DOCUMENT RESUME

ED 033 855 SE 007 606

AUTHOR	Puri, O. P.
TITLE	Concepts in Physical Science.
INSTITUTION	Cooperative General Science Project, Atlanta, Ga.
Pub Date	Aug 69
Note	194p.
EDRS Price	EDRS Price MF-$0.75 HC-$9.80
Descriptors	Astronomy, *College Science, Earth Science, Electricity, Energy, Experimental Curriculum, *Instructional Materials, *Liberal Arts Majors, Matter, Nuclear Physics, *Physical Sciences, *Scientific Concepts
Identifiers	Atlanta, Clark College, Cooperative General Science Program, Georgia

Abstract

Contained in this experimental test are instructional materials for a one-semester course designed to give liberal arts students an appreciation of (1) the nature of science, (2) the development of science, (3) the contributions of scientists, (4) the impact of scientific discoveries on mankind, and (5) the possible future effects of science. The historical and philosophical approach is used to develop basic scientific concepts. Mathematical details and analyses are minimized. The test is divided into the following independent sections: (1) concepts of motion, (2) concepts of matter, (3) concepts of energy, (4) atoms in motion, (5) sound, (6) light, (7) electricity and magnetism, (8) the earth in time and space, and (9) the new physics. Photographs and illustrations accompany the descriptions of various scientific principles in each section. Suggested outside readings and study questions are included at the end of each section. (LC)

CONCEPTS IN

PHYSICAL SCIENCE

Concepts in PHYSICAL SCIENCE

CGSP Staff

CLARK, MOREHOUSE, MORRIS BROWN AND SPELMAN COLLEGES

O. P. PURI, Director
Cooperative General Science Program
Clark College, Atlanta, Ga. 30314
U.S.A.

The following is a list of administrators, staff, and consultants who have helped develop the program of the Cooperative General Science Project. Affilations shown are those prior to their association with the Project.

ADMINISTRATION

Dr. V. W. Henderson, President, Clark College
Dr. Hugh Gloster, President, Morehouse College
Dr. John Middleton, President, Morris Brown College
Dr. Albert Manley, President, Spelman College

DIRECTOR OF COOPERATIVE GENERAL SCIENCE PROJECT

O. P. Puri, Dept. of Physics, Clark College

ADVISORY COMMITTEE

A. S. Spriggs, Dept. of Chemistry, Clark College
J. Mayo, Dept. of Physics, Morehouse College
J. Penn, Dept. of Biology, Morris Brown College
B. Smith, Dept. of Biology, Spelman College

STAFF OF COOPERATIVE GENERAL SCIENCE PROJECT

I. Brown, Dept. of Biology, Spelman College
E. Burt, Macalaster Scientific Company
S. S. Bush, Dept. of Physics, Clark College
A. Dietz, Dept. of Biology, University of Maryland
H. P. Gilman, Dept. of Physics, University of Massachusetts
J. Hannah, Dept. of Science Education, State University of New York, Buffalo
S. Heath, Dept. of Physics, University of California, Berkeley
R. F. Jackson, Dept. of Chemistry, Morehouse College
C. E. Johnson, Dept. of Biology, Univ. of W. Va.
C. S. Kiang, Dept. of Physics, Georgia Institute of Technology
W. Larson, Dept. of Biology, Illinois Institute of Technology
W. B. LeFlore, Dept. of Biology, Spelman College
J. Mayo, Dept. of Physics, Morehouse College
A. G. McQuate, Dept. of Biology, Heidelberg College
R. Miller, Dept. of Biology, State University of New York, Plattsburg
J. F. Moinuddin, Dept. of Health, State of Georgia
S. H. Neff, Dept. of Physics, Morehouse College
J. L. Padgett, Dept. of Chemistry, Morris Brown College
W. F. Payne, Dept. of Biology, Morris Brown College
C. Prince, Dept. of Physics, Brown University
H. Rogers, Dept. of Physics, Emory University
M. Scherago, Dept. of Microbiology, University of Kentucky
K. Shaw, Dept. of Biology, University of Wisconsin
B. T. Simpson, Dept. of Chemistry, Clark College
A. S. Spriggs, Dept. of Chemistry, Clark College
W. P. Thompson, Dept. of the History and Philosophy of Science, University of Indiana
K. K. Vijai, Dept. of Chemistry, Morehouse College

The Cooperative General Science Project has been financially supported by several grants from the United States Office of Education.

G. H. Walker, Dept. of Physics, Clark College
J. Walker, Dept. of Chemistry, Atlanta University
W. Watkins, Dept. of Biology, Howard University
J. D. Wise, Dept of Physics, Clark College
F. Yen, Dept. of Biology, University of Indiana

VISITING LECTURERS AND CONSULTANTS

M. Bell, University of Tennessee
R. Beyer, Brown University
H. Branson, Howard University
N. Byers, University of California, Berkeley
D. Cope, University of Tennessee ORNL
H. Cramer, Emory University
V. Crawford, Georgia Institute of Technology
H. V. Eagleson, Howard University
L. Frederick, Atlanta University
C. Hane, Indiana University
J. L. Haynes, Litton Industries
D. Holden, South Dakota State University
M. T. Jackson, Indiana State University
N. Kowal, Clark College
W. Krogdahl, University of Kentucky
M. Manhas, Steven's Institute of Technology
P. G. Martin, University of Tennessee ORNL
L. Maxwell, U. S. Naval Ordnance, Maryland
K. Nandy, Emory University
R. H. Rohrer, Emory University
F. Rusinko, Clark College
V. Thoren, Indiana University
W. Watson, Yale University
E. Weaver, Atlanta University
L. White, Southern University
S. Winter, State University of New York, Buffalo

SECRETARIAL STAFF

V. L. Merriweather
P. L. Noble
S. A. Spriggs
N. R. Thompson
A. J. Thornton

TABLE OF CONTENTS PAGE

PREFACE Formerly, all of the natural sciences were taught under a course entitled "Natural Philosophy." As our knowledge has increased we have found it necessary to specialize in courses to give an adequate account of the knowledge we have gained. Under the broad headings of physical science and life sciences we now have courses entitled physics, astronomy, geology, biology, botany, zoology, organic and inorganic chemistry and have further specialized these categories with such offerings as meteorology, oceanography, aeronautics, bacteriology, biochemistry, etc.

There are essentially two purposes to be served in offering science courses in college. One is to prepare individuals to function in a specialized field for the benefit of our society; the other is to transmit sufficient understanding of basic scientific concepts to those individuals preparing for non-science vocations that they can be considered educated in our highly technical society. It has become increasingly clear that the conventional science courses offered by most colleges and universities do not fit the needs of many students who are not majoring in the sciences. As our scientific knowledge has increased, there has been a tendency for the departments to become more technical in their preparatory courses in order to give students majoring in a science a firmer foundation in that particular field. Such courses are fine for those students who are majoring in a science, but they do not fit the needs of many students majoring in other fields.

It is of less importance for the non-science majors to master the mathematical approach to the explanation of scientific principles than it is to grasp a verbal and intuitive understanding of the world we live in and the natural laws and principles we have discovered. Many of our leaders in scientific fields have emphasized the need to combine the sciences with the humanities in courses designed for liberal arts students.

Recognizing the need for such a course designed specifically for liberal arts students, we have attempted to put together materials which can fulfill these needs presented in a manner which can be an interesting and rewarding experience for the students. We have concentrated on presenting the basic concepts of science utilizing the historical and philosophical approach. To fully comprehend the significance of our scientific knowledge and its effect on each of our lives requires an understanding of the historical development of these concepts, for we can never truly separate the sciences from the manner in which they have been developed. The historical approach can give us a better understanding of the present and, if we use it as a learning experience, can aid us in preparing for the future.

Many students have developed a mental block against science when it should be realized that science is less a subject area than it is a method of approach in analyzing and understanding the world in which we live. The scientific method can be utilized in many fields whether we are concerned with solving social problems, repairing an automobile engine, learning to play bridge, or improving our sales performance. Possibly, one of the major reasons that a mental block has been formed in many individuals toward science is due to the introduction of new terminology to symbolize certain well defined concepts. This new terminology can be considered a foreign language until the individual becomes familiar enough with it to make it a part of his established vocabulary.

