experimental arrangement for the Cavendish experiment.

In the experiment as devised in 1781 by Cavendish of Cambridge, England, a light horizontal rod AB supporting two small gold balls at its ends is suspended by a fine quartz fiber as shown in Fig. 23.1a. The upper end of the fiber is attached to a graduated screw head as shown. A mirror M is attached to the rod AB as shown. A plan of the rod is shown at AB in Fig. 23.1b. In Fig. 23.1b, C and D represent massive lead spheres whose centers are on the same horizontal plane as the centers of the gold balls A and B. The lead balls C and D are supported on a wooden carriage which runs on rollers. The distance apart of the centers of C and D equals that of the centers of A and B. The lead balls C and D can be moved by means of the carriage from the position C_1D_1 to the position C_2D_2, or they can be removed altogether. When the lead balls C and D are entirely removed the position of the light rod is AB as in Fig. 23.1b. When the lead balls are brought to the position C_1D_1 the ball C attracts the ball A and similarly the ball

D attracts the ball B and the rod AB is pulled over into the position A_1B_1. There is of course an attraction between the lead ball at D_1 and the gold ball at A, but due to the much greater distance between the centers of the ball at D_1 and of the ball at A we shall neglect this force. Likewise we shall neglect the force between the lead ball at C_1 and the gold ball at B. As the rod turns from AB to A_1B_1, the mirror on AB is turned. By means of a telescope, lamp and scale the amount of turning may be determined. Instead, however of determining the amount of turning, the screw head in Fig. 23.1a is turned and the fiber twisted until the reading through the telescope shows that the rod is back in the position, AB, Fig. 23.1b. The angle through which the screw head has been turned is read. Let this angle be θ. The twist of the fiber is therefore θ. According to Hooke's law the torque necessary to twist a fiber or wire through an angle θ is proportional to θ, and in this case the torque is $k\theta$, where k is the torsional constant of the fiber. The distance between the center of the lead ball at C_1 and the center of the gold ball at A is $2l \sin \alpha$, where $2l$ is the length of $AB(= C_1D_1)$ and 2α is the angle AMC_1, it being supposed that the line C_1D_1 passes through M, Fig. 23.1b. Hence the force of gravitational attraction between the lead ball at C_1 and the gold ball at A is

$$\frac{GMm}{4l^2 \sin^2 \alpha} \tag{23-4}$$

where M is the mass of a lead ball and m that of a gold ball. There is an equal force between the lead ball at D_1 and the gold ball at B. These two forces constitute a couple or torque acting upon the rod AB whose magnitude is

$$\frac{GMm}{4l^2 \sin^2 \alpha} \times 2l \cos \alpha, \tag{23-5}$$

as is seen by reference to Fig. 23.2. When equilibrium is established, this torque must equal the opposite torque due to the twist in the fiber, so that

$$\frac{GMm \cos \alpha}{2l \sin^2 \alpha} = k\theta \tag{23-6}$$

All the quantites in (23–6) can be measured with the exception of G and k. The lead balls may be swung into the position C_2D_2 in order to obtain a check upon the measurements.

In order to find G we must first find k. This is accomplished by removing the lead balls and then setting the rod AB into rotational vibration about the fiber F as axis. Let θ be the angle which the rod AB makes at an instant t with its position when the fiber is not twisted. The restoring torque acting upon the rod is then $k\theta$. From (3–8), the period

$$P = 2\pi\sqrt{I/k} \tag{23-7}$$

where I is the moment of inertia of the rod AB and the two gold balls. The period of the oscillations can be found and the moment of inertia of the rod and gold balls can be calculated, so that k can be found from (23–7). Substituting this value of k in (23–6), the value of G can be determined. Its value is

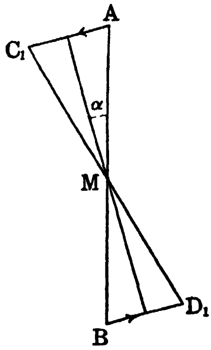


FIG. 23.2.

$$G = 6.66 \times 10^{-8}\ \text{cm}^3\ \text{gm}^{-1}\ \text{sec}^{-2} \tag{23-8}$$

23.2. Masses of the Earth and Sun.—Knowing the value of G, we can immediately determine the mass of the earth. Let the radius of the earth be R so that the force acting upon a mass μ at the surface of the earth is $Gm\mu/R^2$, where m is the mass of the earth. The acceleration produced by this force in the mass μ is g, so that

$$\frac{Gm\mu}{R^2} = \mu g \tag{23-9}$$

whence the mass of the earth is

$$m = \frac{R^2 g}{G} \tag{23-10}$$

Putting $R = 3960$ miles $= 6.37 \times 10^8$ cm, $g = 980$ and $G = 6.66 \times 10^{-8}$, we obtain

$$\text{Mass of the Earth} = 6 \times 10^{27}\ \text{gm} \tag{23-11}$$

From its volume the average density of the earth is found to be 5.5 gm/cm³.

The ratio of the mass of the sun and earth to that of the earth and moon is, in virtue of (23–3), given by

$$\frac{M+m}{m+\mu}=\frac{r_1^3}{r_2^3}\cdot\frac{P_2^2}{P_1^2} \tag{23–12}$$

where M, m and μ are the masses of the sun, earth and moon respectively, r_1 and r_2 are respectively the distance of the center of the earth from the center of the sun and that of the center of the moon from the center of the earth (all orbits being supposed circular), and P_1 and P_2 are respectively the periods of the earth about the sun and of the moon about the earth. Putting in the values $r_1 = 9.3 \times 10^7$ miles, $r_2 = 240{,}000$ miles, $P_1 = 365.24$ days and $P_2 = 27.32$ days, we find $(M+m)/(m+\mu) = 325{,}000$. But the mass of the moon is about one-sixtieth that of the earth, so that $(M+m)/m = 325{,}000 \times 61/60 = 330{,}000$. Knowing the mass of the earth, we obtain

$$\text{Mass of the Sun} = 2 \times 10^{33} \text{ gm} \tag{23–13}$$

It is interesting to note that the mass of the earth is 6×10^{27} gm, while the mass of an electron is 0.91×10^{-27} gm, so that the number of electrons in a gram of electrons is of the same order of magnitude as the number of grams in the mass of the earth.

23.3. The Solar System.—The sun is surrounded by the planets which revolve about it. The orbits of these planets although not exactly in the same plane are fairly close to the plane of the ecliptic which is the plane of the earth's orbit about the sun. Furthermore, the planets all revolve about the sun in the same sense—that is, if the earth as viewed from a point on a perpendicular drawn through the sun to the plane of the earth's orbit revolves in a clockwise direction about the sun, the other planets as viewed from this same point also revolve in the clockwise direction. The sun and the planets rotate about their respective axes and these rotations are in the same sense as the revolutions. Again, the moon revolves about the earth in the same sense as the rotation of the earth about its own axis and the same is true of the revolutions of most of the other satellites about their respective planets. Our own moon rotates about its own axis as it revolves about the earth in such a way that it rotates once during the same time that it revolves once about the earth. This agreement of the period of rotation of the moon with the period of its revolution about the earth causes the moon always to present the same face to the earth. The earth's moon consequently rotates in the same sense as the earth and the sun. Thus it is seen to be the general rule that the bodies of the solar system revolve and rotate in the same sense.

The planets Mercury and Venus which are closest to the sun are not accompanied by moons, the earth has one moon, Mars two, Jupiter eleven, Saturn ten, Uranus four, and Neptune one moon. Between Mars and Jupiter there are thousands of small bodies which revolve about the sun. These small bodies are known as asteroids and are believed to be the fragments of a planet which formerly revolved about the sun in an orbit between the orbits of Jupiter and Mars and which has since been broken up. The relative distance of each planet from the sun is shown in Table 23.1. The distance of the earth from the sun is taken as 10.

TABLE 23.1

DISTANCES OF THE PLANETS FROM THE SUN

	Actual distance	Mass compared to the earth
Mercury	3.9	0.04
Venus	7.2	0.81
Earth	10.0	1.00
Mars	15.2	0.11
Asteroids	26.5	....
Jupiter	52.0	317
Saturn	95.4	95
Uranus	191.9	14.7
Neptune	300.7	17.2
Pluto	398.0	....
Sun	....	330,000

The distance of the asteroids is the average distance. Ceres, the largest asteroid, has a diameter of 480 miles, while only four asteriods have diameters of more than 100 miles. The asteroids began to be discovered about 1775. In 1781 Herschel discovered the planet Uranus.

23.4. Irregularities in the Solar System.—Kepler stated that the orbits of the planets are ellipses with the sun at one focus, and Newton showed that these elliptic orbits are a consequence of the law of gravitation, which states that the force of attraction between two bodies varies inversely as the square of the distance between the bodies. However, on very careful examination of the motion of the planets it is found that, although the motion of a planet is very nearly along an elliptical path, yet there are small discrepancies between the actual motion of the planet and its theoretical motion according to Kelper's laws. These discrepancies are known as perturbations. The perturbations of the motion of a planet are due to the small but not negligible attractions of all the other planets for the planet in question. We have mentioned in Sec. 16.10 that the general solution of the many body problem of the solar system has never been obtained. Fortunately, although an exact solution for the solar system is not possible, an approximate solution can be found in the particular case when one of the bodies is very much heavier than the remaining bodies. In the case of the solar system the sun weighs 330,000 times as much as the earth, while all the planets taken together weigh only about 460 times as much as the earth, so that an approximate solution can be obtained for the orbit of any planet.

About a century ago it was found that the motion of Uranus was not given exactly by the gravitational pulls of the sun and the other planets then known. It began to be suspected that there might be another planet which was perturbing the motion of Uranus. The position of this theoretical planet was worked out independently in 1846 by Leverrier, a Frenchman, and by Adams, an Englishman. Leverrier communicated his calculations to Galle in Berlin. Galle immediately found the new planet in the part of the sky indicated by Leverrier's calculations and the new planet was named Neptune.

Recently the situation with regard to the discovery of Neptune has repeated itself. Further study of the motion of Uranus showed that there was still a discrepancy between its actual motion and the theoretical motion when the effect of the pull of the sun and the pulls of the other planets, including Neptune, had been taken into account. Lowell, an American, calculated that these perturbations of the motion of Uranus could be explained by assuming the existence of another planet still further out than Neptune. In 1930 the planet was discovered fairly close to the position predicted by Lowell. The name of the new planet is Pluto. This planet is about forty times as far from the sun as the earth and it is so far out in space that it takes about 250 years to complete one revolution about the sun.