Mathematical symbols and expressions in science are shorthand representations of thought concepts which many people do not take the time to become sufficiently familiar with to automatically associate with these concepts. Instead, they laboriously must translate each symbol into more familiar language just as a person learning a foreign language must do at first. In the same manner, new terminology can be considered a new language, whether it is introduced in a biology course or a course in business administration.

Recognizing the reason for established mental blocks against mathematics and the sciences should be the first step in overcoming these mental blocks. Even though we compare these new words and mathematical symbols to a new language, we must recognize that these symbols are part of *our* language, and if we want to become fully effective in our society then we must make a conscious effort to master our language. These symbols are the major means by which we communicate with each other in our society.

It is not important for many people to master higher mathematics to be fully effective in our society, but it is important that the relationship existing between variables be understood if we are to comprehend the world in which we live. It is important to understand what is meant by directly proportional and inversely proportional, and to comprehend what is meant by an inverse-square law force.

These concepts can be visualized to some extent through the use of graphs and, as an aid in helping the students establish visual models of these relationships, we have made extensive use of graphs in illustrating many of the established principles in physical science. Each individual in our society will probably be exposed to graphs at sometime and it will be beneficial to learn to interpret these graphs since they are a well established method of communicating information.

It is the sincere desire of the staff who developed this text that its use by the students will be an enriching and rewarding experience, and will cause the students to become a little more aware of the exciting nature of the natural sciences which affect our lives each and every day.

ACKNOWLEDGEMENTS This experimental version of *Concepts of Physical Science* represents the efforts and help of many individuals on the staff of the Cooperative General Science Program. While it is realized that some of the staff members contributed more to the development of this work than others, it becomes difficult to give special acknowledgements to a few individuals without slighting others. I wish to express my appreciation to all members of the staff for the interest they have exhibited in this work and the efforts they have expended in its development. A book such as this could not be completed without help from many individuals other than on this staff, and to those individuals I express my grateful appreciation, also.

I would like to express special thanks to the United States Office of Education for the financial support of this project since 1965. It has contributed greatly to the improvement of our academic program and it is hoped that our efforts in producing these materials will be of benefit to other educational institutions across the nation and abroad.

I am grateful, also, to the administrations of the participating colleges - to Presidents Vivian Henderson, Hugh Gloster, John Middleton, and Albert Manley - for their foresight in establishing this program. We owe special thanks to the various chairmen and faculty of the departments who encourage this program and to deans of several colleges for their cooperation.

O.P. Puri
DIRECTOR

Clark College
Atlanta, Georgia
August 1, 1969

FOREWORD TO THE INSTRUCTOR The materials in this text developed from course materials taught over the past three years to liberal arts students in the Cooperative General Science Project. The philosophy of the course has been to engage the student's interest on a conceptual level using historical and philosophical materials to which he could most easily relate and to minimize mathematical details. Equations are only introduced when it is felt that they serve to clarify the concepts or whenever they relate directly to laboratory materials.

The text does not attempt to cover the traditional subject matter usually found in texts of physical science since the course is currently only a one semester course and one must, of necessity, be more selective in the presentation of materials. It is felt that the type of subject materials presented is of less importance than giving the student an idea of the nature of science itself; how it developed; the type of people who contributed to its discoveries; the impact of their discoveries on future generations; what science is doing today; and how it may affect our lives in the future.

In this sense, one may consider this text a guide to the type of course to be presented to liberal arts students and not the course itself. Some instructors may wish to include other materials and delete sections to fit the interests of their students. Each section is designed to stand primarily by itself and should require little additional explanation to convey the concepts contained in each section. It is the consensus of the instructors of the CGSP staff that our students benefit most from the conceptual approach and that mathematical details and analysis can best be handled in the laboratory when the student is engaged in a more quantitative investigation of physical phenomena.

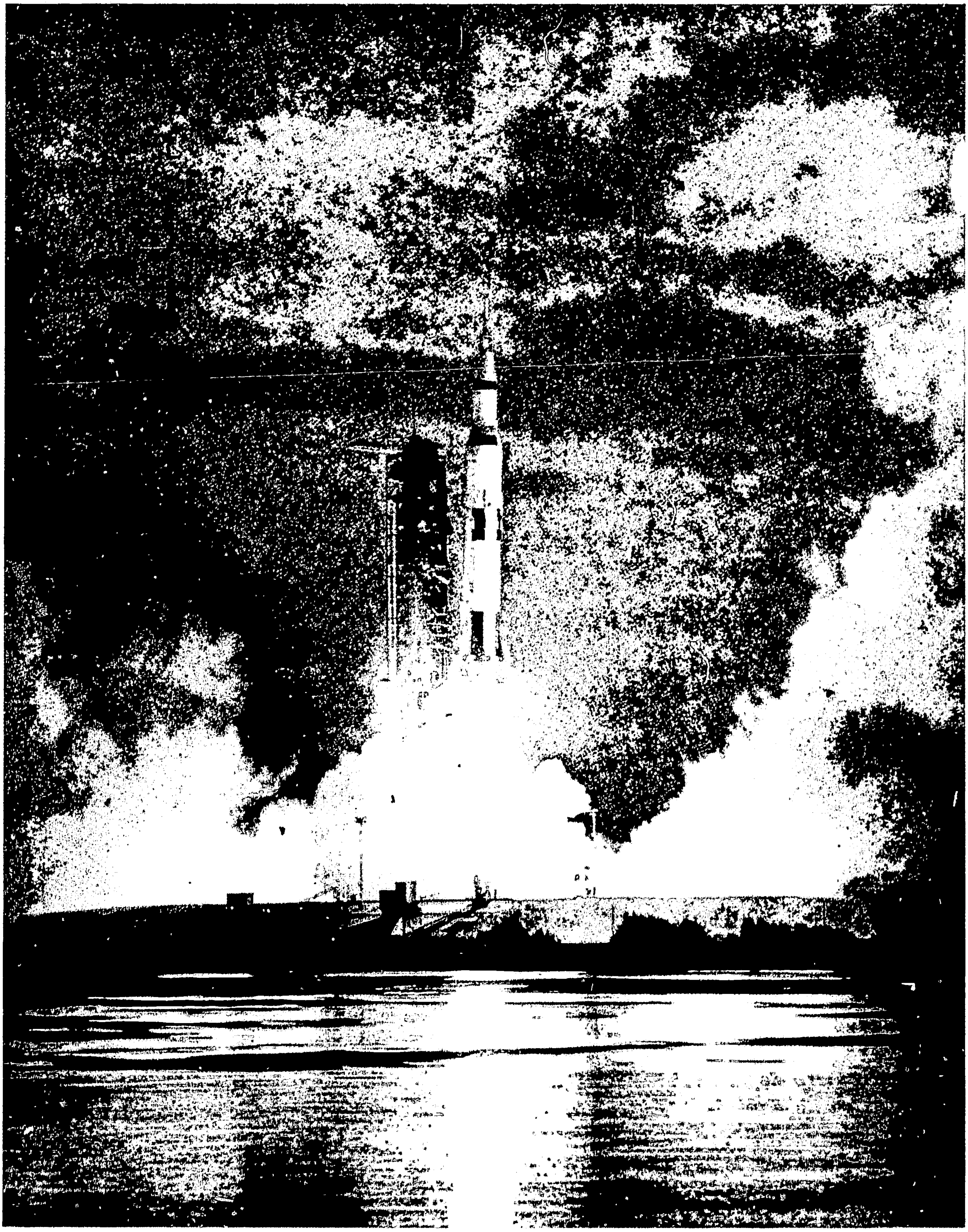
At present, the course consists of two lecture sessions of one hour each, one discussion session of one hour, and one laboratory session of two hours duration. Students who take this one semester course also take a one semester course in CGSP Biological Science to complete their science requirements. The lecture sessions are devoted to presentation of the subject matter, demonstrations, films and guest lecturers. Extensive use is made of the excellent audio-visual aids, such as filmloops, whenever they fit into the course materials. Each semester a field trip and astronomical viewing session is held which the students find very interesting. Students are given a weekly reading assignment and quizzed over these assignments and physical concepts. They are encouraged to read selections from the Suggested Outside Readings and discuss these with the instructor. This text could also serve as the nucleus of a one year course if the instructor wishes to expand the subject materials for this purpose. Sufficient laboratory experiments are also included in the Laboratory Manual for this purpose.

INTRODUCTION - A Challenge for the Future To be alive in the latter half of the twentieth century is to share in some of the greatest achievements of mankind. As this book goes to press a human being has already set foot on the alien, forbidding landscape of the moon, thereby fulfilling a dream of men since time immemorial. Today, we are witnesses to achievements that thirty, twenty, or even ten years ago would have been considered "miracles" by most people of the earth. Laser beams restore eyesight, heart and other organ transplants are almost commonplace, and most diseases that have plagued mankind for ages are either eradicated or brought under control in much of the world. Worldwide communication systems bring international events into millions of homes around the world, and gave millions a front-row seat as an astronaut stepped upon a celestial body. Almost all of these and countless other "miracles" of modern science and technology have occurred in such a relatively short time following World War II that, altogether, they seem to stagger the imagination. The science fiction of a decade ago is reality today, and everyone is influenced in some way by these achievements for either good or bad. Since World War II, we have seen one technological achievement after another being rendered obsolete by yet another. The computer, a child of the late forties, is today in its third generation, and has assumed a position in industry and business undreamed of at its conception. Much of today's technology would be impossible without its development. The computer has revolutionized industry by automation and brought many new industries and businesses into being. Ten years from now it is very probable that you may be working in a field or industry that doesn't even exist today.

How have we progressed so far in science in such a short time, and what will be the effect of this exponential increase in basic knowledge and technology? It is said that the total sum of man's knowledge will double in the next ten years, and double again in less time than that. Someone has said that if a research scientist in any field took time out to read all of the articles appearing in scientific journals in his field and did nothing else, that at the end of a year he would be three months behind. If a research scientist cannot keep abreast with every development in his field, what chance is there for the nonscientist to learn and comprehend anything concerning science today? It should be obvious that he cannot understand all scientific advances, but to be truly educated in this scientific and technological oriented age, the nonscientist must acquire a basic understanding of what science is, what it is doing, and how it affects his life and society. Today we are faced with a formidable multitude of problems, any one of which, if not solved in time, could mean catastrophe for the human race: the problems of nuclear disarmament, overpopulation, air and water pollution, waste disposal, inexpensive sources of energy, adequate housing, and a multitude of related environmental and sociological problems.

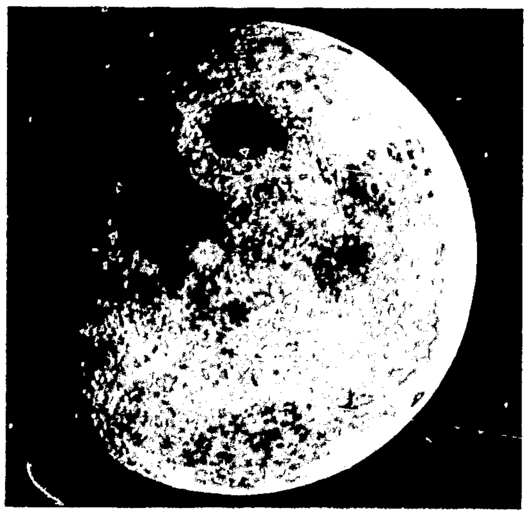
It is often said that a nation that can place a man on the moon should be able to solve all these problems. It has also been said that if a monkey is given a typewriter and an infinite amount of time that he could reproduce all the great literary works by random selection of keys. Unfortunately, we do not have an infinite amount of time in which to solve our problems. In the long run, it is not the scientists that are responsible for the solution of these problems, but the average citizen who votes and voices his opinion on priorities to his legislative representatives. Ultimately, he must decide what programs and what problems should have the highest priority. So it is essential and obligatory that he acquire some basic comprehension of science and its impact on society at large.

It is not surprising that such enormous progress has been made in science during this century, if one considers that approximately ninety per cent of all scientists that have ever lived are alive today! However, today's amazing technological and scientific achievements have not sprung full grown from the head of some mythological being but they have come into being through the combined efforts of many dedicated and ordinary human beings in the past. Each generation of scientists has contributed a small drop of knowledge to that vast reservoir of facts and knowledge that we possess today. They were challenged by the desire to understand and comprehend certain phenomena in nature. These early scientists thought of themselves as *natural philosophers*, they asked questions of nature and sought the answers in a logical manner which we today call the *scientific method*. We did not always have this attitude toward nature; it is the end product of a long evolutionary process that began with the philosophy of the Golden Age of Greece and culminated in the experimentation of the Renaissance. This book is dedicated to the revelation of this development and its impact and influence on today's society.



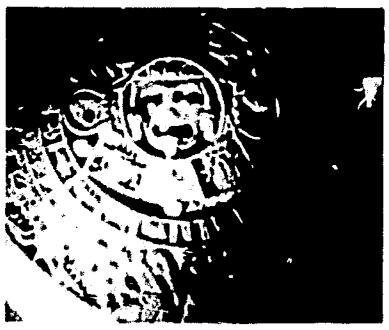
CHAPTER 1 CONCEPTS OF MOTION

APOLLO 8: A Journey to the Moon   On the morning of December 22, 1968, the mighty reaction engines of the Saturn V thundered into life and began to lift millions of pounds of intricate machinery and its cargo of three men into earth orbit in preparation to the first circumnavigation of the moon while millions of people around the world watched spellbound on their home television receivers. Perhaps no other scientific endeavor in history has captured the imagination of the world as has the Apollo mission with its announced goal of placing a man on the moon during this decade. We will not dwell on the event itself, but rather we will attempt to look back and see how we have reached this pinnacle of scientific achievement.



The story of Apollo 8 does not begin with its launch that December day, or even with its conception and design, but rather it begins with the first man who looked into the heavens and wondered what was the nature of the luminous disk he called the moon. Whenever an astronaut sets foot on the moon, it will not be merely an American achievement or a Russian achievement, but it will be a *human* achievement; one in which men over the ages have contributed their share in its fulfillment. Let us look at some of these contributions.

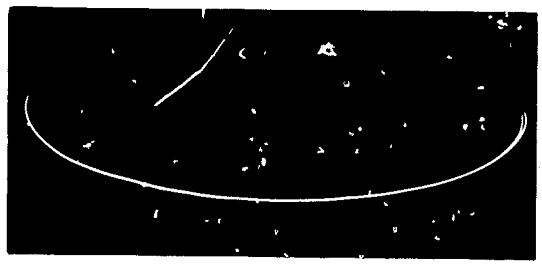
THE BEGINNINGS OF ASTRONOMY   The ancient Babylonians began some of the earliest recorded work in observational astronomy. Their scientific objectives were not aimed at putting a man on the moon, but they contributed even to this in some measure. One of the earliest needs of man was for an accurate means of reckoning the most opportune time for planting crops to insure a good yield. Over the previous centuries man began to acquire a sense of timing geared to the periodic changes of the sun and moon in the sky. He probably began to keep track of certain astronomical events by placing pebbles or sticks in some location to keep an accurate tally of changes occurring in the sky. Some archeologists speculate that Stonehenge in England is one such prehistoric observatory or a device for calculating the seasons. Much of early man's efforts went into such utilitarian endeavors and these often acquired religious significance. Many of the early astronomers were priests who determined the correct date for planting crops and religious events related to good harvests and the like. Numerous examples of such practices can be found in the *Old Testament*. The most significant contributions of the Babylonians was a system of units for the measurement of time and of angles. Today we still use their week of 7 days, the day of 24 hours, the hour of 60 minutes, and the minute of 60 seconds. This system is called a *sexigesimal* scale. Babylonian mathematics abound with such sexigesimal units. The division of the circle into 360 degrees, and the subdivision of the degree into 60 minutes, the minute into 60 seconds of arc is another example. Numerology had a great influence on many ancient peoples and this is illustrated by the system of units in which they measured quantities. There is nothing magical or sacred in the number 60 and its application to temporal and angular measurements today, although some Babylonian astronomer-priests may argue otherwise. Systems of measurement and units differed from people to people and resulted in numerous systems coming into being in different locales and ages. The Babylonian system was one such system. It was no more *correct* or absolute than others, but it has merely stood the test of time as a convenient system. Frequently in the history of science, units and systems of measurement have been revised, rejected, and replaced. The advantage of one system over another is whether it is convenient to use and serves its purpose better than the other. Today's astronauts use the same method of determining time and measuring angles in navigation as a Babylonian astronomer-priest of 5000 years ago, because no one has come up with a better or more convenient system.