23.5. Advance of the Perihelion of Mecury.—To a first approximation the planets revolve about the sun in ellipses. In the case of the earth the ellipse is almost a circle, but in the case of Mercury the orbit is much more elliptical. In an elliptical orbit the sun is at one focus as at S, Fig. 23.3a. The point P where the planet approaches most closely to the sun is called the perihelion, while the point A where the planet is farthest from the sun is called the aphelion. The line *AP* is the major axis of the ellipse. According to Newton's theory of gravitation, the major axis of the orbit of a single planet revolving about a sun remains fixed in space. As we have seen in the previous section the solar system contains several planets. The result of the attractions between the various planets is that the major axis of the orbit of each planet does not remain fixed in space, but rotates slowly so that a planet which is revolving in an orbit such as shown in Fig. 23.3a will after many revolutions be revolving in an orbit such as shown in Fig. 23.3b. This slow rotation of the axis is described by the phrase "advance of the perihelion." In the case of the planet Mercury its perihelion advances at the rate of 574 seconds of arc per century. This rate of advance of the perihelion of Mercury having been established, the astronomers set to work to calculate the perturbations due to the other planets. When the pulls of the other planets were taken into account a rate of advance of the perihelion of 532 seconds per century was accounted for. This leaves a discrepancy of 574 − 532 = 42 seconds of arc per century unaccounted for. Following the historical custom established by the discovery of Neptune, a new planet was predicted. Since of the known planets Mercury is the planet nearest the sun, the only place for the new planet is still nearer the sun. The position of the new planet was calculated, and it was even named Vulcan. However, Vulcan has steadfastly refused to be discovered and astronomers have been forced to search for another explanation of the 42 seconds of arc per century which are still outstanding. As we shall see later this explanation is given by the General Relativity Theory.

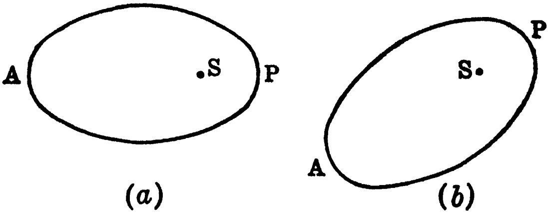


FIG. 23.3.—Planet first revolves in orbit (a). The major axis slowly turns so that after many revolutions the orbit is as in (b).

23.6. The Sun.—In Sec. 23.2 we described how the mass of the earth is obtained and how from the mass of the earth the mass of the sun is found. The mass of the sun is 2×10^{33} gm or 2.2×10^{27} tons. The so-called fixed stars in the sky are also suns. It is found that many of these stars occur in pairs which appear to revolve about each other. We have shown in Sec. 23.1 that, if the distance d of one component of the double star from the other can be measured and the period of revolution P can be found, the sum of the mass of the two components of the double star as it is called can be found from the formula

$$M + m = \frac{4\pi^2 d^3}{GP^2} \tag{23–14}$$

where G is the gravitational constant. By this method the masses of many stars have been calculated and it has been found that our sun has about the mass of the average star. Further, it has been found that the masses of other stars do not differ greatly from that of the sun. It is very rare to find a star whose mass is more than ten times that of the sun and no star is known whose mass is less than one-tenth that of the sun. A few stars approach a mass which is 250 times that of the sun, that is, a mass of 5×10^{35} gm or 5.5×10^{29} tons. It is interesting to note that a man weighting 150 lbs has a mass which is 7.5×10^{31} times the mass of the electron which is the lightest body known, and that one of these largest stars or suns has a mass which is 0.73×10^{31} times the mass of a man. Thus it is seen that on a logarithmic scale man stands about midway in mass between the mass of the lightest and that of the heaviest bodies in the universe.

The sun is a vast source of energy. From the heat given out by the sun it has been calculated that every square centimeter of its surface radiates energy at the rate of about 6×10^{10} erg/sec or 50 horsepower for each square inch of surface. The sun's atmosphere contains not only the light gases which constitute the earth's atmosphere but in addition the vapors of heavy metals such as platinum and lead. The temperature of the sun's atmosphere is about 5500° C. However, this is only the surface temperature. At points within the sun the pressure is tremendous. The radius of the sun is 433,000 miles or 7×10^{10} cm. Since the mass of the sun is 2×10^{33} gm the weight of a mass on the surface of the sun is 28.2 times the weight of the same mass on the surface of the earth. It is supposed that the interior of the sun is gaseous throughout. The temperature of the gas in the interior must be very great in order that the gas may not be condensed by the tremendous pressure. It has been estimated by Eddington and Jeans that the temperature at the center of the sun must be of the order of 40,000,000° C. In Chapter VIII on the Kinetic Theory of Gases we learned that the root mean square velocity of the nitrogen molecules of the air at 0° C is 49,300 cm/sec or 0.31 miles/sec. At a temperature of 40,000,000° C the speed would be 1.89×10^7 cm/sec or 117 miles/sec, since the root mean square speed varies as the square root of the temperature. Although high, this speed is by no means as great as the speed of α-particles and is only a little more than six times as great as the speed of the earth in its orbit about the sun. However, it is very unlikely that diatomic molecules exist at a temperature of 40,000,000° C. It is much more likely that the molecules are all monatomic and the speed of oxygen or nitrogen atoms would be about $\sqrt{2}$ times the above speed, or 2.67×10^7 cm/sec.

The thermionic effect is very great at a temperature of 40,000,000° C, and hence, in addition to the atoms rushing to and fro, free electrons also are rushing about with velocities such that the average kinetic energy of an electron is equal to the average of any of the atoms. Since the mass of a nitrogen atom is $14 \times 1845 = 25{,}800$ times that of an electron, the speed of the electrons at 40,000,000° C is $2.67 \times \sqrt{25{,}800} \times 10^7 = 4.29 \times 10^9$ cm/sec. This is the speed given to an electron by a fall through 5150 volts. Since electrons are rushing about with a speed of 5150 volts and since the ionization potentials for the outermost electrons of atoms are much less than 5150 volts, it is probable that the atoms are largely stripped of their electrons. The nitrogen and oxygen atoms are probably stripped of the *L* electrons and a few of the atoms may also be stripped of their *K* electrons, leaving only the nuclei. Atoms with *K* and *L* electrons are constantly being formed but almost as soon as they are formed they are again ionized by an impinging electron which hits the newly formed atom with a speed of 5150 volts.

Not only are stripped and partially stripped atoms and electrons rushing about at this high temperature, but also there is a tremendous energy in the heat radiation. We know that as the temperature of a body is increased from room temperature it first appears to be a dull red, then as it becomes hotter it glows with a bright red heat, then becomes white hot and so on. The temperature of white heat is about 1500° C. At this temperature the radiation contains all the wavelengths of the visible spectrum. If the spectrum of the radiation given out by a body is examined, and the intensity of radiation in a wave-length range $d\lambda$ for various values of the wave-length λ is determined, it is found that the intensity is a maximum for a certain wave-length and that the value of this wave-length depends upon the temperature. Wien's displacement law states that the relation between this wave-length of maximum intensity which is designated by λ_m and the absolute temperature T is

$$\lambda_m T = \text{constant} \tag{23–15}$$

If λ_m is measured in angstroms, the value of the constant is 2.88×10^7. The radiation which reaches us from the surface of the sun has a maximum intensity in the yellow or at about $\lambda_m = 5000$ angstroms, so that the temperature of the surface of the sun is about 5750° C according to this law. At a temperature of 40,000,000° C, the wave-length of maximum intensity in the heat radiation is $(2.88 \times 10^7)/(4 \times 10^7) = .72$ angstrom. Hence at the temperature which exists at the center of the sun the heat radiations consist mainly of x-rays—instead of being white hot the center of the sun is "x-ray hot." The x-rays aid the free electrons in ionizing the atoms and stripping them of their electrons. At the center of the sun there are partially stripped atoms, nuclei, electrons and x-ray quanta all rushing about and intermingling. The impacts of all these support the enormous pressure due to the weight of the matter from the center of the sun outwards.

We have spoken of the pressure of light in Sec. 5.1. At the center of the sun the temperature radiation exerts an enormous pressure outwards and tends to cause the sun to expand. We have previously mentioned that the stars all seem to have masses which are not far removed from that of the sun. This is explained by the theory that the temperature at the center of a star is determined in part by the mass of the star. As the mass increases the temperature at the center increases. However, as the temperature increases the pressure of the radiation increases and there comes a situation where the outward pressure of the radiation is greater than the inward pressure due to the gravitational weight of the outer layers. Stars with a mass greater than a certain critical value cannot be formed.

The x-rays which exist at the center of the sun or a star gradually penetrate outwards. They are absorbed by atoms and are re-emitted. Also they make elastic collisions with electrons and have their energies reduced by the Compton effect. As the radiation proceeds outward from the center, the temperature becomes less and as a result the radiation which is emitted by the matter at higher levels has a longer wavelength. Finally some of the radiation reaches the surface of the sun or star but now corresponds to the temperature of the surface, which in the case of the sun is about 5500° C.

23.7. Temperature at the Center of the Sun.—In order to estimate the temperature at the center of the sun, we shall for simplicity neglect the radiation pressure in the interior of the sun. Consequently, the estimate will be only a rough approximation. Consider a ball of fluid as in Fig. 23.4. The ball has a radius a and mass M. It can be shown that the force of gravity on a mass of one gram at a distance r from the center of the ball is that due to the mass inside the sphere of radius r, the shell between $r = r$ and $r = a$ having no effect. The mass of the sphere of radius r is Mr^3/a^3. We shall make the simplifying assumption that the density ρ is a constant throughout the fluid. There is some justification for this assumption, since the density of the whole sun is 1.41 gm/cm³ and this is about the density of solid substances at ordinary pressures and temperatures. Consider the shell of thickness dr. Unit area of this shell has a mass $\rho\, dr$ and the force of attraction on this unit area of the shell to the center of the sphere is

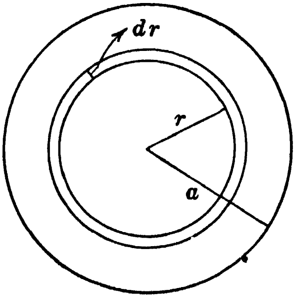


FIG. 23.4.

$$(G \times \rho\, dr \times Mr^3/a^3)/r^2 = GM\rho r dr/a^3 \qquad (23\text{–}16)$$

where G is the gravitational constant. Since this force is applied to unit area of the boundary of the fluid inside the sphere of radius r, it is the pressure developed in the sphere by the attraction of the mass of fluid on the spherical shell dr. If now we consider the ball to be made up of successive shells, each of which contributes to the total pressure at the center, we obtain

$$p = \int_0^a GM\rho r dr/a^3 = GM\rho/2a \qquad (23\text{–}17)$$

for the pressure at the center, where $r = 0$. Since ρ = mass/volume = $M/(4/3)\pi a^3$, this pressure is

$$p = 3GM^2/8\pi a^4 \qquad (23\text{–}18)$$

Let W be the average atomic weight of the atoms of the rocks and soil at the surface of the earth and let Z be the average atomic number, then there are $(Z + 1)$ particles (nucleus plus electrons) for each average atom. Since hydrogen is very scarce, the value of $W/(Z + 1)$ is about 2. We next assume that this also applies to the sun. The average atomic weight per particle of the sun is then 2, so that the average mass per particle is $2 \times 1.66 \times 10^{-24} = 3.32 \times 10^{-24}$ gm. The number of particles in 1 cm³ of mass ρ is therefore $\rho/(3.32 \times 10^{-24})$. This is the value of n in the kinetic theory formula

$$pv = nkT \qquad (23\text{–}19)$$

Putting $v = 1$ cm³, $G = 6.66 \times 10^{-8}$, and eliminating p from (23–17) and (23–19), we obtain

$$M/a = 1.24 \times 10^{15} \times T \qquad (23\text{–}20)$$

For the sun $M = 2 \times 10^{33}$ gm and $a = 433{,}000$ miles, whence at the center $T = 23{,}500{,}000°$ Abs. Although this value is based upon the simplifying assumptions that the density is the same throughout the ball and that the radiation pressure may be neglected, it agrees in order of magnitude with Eddington's estimated value of 40,000,000° Abs given in Sec. 23.6. Estimates more recent than Eddington's make the temperature at the center of the sun about 20,000,000° Abs. It is sometimes stated that the temperature of exploding U^{235} plus the tamper is about 10 times the temperature at the center of the sun. For the earth $M = 6 \times 10^{27}$ gm and $a = 3956$ miles, whence at the center $T = 7020°$ Abs. The temperature at the center of the earth is therefore not much above the temperature at the surface of the sun.