Aztec Calendar

The calendars of the ancient Babylonians and Egyptians were quite precise and have come down to us today in a modified form. Other ancient cultures also developed calendars of one type or another. The Aztecs of South America developed a calendar based on eighteen months of twenty days each. The American Indian reckoned time by the number of "moons" - the lunar month. Today these concepts of time are so ingrained in our culture that we tend to take them for granted without questioning their origins or validity.

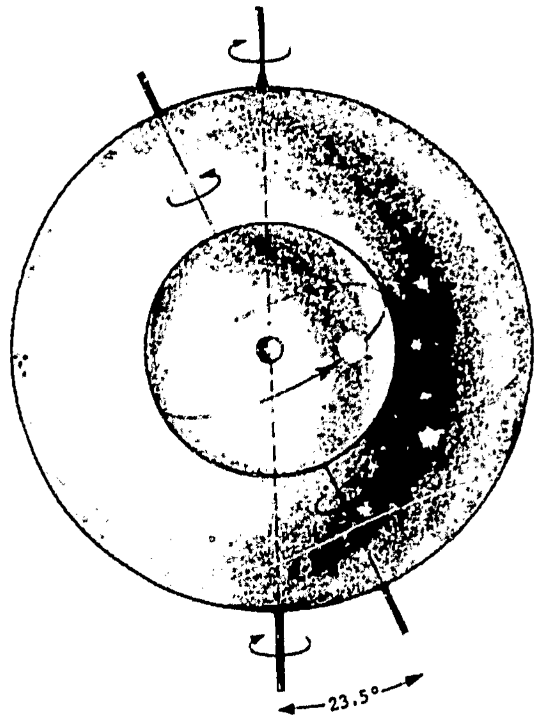
The science of the Babylonians and Egyptians was of a strictly utilitarian nature so we must move on to the Greeks for an explanation of motion in the heavens. The Greeks were given more to speculation and philosophy than were the Egyptians. They were more interested in how the heavenly bodies behaved and why they behaved the way they did. They were keen observers of nature but their love of argumentation and philosophizing made them poor scientists. In mathematics, logic, and the arts, the Greeks showed amazing creative genius, and it is for these things that we are indebted. Even though the astronomy and science of the Greeks were not experimental in nature, they did contribute a great deal to our cosmological concepts.

The Greeks were fascinated by the idea of perfection in nature and in mathematics. To them the circle represented the ultimate perfection. Careful observation of the heavens led them to believe that the stars revolved about the earth in uniform circular motion. The sun and moon also appeared to move in a circular motion, however; their motion differed from that of the stars. The planets appeared to behave most illogically of all the heavenly bodies and on occasions even appeared to move backwards! This type of motion is referred to as *retrograde* motion.



Retrograde motion seen from an angle.

The famous Greek philosopher, Plato, in the fourth century B.C. asked his students if they could devise a theory or explanation to explain this erratic planetary motion using some form of circular motion. Being keen observers, the Greeks came up with the most logical and obvious conclusions; namely, that the earth was the center about which the sun, the moon, planets, and the stars rotated. This model of the universe is called a *geocentric* or earth-centered model. It satisfactorily explained the daily motion of the stars and sun by assuming that they were attached to invisible crystalline spheres that rotated about the earth. The axis of the sphere of the sun was tilted with respect to that of the stars to account for the variation of the sun's height at zenith with the various seasons. Since the sun appears to move through the stars and was brighter, it was assumed to be nearer to the earth than the stars. The spheres of the moon, Mercury, and Venus were placed within the sphere of the sun while those of Mars, Jupiter, and Saturn were placed outside the sphere of the sun but within the sphere of the stars.

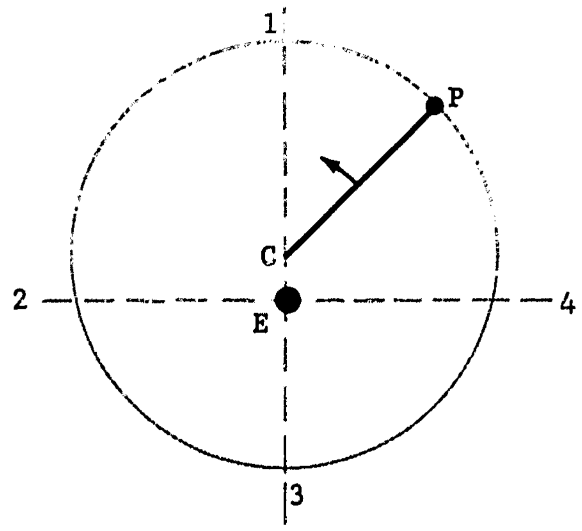


Aristotelian concept of the universe. The sphere of the sun revolves about the earth with its axis inclined at an angle of 23.5° to that of the earth. The sphere of the stars revolves about the earth once daily.

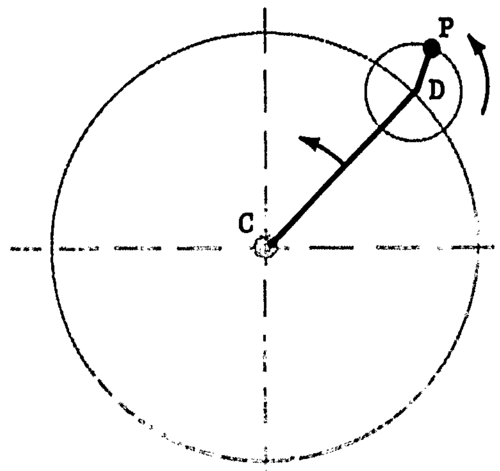
By selecting the correct rotational rate, angles of inclination, and distances, the proper motion could be closely approximated. Still, the model would not account for all planetary motions accurately enough.

Euxodus, a contemporary of Plato, improved on this system by increasing the number of spheres to 27 to account for variations in the observed motions. Later, Aristotle added about 29 more spheres to gain increased accuracy, but still, inconsistencies arose after a long while.

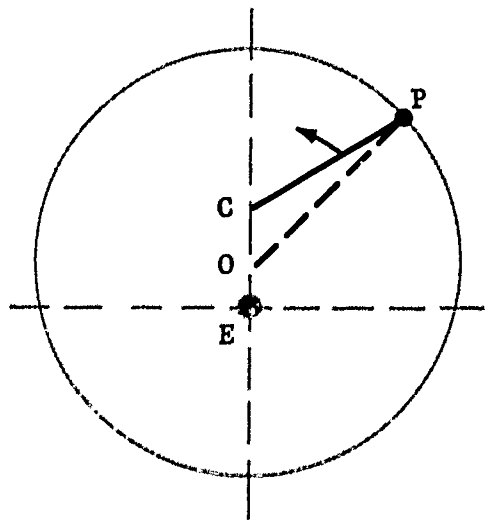
In the third century B.C., Aristarchus came up with an alternative model for the observed motions which placed the sun at the center of rotation with the earth, moon, planets, and stars revolving about it. This type of model is called a *heliocentric* or sun-centered model. Aristarchus' model had the advantage that it could easily explain the retrograde motion of the planets and the daily motion of the sun and stars if he further assumed that the earth rotated once daily on its axis. However, instead of being accepted by his contemporary philosophers, it was bitterly



An eccentric. The planet P revolves about the center C, with the earth, E, off center.



An epicycle. The planet P revolves about the center of the epicycle D, which revolves about C.



An equant.

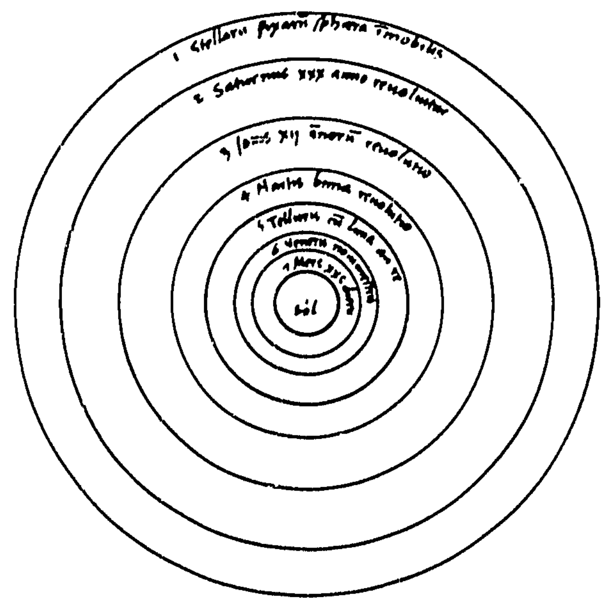
attacked because it implied that the earth moved which most Greeks considered an absurdity. It also conflicted with their preconceived idea of perfection so it died a premature death only to be resurrected about 1800 years later by the Polish astronomer Copernicus.