23.8. Distances of the Stars.—The mean distance of the sun from the earth is 92,870,000 miles. The diameter of the earth's orbit is therefore about 186,000,000 miles. When the nearer stars are viewed first from one end of a diameter of the earth's orbit and six months later from the other end of the diameter it is noticed that the positions of certain stars relative to the rest of the stars change slightly. This is because these certain stars are nearer to us—the nearer stars appear to move against the background of more distant stars when we change our position of observation by 186,000,000 miles. The nearest star, Proxima Centauri, changes its position relative to the background of very distant stars by about 1.5 seconds of arc when viewed from opposite ends of a diameter of the earth's orbit. Half of this change of position is known as the parallax of the star, so that the parallax of Proxima Centauri is 0.75 second of arc. Since the stars are at such great distances a new unit of distance has come into use—namely, the light-year. A light-year is the distance that light traveling at a speed of 186,000 miles/sec travels in one year. This distance is 5.85×10^{12} miles. The relation between the parallax p of a star and its distance D in light-years is

$$D = \frac{3.26}{p} \tag{23-21}$$

where p is measured in seconds of arc. Another unit of length in astronomical measurements is the parsec. The parsec is the distance of a star whose parallax is 1 second of arc. From (23-21)

$$1 \text{ parsec} = 3.26 \text{ light-years} = 1.91 \times 10^{13} \text{ mi} \tag{23-22}$$

The distance of Proxima Centauri is therefore 4.27 light-years. Sirius, the brightest star, is at a distance of 8.65 light-years. Most of the stars are much further away than Proxima Centauri and Sirius. Many stars are known to have a parallax less than 0.005 second, which makes their distance greater than 600 light-years. Distances of stars can be measured by methods other than by their parallaxes. These other methods show stars to be at distances of 10,000 to 1,000,000 light-years.

23.9. Brightness of Stars.—When speaking of the brightness of a star we must distinguish between "intrinsic" and "apparent" brightness. The amount of light received from a star depends upon its distance from the earth and the amount of light from a star received by a telescope on the earth gives the apparent brightness. The intrinsic brightness or candlepower of a star is obtained by multiplying its apparent brightness by the square of its distance from the earth. Unlike their approximate equality in mass, the stars differ enormously in intrinsic brightness or candlepower. Thus Sirius has a candlepower of 26.3 times that of the sun, while another star called the Companion of Sirius has a candlepower of 0.0026 times that of the sun. The intrinsic brightness of Sirius is thus 10,000 times as great as that of its Companion.

23.10. Surface Temperature of Stars.—Ordinary photographic plates are more sensitive to blue light than to red light. When the sky is photographed certain stars are much more pronounced in the picture than others. The photographic "brightness" does not agree with the brightness as seen by the eye, for the eye is most sensitive to yellow light. Astronomers have come to speak of the color of stars. The stars are of different colors because their surface temperatures are different. Red stars are the coolest and white stars the hottest. An accurate determination of the surface temperature of a star can be made by passing the light from a star into a spectroscope and finding the wave-length whose intensity is a maximum. Wien's displacement law as expressed in (23–15) is applied and the temperature determined. The surface temperatures of different stars range from about 1400° C to 40,000° C.

23.11. Diameters of Stars. –The diameter of a planet can be measured because the planet as observed by a telescope appears as a disc of definite size. Stars, however, appear as points of light. Nevertheless, the diameter of many stars can be calculated. The apparent brightness, color, and distance of a star are determined. From these, its candlepower and temperature can be found. From the candlepower we know the rate at which energy of radiation is pouring out from the surface and from the temperature by applying the Stefan-Boltzmann law we can find the rate at which energy is leaving each square centimeter of the surface. Thus, knowing the total radiation from the whole surface and the radiation from each square centimeter of surface, we can immediately find the diameter of the star. According to Sec. 9.11 the Stefan-Boltzmann law states that the rate at which heat radiation is emitted by unit area of the surface of a black body raised to a temperature T° Abs is given by

$$R = \sigma T^4 \qquad (23\text{–}23)$$

where R is the energy radiated per second by 1 cm^2 of surface and σ is a constant of proportionality. The value of σ is 5.714×10^{-5} erg cm^{-2} deg^{-4} sec^{-1}.

In 1920 Michelson, working at Mount Wilson Observatory, California, and using the methods of interferometry, measured the diameters of several stars. He found that Betelgeuse, a red giant star, subtends an angle of 0.045 seconds of arc at the earth. This corresponds to a diameter of about 240,000,000 miles, which is nearly 300 times the diameter of the sun. Michelson's value is about 10 per cent less than that obtained by the method described in the first paragraph of this section. In view of the difficulties of measurement, this result validates the method of the first paragraph.

23.12. Sirius and Its Companion.—These two stars revolve about each other, the Companion being visible in modern telescopes. From the period of revolution and the distance apart of the two stars, the total mass of the two stars has been found from (23–14) to be 3.27 times the mass of the sun. The two stars revolve about their common center of gravity and this point divides the line joining the centers of the stars in the inverse ratio of their masses. From astronomical measurements it has been found that the common center of gravity divides the line adjoining the centers in the ratio of about 3 to 1 so that mass of Sirius is 3 times that of its Companion. The mass of Sirius is thus $3 \times 3.27/4$ or 2.45 times that of the sun, and the mass of the Companion is $1 \times 3.27/4$ or 0.82 times that of the sun. From its candlepower, color and distance the diameter of Sirius is found to be about 1.58 times that of the sun. This makes its volume nearly 4 times that of the sun, so that its density is $2.45/4 = 0.61$ times the density of the sun, which is 1.4 gm/cm^3. On the other hand, the Companion is of nearly the same color and therefore has about the same surface temperature as Sirius and yet its candlepower is 1/10,000 times that of Sirius. Its diameter is calculated to be about 1/50 times that of Sirius, so that its volume is $1/50^3 = 1/125{,}000$ that of Sirius. The density of the Companion is therefore $(0.82/2.45) \times 125{,}000 = 42{,}000$ times the density of Sirius. Since the density of Sirius is $0.61 \times 1.4 = 0.85$ gm/cm^3, the density of the Companion is $42{,}000 \times 0.85 = 36{,}000$ gm/cm^3. A cubic inch of the matter in the Companion of Sirius has a mass of nearly two-thirds of a ton. The surface temperature of both Sirius and its Companion is about 11,000° C. The Companion of Sirius is thus very hot and very small and is on this account known as a white dwarf star.

In the case of the sun most of the atoms at the center of the sun retain their K electrons and, since the radius of the K shell in the atoms of nitrogen is 0.02 angstrom, the diameter of these atoms in the sun is of the order of 4×10^{-10} cm. In the case of the Companion of Sirius it is believed that the temperature at its center is so high that most if not all of the atoms have become stripped of their K electrons and consist only of nuclei. From Rutherford's experiments on the scattering of α-particles, it is believed that the radius of the nucleus is not larger than about 10^{-12} cm. Also the radius of the electron is about 10^{-13} cm. Ordinary atoms have radii of about 10^{-8} cm, so that electrons and nuclei can be packed much more closely when they are not joined to each other in atoms. If a star consisted only of nuclei and electrons, it could have a density of 10^{12} times that of ordinary matter, so that even the high density of 36,000 gm/cm^2 which is found in the Companion of Sirius is still far short of the possible density of matter consisting of free nuclei and electrons which are packed so as to be in contact.

23.13. Velocities of Stars.—The component of the velocity of a star in the line of sight can be obtained from the Doppler effect, for which the formula is

$$\nu = \nu_0\left(1 + \frac{v}{c}\right) \tag{23-24}$$

where ν_0 is the natural frequency of a spectrum line which is given out by a hot gas and ν is the observed frequency when the gas is moving toward the observer with a velocity v and c is the velocity of light. In terms of wave-length the Doppler formula is

$$\delta\lambda = -\frac{\lambda_0 v}{c} \tag{23-25}$$

where $\delta\lambda$ is the shift of the spectrum line from its natural wave-length λ_0. If v is positive when towards the observer $\delta\lambda$ is negative—that is, the wave-length is smaller and the line is shifted towards the violet end of the spectrum.

When the light from the sun is examined by means of a spectroscope it is found that the continuous spectrum of white light is crossed by numerous relatively dark lines, which are known as Fraunhofer lines. If the white light from an electric arc is sent through a sodium flame and then examined by a spectroscope it is found that the continuous spectrum is crossed by a dark line in the yellow corresponding to the emission spectrum of sodium. The sodium vapor absorbs the light which it under other circumstances emits from the white light and we have what is known as the absorption spectrum of sodium. The dark lines in the spectrum of the sun are absorption lines produced by absorption in the cooler vapors which surround the sun. By making use of the formula for the Doppler effect the equatorial speed of rotation of the sun can be found. If the light from the east limb of the sun and then the light from the west limb of the sun is examined by means of the spectroscope, it is found that there is a shift in the Fraunhofer lines. The west side of the sun gives Fraunhofer lines which are shifted toward the red and so it is receding from us. The speed of rotation is such that its period is about 26 days.

Some stars give bright line spectra while others give dark line spectra. The Doppler shift of either bright line or dark line spectra may be used for determining the velocities of stars in the line of sight. Stellar velocities in the range of 5 to 20 mi/sec are common, while velocities over 60 mi/sec are rare although stellar velocities approaching 250 mi/sec are known.

23.14. Spectroscopic Binaries.—The term binary is applied to two stars such as Sirius and its Companion which are close enough to each other so that they revolve about each other under the force of their mutual gravitational attraction. Many systems of two stars cannot be separated into their components by the telescope as are Sirius and its Companion and their binary nature is suspected only because the Fraunhofer lines of such systems, which are apparently single stars, are doubled. Each Fraunhofer line shows a double shift. If one Fraunhofer line is observed over a period of time it is found that its shift varies rhythmically about a certain mean value. This can be explained only by supposing that the star really consists of two stars revolving about each other so that the velocity of each star in the line of sight varies with the time and so the Doppler shift varies from time to time. By observing the variation of the Doppler shift with the time the period of revolution can be found. However, it is not sufficient to know only the period if it is desired to obtain the sum of the masses of the two stars from (23–14). It is also necessary to know the distance apart of the centers of the two stars. Spectroscopic examination alone will not give the mass of the two stars.

We have spoken only of the Fraunhofer or dark lines in the spectrum of a binary star. However, many stars give out a bright line spectrum the lines of which are doubled, indicating the binary nature of the star.

23.15. Eclipsing Binaries.—Certain stars are known as variable stars because their light is found to diminish in amount at regular intervals and then to return to its original strength. This is explained by supposing that the star consists of two stars which revolve about each other in such a way that first star A is wholly or partially hidden by star B from the observer on the earth and then star B is wholly or partially hidden by star A from the observer. The light of this double star is diminished only during an eclipse; at other times the star shines with constant brightness. In the case of eclipsing binaries, the earth is very nearly in the plane of revolution of the center of each star about the other, so that the orbits are observed almost end-on from the earth. Since the plane of orbits is pointing almost directly at us, star A is at one time moving directly towards us and at another time directly away from us, so that the velocities obtained from the maximum and minimum Doppler shifts for star A give the velocities of star A when it is moving directly away and when it is moving directly toward us. Similarly the velocities of star B can be found. If the two star velocities are v_A and v_B, and if for simplicity we suppose that the two stars as a whole are not moving relatively to us, the circumference of the orbit of A is Pv_A and that of the orbit of B is Pv_B, where P is the period. The radii of the orbits about the center of gravity are therefore $Pv_A/2\pi$ and $Pv_B/2\pi$ respectively, and the distance apart of the two stars is $P(v_A + v_B)/2\pi$. Knowing the period and the distance, the sum of the masses, $m_A + m_B$, can be found. Also the ratio m_A/m_B equals the ratio v_B/v_A so that the separate masses can be found.

23.16. Loss of Mass by the Sun.—We have seen in Sec. 23.6 that each square centimeter of the sun's surface is radiating energy at the rate of 6×10^{10} ergs/sec. We have mentioned in Sec. 13.11 that energy has mass according to the relativity theory, the mass m of an energy W being

$$m = \frac{W}{c^2} \tag{23-26}$$

where c is the velocity of light. Hence, each square centimeter of the sun's surface is radiating a mass of $6 \times 10^{10}/(3 \times 10^{10})^2 = 6.6 \times 10^{-11}$ gm/sec. This is not very great, but, if this is multiplied by the whole area of the surface of the sun, the rate of loss of mass is then $6.2 \times 10^{22} \times 6.6 \times 10^{-11} = 4.1 \times 10^{12}$ gm/sec, or 4,400,000 tons per second. This is a rate of loss of mass of 380,000 million tons per day. However, although the sun is losing mass at this enormous rate, yet its mass is so large that this loss amounts only to a fraction $3.8 \times 10^{11}/2.2 \times 10^{27} = 1.73 \times 10^{-16}$ of its mass per day, or 6.3×10^{-14} of its mass in a year. At its present rate of burning up, the sun will have shrunk to half its present mass in 8×10^{12} years.

23.17. Bethe's Carbon Cycle.—The source of the sun's energy has long been a puzzle to physicists. Bethe, of Cornell, has suggested a series of nuclear reactions which will supply the energy. These reactions are described by the following:

$$ {}_6\mathrm{C}^{12} + {}_1\mathrm{H}^1 \rightarrow {}_7\mathrm{N}^{13} \qquad (23\text{–}27) $$

$$ {}_7\mathrm{N}^{13} \rightarrow {}_7\mathrm{C}^{13} + {}_{+1}e^0 \qquad (23\text{–}28) $$

$$ {}_6\mathrm{C}^{13} + {}_1\mathrm{H}^1 \rightarrow {}_7\mathrm{N}^{14} \qquad (23\text{–}29) $$

$$ {}_7\mathrm{N}^{14} + {}_1\mathrm{H}^1 \rightarrow {}_8\mathrm{O}^{15} \qquad (23\text{–}30) $$

$$ {}_8\mathrm{O}^{15} \rightarrow {}_7\mathrm{N}^{15} + {}_{+1}e^0 \qquad (23\text{–}31) $$

$$ {}_7\mathrm{N}^{15} + {}_1\mathrm{H}^1 \rightarrow {}_6\mathrm{C}^{12} + {}_2\mathrm{He}^4 \qquad (23\text{–}32) $$

Adding these reactions together, we obtain

$$ {}_6\mathrm{C}^{12} + 4{}_1\mathrm{H}^1 \rightarrow {}_6\mathrm{C}^{12} + {}_2\mathrm{He}^4 + 2{}_{+1}e^0 \qquad (23\text{–}33) $$

The total process begins with ${}_6\mathrm{C}^{12}$ being bombarded with a proton and ends with ${}_6\mathrm{C}^{12}$ plus an α-particle and two positrons. The ${}_6\mathrm{C}^{12}$ acts as a catalyst for changing protons into α-particles and positrons. Referring to Table 20.2, we note that the total rest-mass on the left side of (23–23) is 16.03646 amu while the total rest-mass on the right is 16.00896 amu. There is thus a loss of 0.0275 amu. This appears as kinetic energy of amount $0.0275 \times 934 = 25.7$ mev. If the process represented by (23–27) to (23–32) occurs often enough, the heat output of the sun can be explained.

23.18. Formation of the Solar System.—The present accepted theory of the birth of the solar system is known as the tidal theory. We know that the tides in the sea are due to the attraction of the moon for the water of the sea. This attraction has the effect of heaping up the water on that side of the earth which is toward the moon. When two stars pass close to one another each raises tides on the other. If the approach is very close, the tides on each star are very high. It is supposed that many millons of years ago a collision between our sun and another sun or star nearly occurred. Instead of colliding, however, the two suns came close enough to raise tides upon each other. In the case of close approach the tides assume more the aspect of huge mountains of matter moving over the surface of the star. It must be remembered that these tidal mountains in the case of hot stars like the sun consist of gas and not liquid. If the approach of the two suns is still closer, these tidal mountains may be transformed into long filaments of gas protruding from the body of either sun. Mathematical theory shows that the greater damage will be done to the sun with the smaller mass and for this reason it is supposed that the star which approached our sun was the heavier of the two bodies. The filament of gas is supposed to have been many millions of miles long. When this gas found itself removed from the tremendous heat of the sun it began to condense into drops. These drops constitute what we now know as the planets. At first the new-born planets described highly complicated orbits under the gravitational attraction of both our sun and the other star. Finally, the bigger star receded and left the planets revolving about the smaller star—our sun—alone.

It is supposed also that the satellites of the various planets were formed in a somewhat similar way. After the birth of the planets due to the close approach of the big star to the sun, the planets followed erratic orbits. Having partly liquefied, it is likely that at some time a planet might pass so close to the sun that tides would be raised on the surface of the planet due to the attraction of the sun and a filament torn off. This filament subsequently condensed into one or more satellites or moons which revolved about the parent planet.

The planets newly formed from the filament of gas cannot be too close together because, as they grow and become more massive, their mutual gravitational attractions become greater and, if they are too close together, these mutual attractions may become so great as to raise tides upon each other and large filaments of matter are torn off. Roche has made a mathematical investigation of this tidal break-up and has shown that when a small body approaches within a certain distance of a much larger body the smaller body becomes broken up into fragments. This particular distance is known as Roche's limit. In the case of two nearly equal bodies it is probable that both are broken up if they come within a certain limiting distance. It is supposed that a planet once existed in the orbit now occupied by the asteroids and that this planet approached within Roche's limit of the planet Jupiter and so became shattered into fragments.

23.19. Nebulae and the Expanding Universe.—Our own sun is one of many million stars forming a galaxy which we call the Milky Way. We have learned that the orbits of the planets in the solar system are very nearly in the same plane so that the solar system as a whole may be likened to a cartwheel with the sun at the hub and the planets at various distances from the center along the spokes of the wheel. However, the analogy is not perfect because the planets revolve with different angular speeds around the sun—the outer planets revolving with smaller angular velocity than the inner planets. Actually, of course, the solar system revolves about the center of gravity of the system. Similarly the stars of the Milky Way form a sort of cartwheel with the difference that there is no special body like the sun near the center of gravity of the system of stars which revolves about this point. Our galaxy has a radius (the distance along the spoke of a cartwheel) of the order of 130,000 light-years.

Viewed from a distance great with respect to 130,000 light-years the stars of our galaxy would not be individually distinguishable and the whole would merge into a cloudy mass emitting light. Such a cloudy mass is known as a nebula. Telescopic observation shows that there are several million nebulae outside our own galaxy.

From the Doppler shift of the bright line spectrum of a nebula its velocity in the line of sight can be determined. It is a remarkable fact that almost all the nebulae are traveling away from us at enormous speeds. Most nebulae are receding at speeds of more than 1000 mi/sec while some of them approach a speed of 10,000 mi/sec away from us. There is an indication that the more distant the nebula the greater is its speed of recession. Expressed mathematically, the velocity of recession of a nebula at a distance r is therefore

$$v = \gamma r \tag{23-34}$$

where γ is what Eddington calls the cosmical constant. The value of γ is estimated to be 500 kilometers per second per megaparsec (a megaparsec = 10^6 parsecs = 3.26×10^6 light-years = 1.91×10^{19} mi). According to some writers, notably Eddington, the phenomena which are described by (23–34) indicate an expanding or "exploding" universe. However, other investigators in the field of astrophysics accept neither the idea of an expanding universe nor that of a cosmical constant.

The concepts of the expanding universe and the cosmical constant have led Eddington to some interesting speculations. For instance, some idea of the size and mass of the universe can be obtained. At a certain distance R the speed of recession of a nebula would equal the speed of light. From (23–34),

$$R = c/\gamma \tag{23-35}$$

Since c = 300,000 km/sec; R = 300,000/500 = 600 megaparsecs. Light emitted at a distance R from us would have an infinite wave-length or zero frequency when it reached us. Anything beyond the distance R would forever be outside our knowledge and so this is taken as a measure of the radius of the universe. Within a sphere of this radius (we ourselves are at the center of the sphere) there is the total mass M of the universe. We can imagine a quantum (or photon) of light upon the surface of this sphere and encircling the universe. According to Einstein's relativity theory a photon of energy $h\nu$ has a mass $h\nu/c^2$. This is attracted by the mass M toward the center of the sphere of radius R by the gravitational force $(GMh\nu/c^2)/R^2$, where G is the constant of gravitation. This equals the centrifugal force of the photon $(h\nu/c^2)c^2/R$ and we have

$$\frac{GMh\nu}{c^2R^2} = \frac{h\nu}{R}$$

whence

$$M = c^2R/G \tag{23-36}$$

Putting in values, we obtain $M = 2.5 \times 10^{55}$ gm. This is 10^{22} times the mass of our own sun. If we suppose all matter to consist of protons and electrons, the number of protons in the universe is $2.5 \times 10^{55}/(1.66 \times 10^{-24}) = 1.5 \times 10^{79}$, together with an equal number of electrons. The circumference of the universe is $2\pi \times 600 = 3770$ megaparsecs $= 1.23 \times 10^{10}$ light-years. It would thus take a ray of light over twelve billion years to encircle the universe. According to some writers, a ray of light will traverse its original path after these many years.