THE PTOLEMAIC GEOCENTRIC MODEL About the middle of the second century A.D., the Greek astronomer Claudius Ptolemy of Alexandria developed a geocentric model that was to be accepted until the mid-fifteenth century. His model differed from the concentric spheres of Aristotle in that the earth was placed slightly off center to account for the fact that the sun and planets appeared to move through certain angles in the sky in unequal number of days. This type of arrangement, Ptolemy called an *eccentric*. Although Ptolemy did not reject the Aristotelian model entirely, he did compromise it to the extent that the earth was not at the exact center of rotation. This improvement helped to explain small variations in the cycles of the sun and planets, but it did not explain the retrograde motion of the planets. To do this, Ptolemy introduced another innovation, the *epicycle*.

The epicycle was a circle whose center lay on the circumference of a larger circle or eccentric. The planet moved around the center of the epicycle while the center of the epicycle moved around the center of the eccentric. By choosing the correct ratios of rotation of the epicycle and eccentric, various types of planetary motion could be explained. Still, discrepancies began to show up with this arrangement so Ptolemy further modified his model by introducing a device he called an *equant*.

The equant was a modified eccentric which placed the earth off center on one side and the planet rotating off center on the opposite side. By judicious selection of uniform rotational rates, distances, and sizes of epicycles, Ptolemy was able to account for the observed motion of the planets with amazing accuracy. Because his complicated geometrical model did fit the observed data so well, it was not questioned until the time of Copernicus to any great extent. Several reasons why it was accepted so unquestionably were that it fitted the philosophical thought of the times, appealed to the common sense notions, and fitted into the Church's view of man's place in the universe that arose in the Middle Ages.

THE COPERNICAN HELIOCENTRIC SYSTEM In 1543 a Polish monk and astronomer named Nicolaus Copernicus published a book entitled *De Revolutionibus Orbium Coelestium* or *On the Revolutions of the Heavenly Spheres* in which he advanced his heliocentric theory of the universe. Basically, his system was similar to that of Aristarchus but he developed it in considerable mathematical detail as did Ptolemy. The accuracy of the two systems was comparable but Copernicus' system offered extra advantages; it was simpler, and provided a means of calculating the distances of the planets from the sun which Copernicus did with amazing accuracy. Copernicus believed his system in its symmetry and order to be closer to the architecture of God than that of Ptolemy. We will see that this belief in symmetry and order is an underlying premise of modern science and has led to many important discoveries in physics. Some physicists today rank symmetry considerations almost as high as conservation principles.



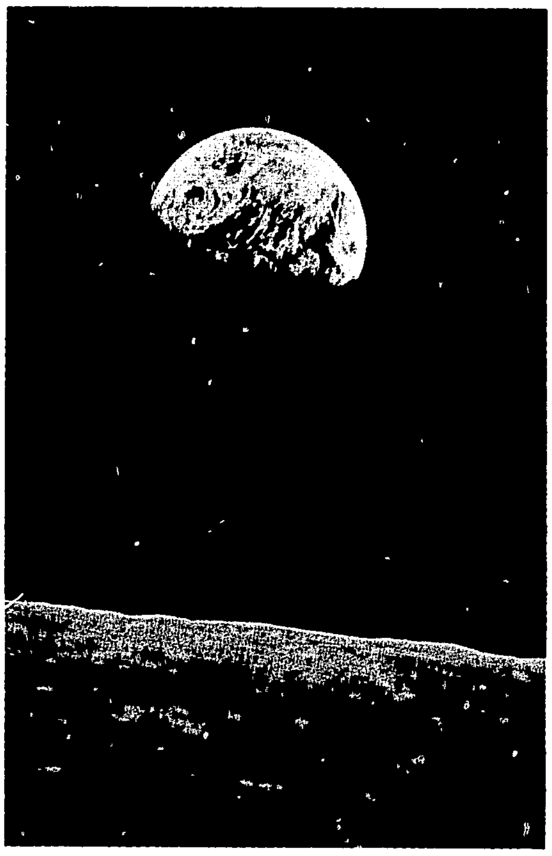
Copernican heliocentric system.

Copernicus' theory was bitterly attacked for philosophical and religious reasons by his contemporaries. Over the centuries, the Church had acquired as part of its tradition the philosophical outlook of Aristotle and had interwoven scientific views with its religious views and dogmas. To suggest that the Ptolemaic-Aristotelian viewpoint was erroneous was paramount to questioning the teaching of the Church and the validity of the Scriptures. For did not Joshua command the sun to stand still when he fought the Amorites:

> "Sun, stand thou still at Gideon, and thou
> Moon in the valley of Ai'jalon." And the
> Moon stayed,
> until the nation took vengeance of their
> enemies.
>
> *Joshua* 10:12-13

Time and time again we will see a similar reaction from religious quarters whenever scientific truth clashes head-on with preconceived pseudo-scientific dogma. In our own time we have witnessed the repeal of the laws against the teaching of the theory of evolution which swept the country in the early 1920's. As G.B. Shaw so aptly put it on an occasion honoring Albert Einstein in London in 1930:

> "Religion is always right. Religion solves every problem and thereby, abolishes problems from the universe. Religion gives us certainty, stability, peace and the absolute. It protects us against progress which we all dread. Science is the very opposite. Science



Earth seen from Apollo 8.

is always wrong. It never solves a problem without raising ten more problems."

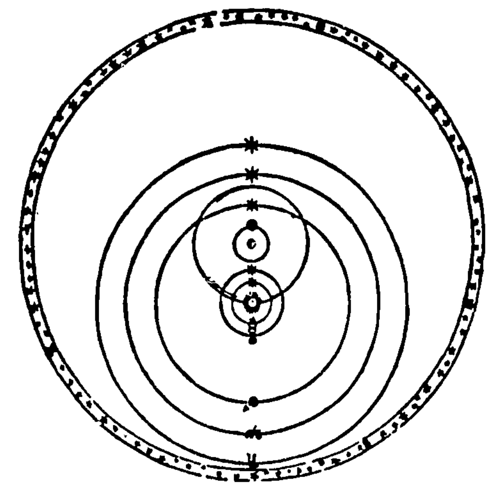
For someone to accept the Copernican system meant that Man was no longer at the center of creation, but relegated to a minor planet of the sun. To Copernicus' contemporaries, this thought was unthinkable and offensive. Contrast that to the magnificent views of the earth taken from Apollo 8, and the crew's expression of awe and humility. Someone has said that the most shattering experience the human race will have as it ventures out into space will be when it comes face-to-face with another intelligent being. We may not see this day, but it probably surely will come.

While Copernicus did not live to see his system accepted as fact, it eventually was, in a slightly modified form, as more accurate observations of the motion of the planets became available for astronomers to make a choice between the Ptolemaic and Copernican systems. For these observations we are indebted to a Danish astronomer, Tycho Brahe, born three years after the death of Copernicus.

TYCHO'S CONTRIBUTION Often in the history of science, one event or discovery stands out as a turning point. The work of Tycho is such a pivotal point. Today one hears much concerning the value of basic research or research for the sake of knowledge alone when it is obvious that there are so many practical and applied technological problems to be solved. Often such basic research is like the proverbial iceberg; only about one-tenth of its value is visible at one time; often its true value is obscured from view, waiting for the right person to reveal it. The work of Tycho is a case in point.