In closing this section it should be pointed out that the above speculations may prove to be quite "wild."

CHAPTER XXIII

PROBLEMS

1. Find the root mean square velocities in cm/sec of electrons, hydrogen, nitrogen, and oxygen atoms at a temperature of 50,000,000° C. Also express the root mean square velocity of the electrons in volts.

2. Electrons which are rushing around with a speed corresponding to 50,000,000° C produce x-rays when suddenly stopped. Assuming that all the electrons are moving with the root mean square velocity, determine the shortest x-rays which will be produced by the sudden stoppage of these electrons upon impact with heavy atoms. Use the Duane-Hunt relation.

3. All electrons are not moving with the same velocity as we have assumed in Problem 2. Many electrons are moving with a velocity greater than the root mean square velocity. When these swiftly moving electrons are suddenly stopped, x-rays of shorter wave-length are produced. The wave-length of the x-rays of maximum intensity is given by Wien's displacement law. Hence, find the wave-length of the x-rays of maximum intensity in the interior of a sun where the temperature is 60,000,000° C.

4. How long does light take to travel from (*a*) the sun to the earth, (*b*) the sun to Neptune?

5. The parallax of a certain star is 0.21 second of arc. Find the distance of the star in (*a*) light-years, and (*b*) miles.

6. The light from a certain star is examined with a spectroscope and it is found that the wave-length of the light of maximum intensity is 3000 A. Find the surface temperature of the star.

7. The wave-length of the D_1 line of sodium is 5896 A. In the spectrum of a star the D_1 line is found to have a wave-length of 5901 A. Find the velocity of the star in the line of sight.

8. Two stars, *A* and *B*, are observed to move back and forth against the background of distant stars. *A* moves back and forth along an elongated ellipse whose major axis subtends an angle of 4 seconds of arc, while *B* moves back and forth along another elongated ellipse whose major axis subtends an angle of 12 seconds of arc at the earth. During these oscillations the two stars pass close to each other, the major axes of the two ellipses being in the same straight line. The period of oscillation of each star is 50 years. The two stars thus constitute a visual binary. The parallax of the star *A* is 0.38 second of arc. Assuming the real orbit of each star to be circular, calculate the mass of each star in terms of the mass of the sun.

9. A certain star shows two sets of spectrum lines, so that the star is a spectroscopic binary. Fixing the attention upon one line, say, the D_1 line of sodium, of each set, it is found that for one set the wave-length of the D_1 line changes with the time according to the values shown in column *B*, while the wave-length of the D_1 line in the second set changes according to the values shown in column *C* of the following table.

Time, days	Wave-length angstroms	
	B	*C*
0	5897.170	5896.030
10	7.034	6.132
20	6.719	6.372
30	6.335	6.663
40	6.014	6.907
50	5.873	7.012
60	5.961	6.947
70	6.250	6.728
80	6.632	6.438
90	6.976	6.178
100	7.160	6.038
110	7.130	6.072
120	6.864	6.262
130	6.490	6.544

The true wave-length of the sodium D_1 line is 5895.932. Determine the velocities at the various times and plot two curves of velocity against time. These two curves cut each other at a velocity which is the velocity of the center of gravity of the binary in the line of sight. Assuming that the spectroscopic binary is also an eclipsing binary and that the orbits of the two stars forming the binary are circular, calculate the mass of each of the stars.

CHAPTER XXIV

RELATIVITY—GENERAL THEORY

24.1. Relative Acceleration.—The theory described in Chapter XIII is known as the Special Relativity Theory. This theory was announced by Einstein in 1905. The Special Relativity Theory has to do with those problems where one frame of reference is moving with constant velocity v with respect to a second frame of reference—or in other words, where one observer and his measuring instruments are moving with constant velocity v with respect to a second observer and the second observer's instruments. In 1915 Einstein announced the General Theory of Relativity. This theory has to do with problems where Frame 2 no longer moves with constant velocity but with acceleration relative to Frame 1. Since the General Theory involves the knowledge of a branch of mathematics called tensor calculus and this is beyond the scope of this book, we shall content ourselves with just a glimpse of the theory. The principal accomplishment of the theory is the explanation of gravitation.

Einstein was led to his ideas by a consideration of centrifugal force. There is really no such force as centrifugal force. A force is needed to make a body move in a circular path. This force is known as the centripetal force and is directed to the center of the circle of the path of the moving body. If a stone at the end of a string is whirled in a circle, a tension equal to mv^2/r is developed in the string. To an observer on the stone it appears that the stone is trying to get away from the center about which the stone is being whirled. If, further, this observer can see only the center and cannot see other objects and therefore does not know that the stone upon which he is placed is moving in a circular path, he will be of the opinion that there is a force of repulsion between the stone and the center of the circle because if he cuts the string he finds that (to him) the stone will fly away from the point which (to us) is the center of revolution. To the observer on the stone there is a force of repulsion while to us there is not. Einstein saw that this apparent repulsive force is due to the curved motion of the stone, and he extended the idea of curved motion evident in centrifugal force to a sort of curved space in order to explain gravitational force.

24.2. Advance of the Perihelion of Mercury.—According to Einstein the space near a heavy mass is curved so that another body which enters this space travels in a curved path. The curvature of the space increases the closer the other body approaches to the heavy mass. The orbits of the planets around the sun are explained by this theory. It so happens that for those planets which are at relatively great distances from the sun and whose orbits are nearly circular, the Einstein law of gravitation reduces to the Newtonian inverse square law. However, in the case of Mercury, the planet is so close to the sun and its orbit is so elliptical that the Einstein law departs sufficiently from the Newtonian law for the two laws to give orbits which differ from one another enough for an experimental test. In Sec. 23.3 we have mentioned the outstanding discrepancy of 42 seconds of arc per century in the advance of the perihelion of Mercury. Einstein, using his theory, calculated an advance of 41 seconds.

24.3. Bending of a Ray of Light.—In Sec. 13.11 we have shown that according to the Special Theory energy has mass. We must now consider two kinds of mass—inertial and gravitational mass. We have stated that the mass of a wound-up watch spring is greater than that of the same spring unwound. According to the Special Theory the increase of mass is inertial—that is, the acceleration developed when a given force acts upon the spring will be less when the spring is wound-up than when unwound. But does the spring weigh heavier when wound-up? Weight involves the gravitational mass of a body. Einstein's General Theory states that the inertial mass of a body is always the same as its gravitational mass, so that the wound-up watch spring is a little heavier than the unwound spring. Now a quantum of light has an energy $h\nu$ and should therefore have both an inertial and a gravitational mass $h\nu/c^2$. If a quantum passes near the sun it therefore should travel in a path somewhat like that of a planet or comet moving with a very high speed—a speed approaching that of light. As a result the path of the quantum is slightly bent towards the sun as it passes the sun. This suggested an experiment to measure the deviation of a ray of light as it passes the limb of the sun. This experiment can be performed only during a total eclipse of the sun. Stars which appear just at the edge of the disc of the sun (the sun being obscured by the moon) are slightly displaced from their true positions by an angle θ which according to Einstein is given by

$$\theta = \frac{4GM}{ac^2} \tag{24-1}$$

where θ is measured in radians, G is the gravitational constant, M and a the mass and radius of the sun respectively and c the velocity of light. Putting in numerical values, θ works out as 1.74 seconds of arc. Expeditions were organized to observe the total eclipses of 1919, 1922 and later years. The results agreed with Einstein's prediction.

24.4. Shift of the Fraunhofer Lines.—The third prediction of the General Theory is that when light moves from a strong gravitational field the wave-length becomes slightly longer. The formula in this case can be obtained by a method which is within the ability of the student. Consider a quantum $h\nu$ traveling from the sun to the earth. This quantum has a mass $h\nu/c^2$ and is attracted to the sun with a force $GMh\nu/c^2r^2$ when at a distance r from the sun. In traveling a distance dr, the work done against the attractive force is $GMh\nu dr/r^2c^2$ and this is done at the expense of the energy of the quantum. The loss of energy is therefore $d(h\nu)$ where

$$d(h\nu) = -\frac{GMh\nu}{c^2} \cdot \frac{dr}{r^2} \tag{24-2}$$

The loss of energy in the quantum as it moves from the surface of the sun to infinity is

$$\delta(h\nu) = -\int_a^\infty \frac{GMh\nu}{c^2} \cdot \frac{dr}{r^2} = -\frac{GMh\nu}{ac^2} \tag{24-3}$$

Canceling out h, we have

$$\delta\nu = -\frac{GM\nu}{ac^2}$$

where a is the radius of the sun.

In terms of the wave-length, we have, since $\nu = c/\lambda$,

$$\frac{\delta\lambda}{\lambda} = \frac{GM}{ac^2} \tag{24-4}$$

There is thus an increase of wave-length and the spectrum line shifts toward the red. The expected shift for a line of 5000 angstroms from the sun is 0.01 angstrom.

The experimental confirmation of this is difficult because of the Doppler shift. If there is relative motion between the sun and the observer, the Doppler shift is given by $\delta\lambda/\lambda = v/c$. It is difficult to correct for the Doppler shift. However, observations have recently been made on the spectrum lines in the light coming to us from the Companion of Sirius. This star as we have noted in Sec. 23.9, is a very dense star of large mass and small radius, so that M/a is large. The expected Einstein shift is some 30 times that for our own sun. It is therefore easier to correct for the Doppler effect and the observations have confirmed Einstein's value.

CHAPTER XXIV

PROBLEMS

1. The mass of the Companion of Sirius is 0.82 times that of the sun, while its diameter is 0.032 times that of the sun. By how much will a ray of light be deviated as it passes close to the surface of the Companion of Sirius?

2. Calculate the shift of a Fraunhofer line of wave-length 4500A coming from a star whose mass is 1.6 times that of the sun and whose diameter is 0.8 that of the sun.

CHAPTER XXV

PHILOSOPHICAL IMPLICATIONS

25.1. Trend to Whole Numbers.—At the beginning of the nineteenth century Dalton announced the atomic theory of chemistry and Young discovered interference of light. The idea of integers or whole numbers is common to the atomic theory and to the explanation of interference by means of the wave theory. Any piece of matter contains a whole number of atoms and for diffraction in particular directions from a diffraction grating the path difference of the rays through adjacent slits must equal a whole number of wave-lengths. Later on, with Faraday's discovery of the laws of electrolysis, it began to be evident that a charge of electricity must consist of a whole number of fundamental units of charge. At the end of the nineteenth century the electron, the fundamental unit of negative electricity, was discovered. With the discovery of α-particles physicists became accustomed to the idea of particles carrying a positive charge whose numerical value was double that of the electron. In the second decade of the twentieth century H-particles or protons were discovered and each of these was found to carry a positive charge equal numerically to the charge on the electron. No particles have ever been found with a charge equal to a fractional number of electrons or protons. Nowadays, we are so accustomed to the idea of any charge consisting of a whole number of protons and electrons that we forget that this idea is quite recent. Again, in the vibration of strings or open organ pipes in sound we find that the string or organ pipe must contain a whole number of loops. In the quantum theory of the atom one of Bohr's assumptions is that the angular momentum of the electron about the nucleus is a whole number times $h/2\pi$. In the theory of the specific heat of a solid, we have used a formula for the number of vibrational modes with frequencies between ν and $\nu + d\nu$, and this number is a whole number. In the Debye-Scherrer circles given by a powdered crystal x-ray photograph of a simple cubic crystal, the circles correspond to angles θ_r given by

$$\sin \theta_r = (\lambda/2d_{100})\sqrt{r}$$

where

$$r = h^2 + k^2 + l^2 \tag{25-1}$$

and h, k and l are whole numbers. Since three whole numbers whose squares add to 7 cannot be found, the circle corresponding to $r = 7$ does not appear in the powdered crystal photograph of a simple cubic crystal.

The examples given in the previous paragraph show the trend towards a physics based upon whole numbers. Whole numbers are frequently connected with counting and there is a growing opinion that much of theoretical physics can be reduced to schemes of counting. Of course if counting is to be done, the units to be counted must be known. We know the fundamental unit or "atom" of electric charge and so we can theoretically obtain the charge on a body by counting the numbers of positive and negative fundamental units of charge. In a similar way there is perhaps a fundamental unit or "atom" of length. It has been suggested that the radius of the electron ($r = 2e'^2/3mc^2 = 1.9 \times 10^{-13}$ cm) might be such an "atom" of length. A more startling idea is that perhaps time proceeds in atomic jerks instead of in a continuous flow. "Chronon" has been suggested for the name of the fundamental unit of time. The quantum theory may be a cumbersome way of taking account of the atomic make-up of time and length. Eddington has suggested that the quantity ($hc/2\pi e'^2$) must be a whole number with the value 137. However, putting in the values of h, c, and e' which have been obtained by experiment, we obtain 137.03 for the value of the quantity.

25.2. Heisenberg's Uncertainty Principle.—In Fig. 17.2 we have seen that the electron which revolves about the nucleus of an atom is from the point of view of wave mechanics to be replaced by a group of waves. This being so, one might inquire as to the exact position of the electron. Is the electron to be thought of as being at the middle of the group? One might answer such questions by saying that the electron is most probably at such and such a position although it might be at some other position. However, Heisenberg in 1927 took the view that all such questions were beside the point and that the position of the electron is actually indefinite over a certain range. He showed that theoretically we can measure either the position or the velocity of an electron exactly, but that we cannot at the same time measure both the position and velocity exactly. If an attempt is made to measure both the position and the velocity of an electron simultaneously, Heisenberg says that there will be an uncertainty dx in the position and an uncertainty dv in the velocity of the electron such that

$$dx \cdot dv \nless \frac{h}{4\pi m} \tag{25-2}$$

The reader must understand that the uncertainties dx and dv are not to be thought of as experimental errors in the measurement of the position x and the velocity v of the electron. The uncertainties dx and dv are to be taken as existing in nature itself. One cannot even think of an electron being exactly at a position x and having a velocity v. The best one can do is to think of an electron being somewhere between $x - dx/2$ and $x + dx/2$ and having a velocity somewhere between $v - dv/2$ and $v + dv/2$. By reference to Fig. 17.2b, the uncertainty of the position of the electron is immediately made obvious.

25.3. Freedon of Will.—Some interesting philosophical implications have been drawn from Heisenberg's uncertainty principle. Until the statement of this principle, many, and perhaps most, physicists believed that, if the position and velocity of every electron and proton in the universe were known, the position and velocity of each electron and proton at any future time could be predicted, although the mathematical processes needed to arrive at the prediction might be extremely difficult and tedious. Thus any future state of the universe was supposed to grow out of the present state of the universe. Coming to human affairs, we had the doctrine of materialism, which is a form of the doctrine of predestination as preached by Calvin at Geneva, Switzerland, about 1550. According to the doctrine of materialism, all human actions are controlled by the electronic and atomic configurations in the body, so that there is no room for freedom of will. Such freedom of will as is observed is only apparent and not real. When one acts in accordance with a wish, the wish is itself a result of material circumstances. This was the belief of many scientists until 1927. However, according to Eddington, the uncertainty principle has changed the situation completely. According to Eddington and his school of thought, the uncertainty principle requires that a future state of the universe is not absolutely determined by its present state and that there is an inherent and natural uncertainty in all predictions of the future. Hence, this school argues, human actions may not be entirely determined by the electronic and atomic configurations of the body, and, if this is so, there may be room for absolute freedom of will within the bounds set by the uncertainty principle.

25.4. Causation and Time.—The principle of causation or the relation between cause and effect has also been called into question by modern science. We have seen in the relativity theory that the times of two events *A* and *B* may be such that *A* happens before *B* to one observer and after *B* to a second observer. If the first observer concludes that *A* is the cause of the effect *B*, what will the second observer conclude? If *A* happens after *B* it is difficult to conceive of *A* being the cause of *B*. This brings in the philosophical conception of time. Just what is time? Why is it that time goes only in one direction? Eddington relates the direction of time to the direction in which the randomness of the universe is increasing. Thus, if we take a pack of cards in which the cards are originally arranged in a given order and if we shuffle the cards, the original arrangement disappears and the randomness of the arrangement increases as the cards are shuffled. The cards are more at random at a later time. In physics there is a quantity known as entropy, and there is a law of physics which states that the entropy always tends to increase with the time. If we imagine a gas with its molecules all moving in one direction, then, after hitting the walls of the containing vessel, the velocities become more random, and, as time goes on, the molecules hit one another and the randomness increases. Time goes on in the direction of increasing randomness. Entropy can be thought of as a measure of randomness. The entropy law can be considered as defining the direction of time. If one can imagine a state of affairs where the randomness could not be increased and the entropy remains stationary, there would be no direction to time and time would cease. The universe is running down as time goes on. When the universe has completely run down so that everything is at a uniform temperature and all motion is random, there will be no time.

25.5. The Operational Viewpoint.—The previous two sections on Freedom of Will and on Causation and Time are perhaps too philosophical for the pure physicist, and we shall now return to the consideration of a viewpoint which has more to do with experimental physics. Recently Bridgman has advanced what is known as the operational viewpoint. He says that a concept, say length, has meaning only when we describe the operations by which it is measured.

Let us consider the concept of length. The length of an ordinary object such as a floor rug is obtained in effect by laying a measuring rod such as a meter stick against the object so that one end of the measuring rod coincides with one end of the object and marking the position of the other end of the rod on the object. The rod is then moved so that its first end coincides with the mark on the object and the position of the second end of the rod marked on the object. The process is repeated until the other end of the object is reached. The length of the body is the number of times the rod is applied. A length measured in this way involves the sense of touch, so that the length may be called a tactual length.

Now suppose we wish to measure the length of a moving object such as a moving street car. So long as the street car can be boarded by the observer, its length can be measured by the simple method just described. However, if the street car is going so fast that the observer cannot board it, a new set of operations for measuring its length must be devised. Such a set of operations was described in Sec. 13.7 of the chapter on Relativity. It is immediately seen that Einstein's method involves the idea of simultaneity. But, as has been pointed out in Sec. 13.8, simultaneity is a relative matter. Two events which are simultaneous to one observer are not simultaneous to a second observer who is moving relatively to the first observer. The length of an object as measured by Einstein's method thus depends upon the velocity of the moving object—in fact, if the object is moving in the direction of its length, this length as measured by a stationary observer grows smaller as the velocity of the object increases. Hence, the length of the body as measured by an observer traveling with the body is different from the length as measured by an observer with respect to whom the body is moving. On Bridgman's viewpoint this is a natural outcome of the two different lengths being measured by two different sets of operations.

We have just discussed the effect of velocity on length. Now let us consider the effect of size of the object. The procedure for measuring large tracts of land is to measure a base line by the method of the measuring rod described above and then to measure angles by means of a surveyor's telescope. The angles measured are angles between beams of light. The lengths as calculated by trigonometry depend on the assumption that light travels in straight lines. The set of operations is different from the previous two sets, and we have a third kind of "length," which, strictly speaking, should be called by a third name. All trigonometrical measurements depend upon the theorem of Pythagoras and upon the assumption that the angles of a triangle add up to

180°. Space in which these two propositions are true is called Euclidean space. If the angles of a very large triangle on the earth were measured optically by means of an accurate surveyor's telescope, the sum of the angles would be found to be slightly greater or less than 180°, depending upon the direction in which the light beam had traveled around the triangle with respect to the rotation of the earth.

In the case of solar and stellar distances the situation is still worse. Here all distances are optical in character and there never can be a comparison of optical with tactual lengths. Also there is no possible way of measuring the three angles of a triangle. A check on the sum of the three angles of a triangle is very important in stellar distances because the distance of a star is measured in terms of its parallax and the parallax at best is less than a second of arc. Any slight error in our assumption that space is Euclidean would throw out entirely the values of the distances of the stars which have been calculated from their parallaxes. We thus see that the concept of length has changed completely in going from the lengths of ordinary objects to stellar distances. To say that the distance of Sirius is eight light-years is actually and conceptually an entirely different thing from saying that the length of a room is ten feet.

The concept of length also changes from the ordinary concept of length when we go to the region of very small distances. Just what do we mean by saying that the wave-length of the molybdenum $K\alpha$ x-ray spectrum line is 0.71 A? The meaning can be found only by examining the operation of measuring the wave-length. The wave-length is obtained from the measurement of the angle θ in the Bragg equation

$$n\lambda = 2d \sin \theta \tag{25-3}$$

and from the calculation of d from the molecular weight and density of rocksalt and the value of Avogadro's number. Further, the concept of the length between two wave crests depends upon the assumption of a wave theory of x-rays. Again, we see that the set of operations for measuring the length of the side of a house is entirely different from the set of operations for measuring the length between the wave crests in a train of molybdenum $K\alpha$ x-rays. Again we can say that the two "lengths" are entirely different and should be called by different names.

25.6. Mass and Atomic Energy.—It is sometimes said that mass is converted into energy in the atomic bomb. This statement is incorrect.