The Ptolemaic and Copernican systems had one characteristic in common; they both attempted to explain the *observed* motion of the planets. In order for either of them to be valid they would have to verify the observed motions and be able to predict future motions of the planets. This both systems did to about the same degree of accuracy and herein lay the crux of the problem. The data of the planetary positions that Ptolemy and Copernicus used to formulate their systems were not accurate. Tycho recognized this and resolved to rectify the situation by gathering as accurate data concerning planetary motions as he could. King Fredrick II, recognizing Tycho's potential, offered him an observatory of his own and sufficient funds to maintain it. Tycho began work in the observatory which he named Uraniborg or "Castle of the Heavens" about 1576 and conducted painstaking measurements of the positions of planets there until about 1597. In order to make better measurements Tycho designed and built many elaborate and accurate sighting instruments and took care to calibrate them to determine his limits of error, or the degree of precision that he could assign to his work.

After the death of King Fredrick II, Tycho moved his work to Prague, where he found finanicial support and a young mathematician named Johannes Kepler to assist him in his work. Kepler was assigned the task of analyzing Tycho's data and trying to fit it into a planetary system. Tycho himself had developed a system that was a hybrid of the Copernican and Ptolemaic systems. He placed the earth at the center of his system, but he had all of the other planets revolving around the sun, which in turn revolved around the earth, which he considered to be stationary.



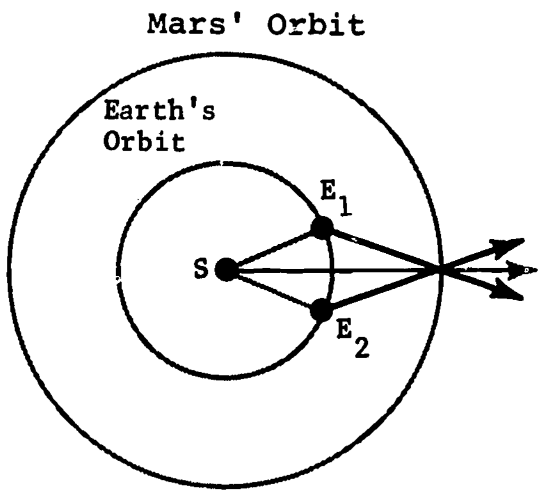
Tycho's compromise system. The sun revolves about the earth with the other planets revolving about the sun.

Some of Tycho's contemporaries accepted his system while others rejected it, mainly for the same reasons they accepted or rejected the Copernican and Ptolemaic systems. Tycho died in 1601 before he saw the fruits of his labors reach maturity, but he could not have left his work to more capable hands than those of Kepler.

KEPLER AND THE NEW ASTRONOMY Quite often in the development of science, one person or other has set out to prove a theory or preconceived idea only to be sidetracked into discovering some greater truth. Such was the case with Kepler. To Kepler, the Greek notion of perfection in mathematics and nature had great appeal and motivated him in his work. The mysticism that the Pythagoreans felt toward mathematics and its ability to provide them an understanding of nature is evident in much of Kepler's work. In his first published work in 1597, *Mysterium Cosmographicum*, he states:

> "The orbit of the earth is a circle; round the sphere to which this circle belongs, describe a dodecahedron; the sphere including this will give the orbit of Mars. Round Mars describe a tetrahedron; the circle including this will be the orbit of Jupiter. Describe a cube round Jupiter's orbit; the circle including this will be the orbit of Saturn. Now inscribe in the Earth's orbit an icosahedron; the circle inscribed in it will be the orbit of Venus. Inscribe an octahedron in the orbit of Venus; the circle inscribed in it will be Mercury's orbit. This is the reason of the number of planets."

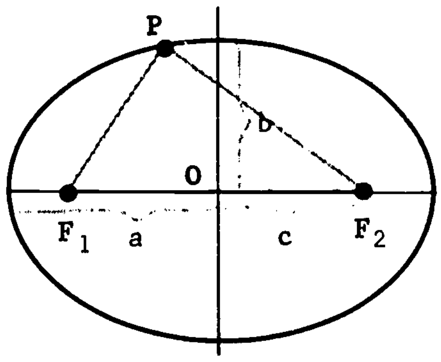
Such mystical reasoning may appear alien for a scientist of Kepler's stature, but we must remember that he was a product of his age as we are products of ours. Kepler believed that God created the world according to some mathematical plan and that nature exhibited these mathematical harmonies in the order of the heavens.



Determination of Mar's orbit by triangulation.

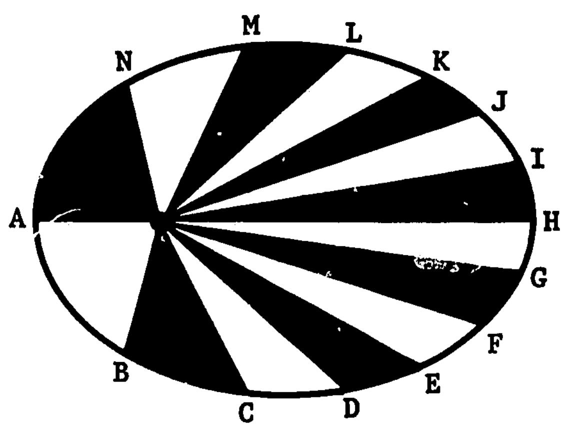
One of his favorite quotes was one attributed to Plato: "that God ever geometrizes." We again see this same thought appearing in this century in the work of Albert Einstein, who created another universe of space-time out of pure geometry and who proved that our universe was a rational and well-ordered universe imbued by a Supreme Intellect. Einstein's own belief in cosmic order is illustrated by his statement: "The Lord does not throw dice."

Kepler's geometrical system fortunately, did not fit the observed data that Tycho so painstakingly accumulated concerning the orbit of Mars. Kepler realized that it was not Tycho's data that was in error, but his attempt to fit it into a preconceived system. Once he realized this, he was able to make progress.



An ellipse.

By correlating Tycho's data concerning the relative positions of the sun, earth, and Mars at different times, Kepler, by a process known as triangulation, was able to derive the orbit of the earth and later to derive the orbit of Mars. To his astonishment, the orbit of the earth turned out to be almost a circle, with the sun slightly off center. The orbit of Mars, by contrast, turned out to be oval-shaped. Kepler now began to search to find the exact type of curve the oval orbit of Mars best approximated. He tried various shaped ovals and finally found that an ellipse gave the best fit. Closer inspection showed that the earth's orbit was also elliptical with the sun located at one focus. This resulted in what today is referred to as Kepler's First Law of Planetary Motion: *the planets move about the sun in elliptical orbits, with the sun at one focus.*



Law of Areas. A planet sweeps out equal areas in equal times.

Prior to the discovery of his first law, Kepler accidently discovered a fact that is today known as Kepler's Second Law of Planetary Motion: *during equal time intervals, a line connecting the planet to the sun sweeps through equal areas.* This law is commonly referred to as the Law of Equal Areas, and means that as a planet approaches closely to the sun, it speeds up. As it moves away from the sun, it begins to slow down. This result was puzzling to Kepler and his successors and remained a mystery until Newton explained *why* the planets behaved the way they did.

One of Kepler's earlier goals was to relate the orbits of the planets to some kind of mathematical regularity or harmony. This search for mathematical order acquired the status of almost an obsession with him until finally, in 1619, he made the discovery of his long, elusive goal - his Third Law of Planetary Motion, the Law of Periods. In his book, *Harmony of the World*, Kepler expounded on this relationship and even went so far as to compose tunes or the "music of the spheres" that he believed they emanated. This theme of the "music of the spheres" is found in various later works of literature, one of the most famous occurring in the Prologue in Heaven of Goethe's *Faust:*

"The sun intones, in ancient tourney with
Brother-spheres, a rival song, fulfilling
Its predestined journey,
With march of thunder moves along."

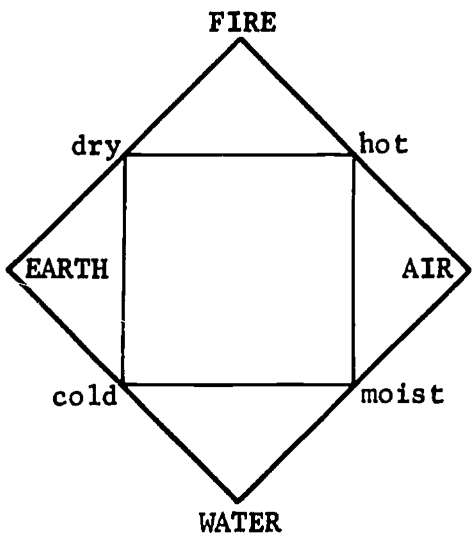
The Law of Periods gives the relationship between the time required for the planets to revolve about the sun and their mean or average distance from the sun. It states: *the square of the times required for the planets to complete one revolution of the sun is directly proportional to the cubes of their mean distance from the sun.* Stated in terms of symbols:

$$T^2 \propto R^3_{av} \text{, or } T^2 = KR^3_{av} \text{,} \qquad (1.1)$$

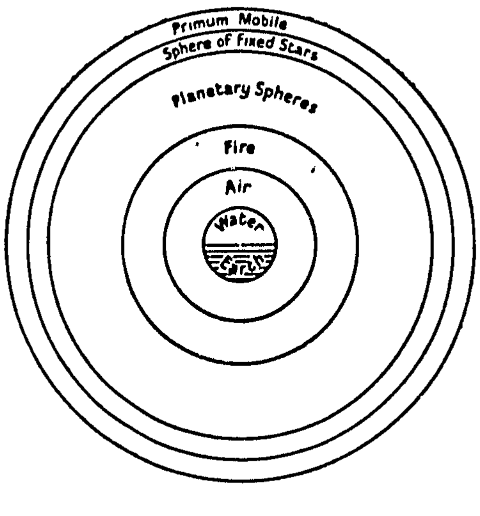
Where T is the period of revolution of the planet and R_{av} is the mean distance of the planet from the sun. The symbol K is a constant of proportionality, and has the same value for all planets. What this relation means is that we can determine the orbit of a planet if we know its period of revolution, or vice-versa.

In Kepler's three laws, we now have the solution to Plato's problem of planetary motion, not in terms of uniform circular motion, but in terms of nonuniform elliptical motion. Still, we do not know *why* the planets move this way; we only know how to *describe* their motion. Kepler's three laws are *empirical* rules based on factual observations which are useful in predicting the future position of the planets, but they cannot by themselves explain why the planets behave this way. Kepler himself attempted to answer the question, "why?" but failed. After reading Sir William Gilbert's work on magnetism, Kepler speculated that perhaps some kind of magnetic force from the sun was responsible for the motion of the planets, but he did not see his ideas reach fulfillment. Before this important step in determining a *causal* relationship could be taken, certain experiments had to be performed by Galileo to lay the necessary foundation in mechanics, or the science of motion. We will turn our attention to these temporarily.

GALILEO AND THE NEW PHYSICS OF MOTION Although Galileo and Kepler were contemporaries, they differed considerably in their scientific methodology and philosophy. We have already discussed Kepler's obsession with numerical mysticism and his Neo-Platonic quest for perfection in the heavens. Kepler was a product of the medieval school of Scholasticism and Neo-Pythagoreanism, while Galileo, although a believer in Neo-Platonism to an extent, was primarily a mathematical experimentalist. To Galileo, Nature was an orderly, mathematical system, but his method of discovering her secrets was one of experimentation and not one of scholastic logic. His method involved the observation of natural phenomenon, its translation into mathematical language, deduction of the behavior of the phenomenon by mathematical reasoning, and verification of these deductions by experimentation. Because he was one of the first to utilize this technique, Galileo is regarded as the founder of modern experimental science.



Before we proceed to relate the experiments of Galileo that were to have such a profound influence upon future scientific development, let us review briefly the scientific views concerning motion that were prevalent in Galileo's day. Beginning with the rediscovery of the works of the Greek philosophers, in the thirteenth century A.D., there developed a desire to assimilate into Western culture the great lost works of the ancients. Aristotelian philosophy and logic were not altogether unknown to the medieval Church which played a major role in preserving some ancient manuscripts and works following the decline of Rome. During this time, however, it emerged as an integral part of Christian theology instead of a separate area of study.

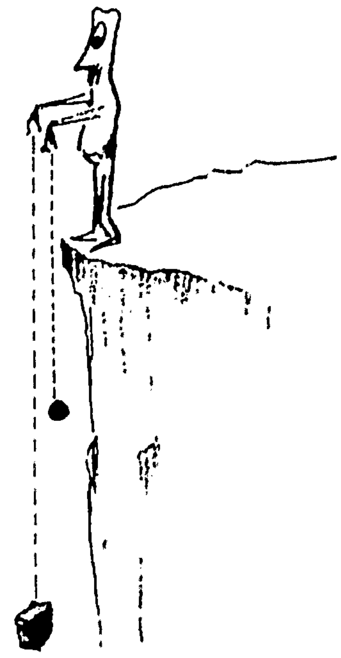


Aristotelian Elements.

In the writings of St. Thomas, Aristotelian thought was incorporated into Christian theology and by the thirteenth century had become the authority for religious and secular philosophers. Against such a backdrop, Galileo enters upon the stage.

The medieval philosophers separated objects into two catagories - those terrestrial and those celestial. Terrestrial matter was believed to be composed of four elements - earth, water, air, and fire. Each element was assigned its "natural place" on the earth, beginning with earth at the center and proceeding outward. These were water, air, and fire. Above fire existed the celestial spheres consisting of the planets, sun, and stars, and above these the Abode of God. Change was characteristic of the terrestrial region while the celestial region was unchanging and eternal. The falling of a terrestrial object was considered to be a natural motion if the object sought its own natural place.

For example, a stone dropped into water sinks to the ground which is its natural place. A stone thrown into the air is not in its natural place, and such motion was termed "violent" and required a violent force to cause it to move upward away from its natural place. To explain such phenomena as the flight of an arrow through the air, Aristotle asserted that the air displaced by the head of the arrow moved to the rear and exerted a force on it, pushing it forward. Aristotle explained the increasing speed of a falling object by asserting that as it approached closer to its natural place, it moved faster. In a similar argument he explains that a heavy object will fall faster than a lighter one because of its desire to reach its natural place. These are basically the Aristotelian concepts of motion that Galileo set out to attack, and in doing so, he eventually toppled the whole Aristotelian cosmology and its theological support.



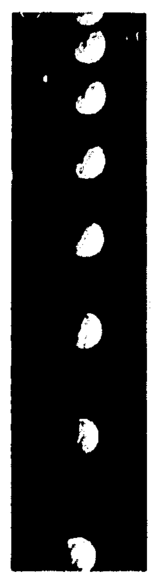
Aristotelian concept of motion. A heavy object falls faster than a light object.

Galileo began his attack by publishing a book entitled: *Discourse and Mathematical Demonstration Concerning Two New Sciences Pertaining to Mechanics and Local Motion,* which we will hereafter refer to as *Two New Sciences*. In the *Two New Sciences*, Galileo employed the Platonic dialogue form to present the Aristotelian viewpoint and his viewpoint to a third party who is uncommitted in his views. Much of the dialogue is aimed at ridiculing the Aristotelian concept of falling bodies which Galileo does superbly.

Later, Galileo defines what he means by concepts such as *uniform acceleration* which he believed falling bodies experienced. Prior to this, this concept had rather vague meanings and now Galileo gives it a precise mathematical definition. Galileo defined *uniformly accelerated motion* to be *motion that a body acquires when starting from rest which gives it equal increments or changes in speed for equal time intervals*. Today we would express this in symbols as:

$$a = \frac{\Delta v}{\Delta t} \qquad (1.2)$$

where a is the uniform acceleration, Δv is the change in speed, and Δt is the change in time during which the speed is changing from some initial value v_i to some final value v_f. If we let $\Delta v = v_f - v_i$, and $\Delta t = t_f - t_i$, then Eqn. (1.2) becomes:



A uniformly accelerated object.