Rest-mass is changed into energy. The mass of an electron moving with velocity v is $m_0/\sqrt{1 - \beta^2}$, where $\beta = v/c$. This is equivalent to an energy $m_0c^2/\sqrt{1 - \beta^2}$. This energy consists of rest-mass energy m_0c^2 and kinetic energy $m_0c^2(1/\sqrt{1 - \beta^2} - 1)$. The mass of this kinetic energy is $m_0(1/\sqrt{1 - \beta^2} - 1)$. If an atomic bomb were exploded in the middle of a huge cubic box with an edge 40 mi long, and if the walls of the box were impervious to all particles and all radiation, including heat radiation, the mass and also the weight of the box would be the same after the explosion as before. There is no loss of mass. If an observer outside the box could take a sudden peep inside through a shutter just before the explosion he would see atoms, electrons, photons, protons, neutrons, and molecules moving about normally; after the explosion on taking a second peep he would see a terrific commotion inside the box but the mass of the box and its contents would remain unaltered. The total energy in the box remains the same. Before the explosion the ratio of kinetic energy to total energy has a certain value; after the explosion the ratio has a higher value. From the cosmic viewpoint, that is all that has happened.

25.7. Physical Theories.—Modern physicists do not take physical theories so seriously as they used to do. We have noted in Sec. 22.11 that Sir William Bragg has said that on Mondays, Wednesdays and Fridays one uses the quantum theory, while on Tuesdays, Thursdays and Saturdays one uses the wave theory. One uses the quantum theory to "explain" a certain group of phenomena and the wave theory to "explain" another group of phenomena. Possibly the real truth lies somewhere between the two theories. The author's own opinion is that the word "explain" in its usually accepted meaning has no place in modern physics. Just what does one mean by "explaining" the Compton Effect in terms of bouncing billiard balls—after all, neither light nor electrons are billiard balls. The author's own view is that a theory is to be judged only by its usefulness. If a theory correlates a large number of observable and measurable phenomena, the theory is profitable. The author inclines to the view that the purpose of a theory is to enable physicists to find good empirical relations between measurable physical quantities. For example, the law of Boyle and Charles, $pv = RT$, is a useful relation connecting the pressure, volume, and temperature of a gas. The relation holds well as long as the measurements of p, v, and T are not too accurate. The law is derivable from the simplest assumptions of the kinetic theory of gases. It may be said therefore that the simple form of the kinetic theory is "adequate" to give a relation which connects the experimental results up to a certain accuracy of measurement. The relation given by the simple theory becomes "inadequate" for a higher accuracy of measurement. However, by adding the concepts of intermolecular forces and molecular size, Van der Waals' equation, $(p + a/v^2)(v - b) = RT$, is obtained. This works well for higher accuracy of measurement. The revised theory is therefore "adequate" on a plane of higher accuracy. For a plane of still higher accuracy, Van der Waals' equation and the theory back of it are inadequate. Consequently, for a still higher degree of accuracy, a further revision of the theory must be made. It is not proper to say that the simple theory is wrong but only that it is inadequate. Likewise Newton's law of gravitation is adequate up to a certain degree of accuracy; Einstein's law of gravitation is adequate to a higher degree of accuracy. It is the object of a theory to "correlate" rather than "explain."

In Sec. 17.4 we saw that Heisenberg in 1925–26 inveighed against such unobservable quantities as the radii of orbits in Bohr's quantum theory. If unobservables are used in a theory, they must not appear in the final equations. These final equations must contain measurable quantities—the readings of thermometers, voltmeters, ammeters, meter sticks, clocks, and so on. However, it is rather difficult in 1946 to inveigh against unobservables; the neutrino, an unobservable particle, is definitely part of the theories of beta decay and mesotron decay. For a time the mesotron as used by Yukawa in his theory of the nucleus was an unobservable particle, but now in 1946 it has been observed. Unobservable quantities may later on become observable and measurable. The author believes that it is the hope of physicists that there shall be no unobservables.

The author closes this book with a note of warning:

> . . . and, considering the constant tendency of the mind to rest on an assumption, and, when it answers every present purpose, to forget that it is an assumption, we ought to remember that it, in such cases, becomes a prejudice, and inevitably interferes, more or less, with a clearsighted judgment.
>
> —Michael Faraday,
> *Philosophical Magazine*, 1844

APPENDIX I

NOBEL PRIZE AWARDS TO PHYSICISTS

1901	Roentgen (G)
1902	Lorentz (H)
	Zeeman (H)
1903	Becquerel (F)
	P. and M. Curie (F)
1904	Rayleigh (E)
	Ramsay (chem) (E)
1905	Lenard (G)
1906	J. J. Thomson (E)
1907	Michelson (A)
1908	Lippmann (F)
	Rutherford (chem) (E; b.N.Z.)
1909	Marconi (It)
	Braun (G)
1910	Van der Waals (H)
1911	Wien (G)
	M. Curie (chem) (F)
1912	Dalen (Swe)
1913	Onnes (H)
1914	Laue (G)
1915	W. H. Bragg (E)
	W. L. Bragg (E;b. Australia)
1917	Barkla (E)
1918	Planck (G)
1919	Stark (G)
1920	Guillaume (Swi)
1921	Einstein (G; now A)
	Soddy (chem) (E)
1922	Bohr (D)
	Aston (chem) (E)
1923	Millikan (A)
1924	Siegbahn (Swe)
1925	Frank (G)
	Hertz (G)
1926	Perrin (F)
1927	A. H. Compton (A)
	C. T. R. Wilson (E)
1928	Richardson (E)
1929	L. de Broglie (F)
1930	Raman (India)
1932	Heisenberg (G)
	Langmuir (chem) (A)
1933	Dirac (E)
	Schroedinger (G; now Ireland)
1934	Urey (chem) (A)
1935	Chadwick (E)
	M. Joliot Mme Curie-Joliot } (chem) (F)
1936	Anderson (A)
	Hess (Austria; now A)
	Debye (chem) (G;b. H; now A)
1937	Davisson (A)
	G. P. Thomson (E)
1938	Fermi (It; now A)
1939	Lawrence (A)
1943	Stern (G; now A)
1944	Rabi (A)
1944	Hahn (chem) (G)
1945	Pauli (Austria; now A)
1946	Bridgman (A)

The first letter in the parentheses after a recipient's name indicates the country in which most of his work was done. If there are other letters, these refer to the country of birth and to his present or last country of residence H denotes Holland; N.Z., New Zealand; Swe, Sweden; Swi, Switzerland.

Alfred Nobel, the Swedish scientist who invented dynamite, died in 1896. The interest on the fortune which he bequeathed is such that an undivided Nobel prize was worth $40,555 in 1935. Nobel prizes are awarded in physics, chemistry, medicine, literature and peace.

APPENDIX II

USEFUL CONSTANTS TO REMEMBER

The values given are approximate. For many purposes these values are sufficiently correct. Should the student desire exact values, he is referred to the International Critical Tables.

Name of Constant	Symbol	Value
Charge of the electron	e'	4.8×10^{-10} esu
Mass of the electron	m	9.1×10^{-28} gm
Mass of unit atomic weight ($O^{16} = 16$)	......	1.66×10^{-24} gm
Mass of the Proton	......	1.67×10^{-24} gm
Rest-mass energy of the electron	......	0.51 mev
Rest-mass energy of unit atomic weight ($O^{16} = 16$)	......	934 mev
Rest-mass energy of the proton	......	941 mev
Planck's constant	h	6.60×10^{-27} erg · sec
Velocity of light in a vacuum	c	3×10^{10} cm/sec
Rydberg constant	R	109,700 cm^{-1}
Avogadro's number	N	6.03×10^{23}
Volume of gram-molecule of a perfect gas at 0° C and 76 cm pressure	......	22,400 cm^3
Gas constant per gram-molecule	R	8.31×10^7 ergs/deg
Boltzmann's constant (R/N)	k	1.38×10^{-16} erg/deg
Atomic weight of H' ($O^{16} = 16$)	......	1.0081
Faraday's constant	......	9642 emu/gm-equivalent
Mechanical equivalent of heat	J	4.18 joules/cal
Gravitational constant	G	6.66×10^{-8} dyne · cm^2/gm^2
λ (in A) $\times$ W (in ev) for a photon	......	12,345 (approximately, but easy to remember)

APPENDIX III

CHEMICAL ELEMENTS

INTERNATIONAL ATOMIC WEIGHTS—BEFORE THE DISCOVERY OF FISSION

	Symbol	Atomic Number	Atomic Weight		Symbol	Atomic Number	Atomic Weight
Aluminum	Al	13	26.97	Molybdenum	Mo	42	96.0
Antimony	Sb	51	121.76	Neodymium	Nd	60	144.27
Argon	A	18	39.944	Neon	Ne	10	20.183
Arsenic	As	33	74.91	Nickel	Ni	28	58.69
Barium	Ba	56	137.36	Nitrogen	N	7	14.008
Beryllium	Be	4	9.02	Osmium	Os	76	191.5
Bismuth	Bi	83	209.00	Oxygen	O	8	16.0000
Boron	B	5	10.82	Palladium	Pd	46	106.7
Bromine	Br	35	79.916	Phosphorus	P	15	31.02
Cadmium	Cd	48	112.41	Platinum	Pt	78	195.23
Calcium	Ca	20	40.08	Potassium	K	19	39.096
Carbon	C	6	12.01	Praseodymium	Pr	59	140.92
Cerium	Ce	58	140.13	Protactinium	Pa	91	231
Cesium	Cs	55	132.91	Radium	Ra	88	226.05
Chlorine	Cl	17	35.457	Radon	Rn	86	222
Chromium	Cr	24	52.01	Rhenium	Re	75	186.31
Cobalt	Co	27	58.94	Rhodium	Rh	45	102.91
Columbium	Cb	41	92.91	Rubidium	Rb	37	85.48
Copper	Cu	29	63.57	Ruthenium	Ru	44	101.7
Dysprosium	Dy	66	162.46	Samarium	Sm	62	150.43
Erbium	Er	68	167.64	Scandium	Sc	21	45.10
Europium	Eu	63	152.0	Selenium	Se	34	78.96
Fluorine	F	9	19.00	Silicon	Si	14	28.06
Gadolinium	Gd	64	156.9	Silver	Ag	47	107.880
Gallium	Ga	31	69.72	Sodium	Na	11	22.997
Germanium	Ge	32	72.60	Strontium	Sr	38	87.63
Gold	Au	79	197.2	Sulfur	S	16	32.06
Hafnium	Hf	72	178.6	Tantalum	Ta	73	180.88
Helium	He	2	4.002	Tellurium	Te	52	127.61
Holmium	Ho	67	163.5	Terbium	Tb	65	159.2
Hydrogen	H	1	1.0078	Thallium	Tl	81	204.39
Indium	In	49	114.76	Thorium	Th	90	232.12
Iodine	I	53	126.920	Thulium	Tm	69	169.4
Iridium	Ir	77	193.1	Tin	Sn	50	118.70
Iron	Fe	26	55.84	Titanium	Ti	22	47.90
Krypton	Kr	36	83.7	Tungsten	W	74	184.0
Lanthanum	La	57	138.92	Uranium	U	92	238.07
Lead	Pb	82	207.21	Vanadium	V	23	50.95
Lithium	Li	3	6.940	Xenon	Xe	54	131.3
Lutecium	Lu	71	175.0	Ytterbium	Yb	70	173.04
Magnesium	Mg	12	24.32	Yttrium	Y	39	88.92
Manganese	Mn	25	54.93	Zinc	Zn	30	65.38
Mercury	Hg	80	200.61	Zirconium	Zr	40	91.22

Courtesy of the Journal of the American Chemical Society

ANSWERS TO PROBLEMS

Wherever possible the answers to the problems have been worked out with the aid of a slide rule. The answers are therefore approximations. The student will save himself much time if he supplies himself with a slide rule.

CHAPTER II

1. $5x^4$, $20x^3$; $(3/2)\sqrt{x}$, $(3/4)\dfrac{1}{\sqrt{x}}$; $(1/2)\dfrac{1}{\sqrt{x}}$, $-(1/4)\dfrac{1}{\sqrt{x^3}}$; $-(1/2)\dfrac{1}{\sqrt{x^3}}$, $(3/4)\dfrac{1}{\sqrt{x^5}}$; $-\dfrac{1}{x^2}$, $\dfrac{2}{x^3}$; $-(3/2)\dfrac{1}{\sqrt{x^5}}$, $(15/4)\dfrac{1}{\sqrt{x^7}}$.

2. $-3\sin 3x$, $4\sin^3 x\cos x$, $\cot x$, $3(x-4)^2$, $12x^2(x^3+9)^3$, $3\cos(3x+4)$, $-3e^{-3x}$, $2\dfrac{\tan 2x}{\cos 2x}$.

4. $e^{-at}\{b\cos(bt+c) - a\sin(bt+c)\}$.

5. $Ae^{-b/T}\{b+2T\}$.

8. 5, 306° 50′; 65.0, 104° 15′; 2, 90°; **9.** 5, 3.66, 6.04. **10.** 25.2. **11.** 7.9.

CHAPTER III

1. 19,600 dyne/cm; 2.23 vib/sec. **2.** 3940 dyne-cm. **3.** 22,000 ergs; 0.391 cm. **4.** 3500 cm/sec; 38.6 ergs. **5.** $3U/2$; 144.4 cm/sec. **6.** 533.3 vib/sec. **7.** 64 (table), 67.86 (formula, using $\lambda = 1\frac{2}{3}$ cm.)

CHAPTER IV

1. 4.12 amp, 82.4 volts, 73 volts, 0 amp. **2.** 6.51 amp, 195.3 volts, 102.2 volts, 7.33 amp. **3.** 11.0 amp, 110 volts, 145.2 volts, 146.0 volts, 0 amp. **4.** 5.5, 5.0, 7.43 amp. **5.** 11.3 amp (through R and L), 12.4 amp (through C), 3.28 amp (total), 44 amp (d.c. through R). **6.** Max. current when $L = 338$ microhenrys. **7.** 4.15 amp, 0 watt; 4.07 amp, 83 watts. **8.** 109.3 volts. **9.** 57.8 volts. **10.** (*a*) 0.442 amp; 0; 0 watt. (*b*) 0.608 amp; 0.685; 500 watts; 5 amp. **11.** 5 ohms.

CHAPTER V

1. 942.5 lines, 0.03 lines/cm², 0.03 esu. **2.** 398 cm. **3.** 3980 esu, 19,900 ergs, 0.995 erg/mm³. **4.** 3980 dynes/cm². **5.** 3.33 esu. **6.** 1.603×10^{10} cm/sec.

CHAPTER VI

1. 10 cm, 15 esu, 1125 ergs. **2.** 4.46×10^{-19} gm. **3.** 160 emu. **4.** 8.00×10^{-11} dyne. **5.** 2×10^{26} cm/sec^2, 228 erg/sec, 2.28×10^{-15} erg.

CHAPTER VII

1. 22.6 cm. **2.** 29.9 meters/sec (viscosity coeff. at 23° C = 1.823×10^{-4} gm · cm^{-1} sec^{-1}). **3.** 0.0057 cm/sec (viscosity coeff. at 0° C = 1.709×10^{-4}), 19.35 volts.

CHAPTER VIII

1. 47,700 cm/sec. **2.** 215,000 cm/sec, 176,000 cm/sec. **3.** 3.14 miles/sec. **4.** 1680 ft/sec. **5.** 0.0088. **6.** (*a*) 0.09, (*b*) 0.33. **7.** 20,600 cm^3, 1.89×10^{23}, (*a*) 1.34×10^{22}. (*b*) 1.93×10^{22}. **8.** 6.21 kilometers. **9.** 6.47×10^{23}. **10.** 0.016, 0.037, 0.066, 0.116 mm.

CHAPTER IX

1. 0.075 cal/gm, 0.125 cal/gm. **2.** 0.156, 0.218 cal/gm. **3.** 0.136, 0.181 cal/gm; 326 meters/sec. **4.** 4.57, 2.27×10^5 cm/sec; 336° Abs; 0.002, 0.076, 0.085, 0.091 cal/gm. **5.** 0.044, 0.049 cal/gm (density of tin, 7.31 gm/cm^3).

CHAPTER X

1. 8.4, 18.8, 42, 94.2, 594×10^7 cm/sec. **2.** 2.38×10^{-9} erg. **3.** 2.79×10^{-8} microfarads. **4.** 1.5×10^{-11}, 3.4×10^{-6}, 3.7×10^{-8}, 0.35 amp. **5.** 4.52 volts. **7.** 0.115 amp.

CHAPTER XI

1. 2.59, 4.12, 2.06, 1.82, 5.37 volts. **2.** 4.09, 2.56, 4.62, 4.86, 1.31 volts; 0.94, none, 1.47, 1.71, none volts. **3.** 3390, 3050, 5800 A; 8.85, 9.83, 5.17×10^{14} sec^{-1}. **4.** 6.59, 5.38, 9.85×10^7 cm/sec.

CHAPTER XII

1. 1.94×10^{-10} ohm^{-1} volt$^{-2.5}$, 584 ohms, 59.2 ohms. **2.** 2500 ohms, 73.2 volts. **3.** Approximately 380 volts. **4.** 250 μμf, 0.00115 sec. **5.** 508 μh. **6.** 500 kc; ∞, 5.04×10^4, 2.45×10^4 ohms. **7.** 1210, 385. **8.** 150 ohms. **9.** 88.9 μf, $(0.038 - 2.39j)$, 2700 ohms, 11.1 ma, 13.3 volts (amplitude), 0.177 ma (amplitude), 66.7 ma. **10.** 0.145 watt, 5.33 watts. **11.** 0.998, 0.024, −0.002. **12.** 1572 ft/sec.

CHAPTER XIII

1. 2.975×10^{10} cm/sec. **2.** 2.75, 6.55, 35.8×10^{-7} erg; 1.73, 4.12, 22.5×10^5 volts. **3.** 55.3 cm. **4.** 1.81 hours. **5.** 11.37×10^{-28} gm. **6.** 2.23×10^{10} gm · cm^2 1.22×10^{14} ergs, 1.355×10^{-7} gm.

CHAPTER XIV

1. 0.99, 1.43, 1.65×10^{10} cm/sec. **2.** 0.187 cm. **3.** 2.77 gm^{-1}. **4.** 0.485, 0.497, 0.561 A. **5.** 3.14 A. **6.** 58° 36′, 50° 18′, 50° 2′. **7.** 29° 9′, 24° 58′, 24° 12′, 21° 10′. **8.** 1.022, 0.908, 0.891 A. **9.** 0.145 A. **10.** 31,800 volts. **11.** (*a*) 32.2, (*b*) 107.4. **12.** 16.9 gm^{-1} ($K\alpha$), 84.3.

CHAPTER XV

1. (*a*) 28° 39′, 25° 17′, 24° 43′; (*b*) 36° 00′, 31° 30′, 30° 48′. **2.** no (1, 1, 1) reflection since $1 + 1 + 1 = 3$, an odd number; 38° 40′; no circle; 22.2 cm. **3.** 5° 16′; none since 2, 1, 0 are neither all odd nor all even; 16° 37′. **4.** 62.3 min. **5.** 0.0054A^2; reduced to 0.416. **6.** 0.000195r.

CHAPTER XVI

1. 1.43, 3.33, 20, 500, 3,330, $14{,}100 \times 10^4$ cm^{-1}; 4.29, 10, 60, 1.500, 10,000, $42{,}300 \times 10^{14}$ sec^{-1}. **2.** 8.22, 9.74, 10.29, 10.52, 10.67, 10.73×10^4 cm^{-1} (Lyman); 1.523, 2.055, 2.30, 2.43, 2.495, 2.57×10^4 cm^{-1} (Balmer); 5.33, 7.78, 9.13, 9.94, 10.47, 10.83×10^3 cm^{-1} (Paschen); Limits: 911 A (Lyman), 3650 A (Balmer), 8200 A (Paschen). **3.** 5.42, 4.76, 0.675, 0.66×10^8 cm^{-1}. **4.** −54.4, −13.6, −6.05, −3.4, −2.18, −1.51 volts. **5.** 304, 256, 242, 236 A; 1635, 1210, 1081, 1022 A.

CHAPTER XVII

1. 1.43 A, 13° 10′. **2.** 425 ev, 1.22×10^9 cm/sec. **3.** 5.76×10^{-11} cm, 1.04×10^{-20} cm^2, 1.15×10^4 ev.

CHAPTER XVIII

1. 11,420 (8950), 30,000 (25,450), 74,500 (69,300), 91,500 (89,100) ev. **2.** 0.0049 A. **3.** 8.75×10^8 cm/sec, 0.0645 A.

CHAPTER XIX

1. 1.12×10^6 volts. **2.** 0.0187 A. **3.** 3710 years. **4.** 10 minutes. **5.** 0.93 cm. **6.** 2.4×10^7 cm/sec, 17, 8.0 cm. **7.** 6562.83 A.

CHAPTER XX

1. 38.3 cm, 1.125 mev. **2.** 9.0 mev, 85° 1′. **3.** 4.84 mev. **4.** 13.0098.

CHAPTER XXI

1. 0.486 ev. **2.** 17 (nearest integer). **3.** 13,100 barns; 9.45 cm^2 gm^{-1}; 3240 barns. **4.** 28.2 mev; 6.8×10^{21} ergs. **5.** 2.86×10^8 tons. **6.** 167,800 ton; 118 lb.

CHAPTER XXII

1. About 35°. **2.** 2260 mev, about 46°, 0.174. **3.** 6×10^5, 8.88 cm. **4.** 3.17×10^{-18}. **5.** 0.156, 287.

CHAPTER XXIII

1. 472, 11.1, 2.98, 2.79×10^7 cm/sec; 6,465 volts. **2.** 1.910 A. **3.** 0.48 A. **4.** 499 sec, 4 hr 10 min. **5.** (*a*) 15.5 light-years, (*b*) 9.09×10^{13} miles. **6.** 9,600° Abs. **7.** 254 km/sec away from the earth. **8.** 2.76, 0.92 times the mass of the sun. **9.** 1.81, 2.40×10^{33} gm.

CHAPTER XXIV

1. 44.6 seconds of arc. **2.** 0.019 A.

INDEX