$$a = \frac{v_f - v_i}{t_f - t_i} \qquad (1.3)$$

Unfortunately, Galileo was unable to experimentally verify this relation by direct measurement. He lacked the precise instruments that we have today to determine small time intervals and it is extremely difficult to determine the final speed of the object before it struck the ground. Galileo, however, was foremost an experimentalist. So he attempted to approach the verification of this relation from a different viewpoint. Since he could not directly measure the final speed of the falling object, he attempted to develop other relationships between variables that he could measure and the final speed. These other variables were the *distance* the object falls and the *time of fall*. In order to do this, he first had to define some other quantities and derive relationships between them and the final speed. Galileo defined a quantity called the *average speed*, v_{av} as the total distance, *d*, that the object falls (or travels) divided by the time of fall (or travel), t. Symbolically:

$$v_{av} = \frac{d}{t} \qquad (1.4)$$

If an object is moving with uniform acceleration, one can also determine its average speed by summing its initial speed and its final speed and dividing by two, or

$$v_{av} = \frac{v_f + v_i}{2} \qquad (1.5)$$

This expression for determining average speed is similar to one for determining the average value of any two quantities. For example, if we wish to determine the average price of two books, one costing $2.00 and the other $3.00, we sum the two prices and divide by two, or:

$$\text{Average cost} = \frac{\$2.00 + \$3.00}{2} = \$2.50$$

If the falling object starts from rest, its initial speed, v_i, is zero, so Eqn. (1.5) would reduce to the expression:

$$v_{av} = \frac{1}{2} v_f \qquad (1.6)$$

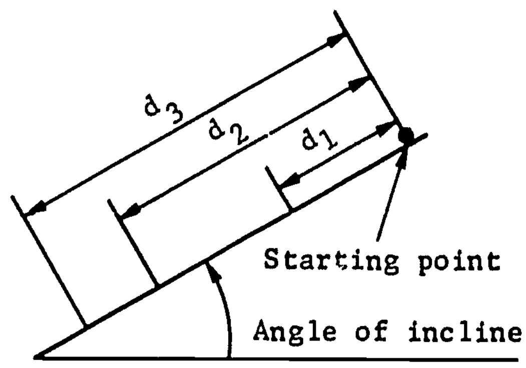
Now, by simple algebraic substitution of Eqn. (1.6) into Eqn. (1.4), and the substitution of the modified Eqn. (1.4) into Eqn. (1.3), one can eliminate the final speed v_f from Eqn. (1.3) and arrive at the desired result that Galileo was seeking:

$$d = \frac{1}{2}at^2 \qquad (1.7)$$

The derivation of Eqn. (1.7) is left to the student as an exercise. This expression relates the uniform acceleration of a falling body to the distance it falls and its time of fall.

Galileo still lacked an accurate means of measuring the time interval for a freely falling body, so he did something that most scientists today probably do without thinking about; he performed an experiment for a *special* case and extrapolated or extended it to the *general* case. He reasoned that since a freely falling object experienced a maximum uniform acceleration and an object resting on the earth experienced no acceleration, that an object rolling down a smooth plane, inclined at some angle to the horizontal, must experience an acceleration which lies between these two limiting cases, and should be a function of the angle of inclination.

Galileo performed this experiment for a given angle of inclination and found that the ratio of the distance that the ball rolls down the plane to the square of the time required for it to travel this distance was a constant for various distances of travel, or:



$$\frac{d}{t^2} = \text{constant.} \qquad (1.8)$$

Today, we would say that the distance is directly proportional to the square of the time, or:

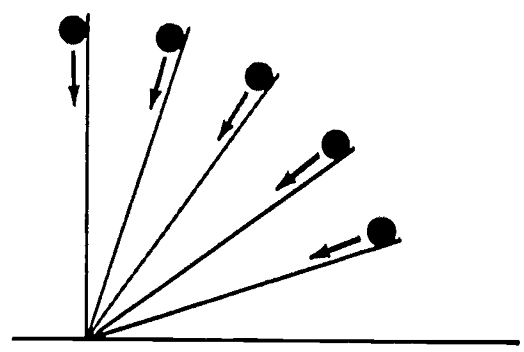
$$d \propto t^2 ,$$

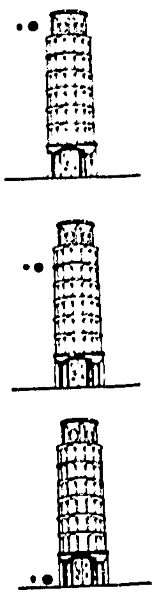
or

$$d = kt^2 , \qquad (1.9)$$

where k is a proportionality constant. We see by comparison of Eqn. (1.9) with Eqn. (1.7) that k must be equal to $1/2\ a$, or that the acceleration is twice the proportionality constant.

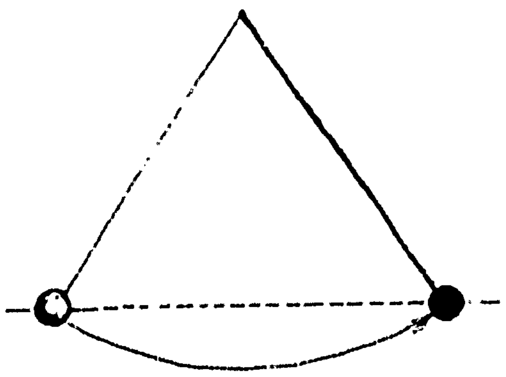
Galileo performed this experiment for various angles of inclination and discovered a new proportionality constant for each angle. He observed that as the angle of inclination with the horizontal increased, the acceleration also increased, and he concluded that when the angle of inclination is equal to 90 degrees the acceleration would be the same as for a freely falling body. This technique of analysis by extrapolation has become one of the most useful methods of determining relationships between variables that are inaccessible to direct measurements.



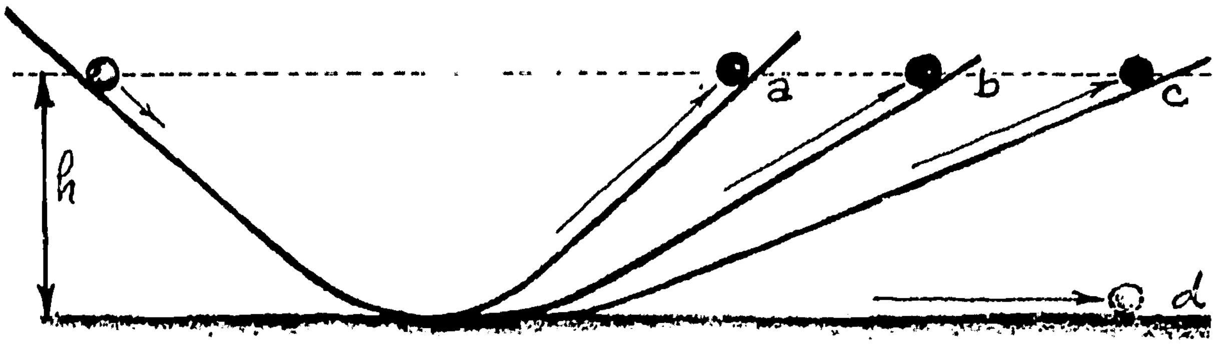


The importance of Galileo's work with falling bodies is found in its proof that all bodies, regardless of their weight, fall with the same uniform acceleration. Whether or not Galileo actually dropped two bodies of different weight from the Leaning Tower of Pisa to prove this point is still open to debate; however, his inclined plane experiments offered enough decisive proof to eventually refute the Aristotelian physics of motion.

Galileo also studied other types of motion such as the motion of the pendulum and projectiles. It is said that while praying in the cathedral of Pisa, his attention was drawn to the swinging of a suspended lamp. He timed its period of oscillation with his pulse and found it to be approximately a constant. It was his study of the pendulum that led him to the conclusion that the Aristotelian notion that a force must be continously applied to an object to keep it moving was erroneous. He observed that the pendulum bob rose to approximately the same height at the end of its swing as the height from which it was released. Using this observation, Galileo devised an imaginary experiment in which he envisioned a ball rolling down a frictionless plane and up another, inclined at the same angle to the first, in analogy to the motion of the pendulum bob. He concluded that the ball would rise to the same height each time regardless of the angle of inclination of the second plane. For the case when the second plane is horizontal, Galileo concluded that it would roll on forever in a straight line since it could not reach the original height from which it was released. This is the beginning of the concept of inertia of a body, or the tendency of a body to persist in its state of motion. We will see that Newton will develop this concept more fully later.



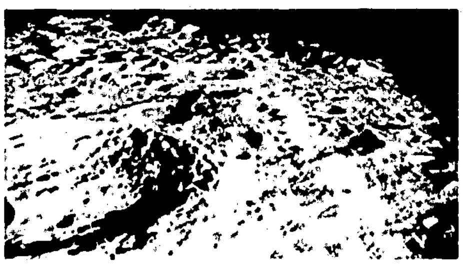
A pendulum bob rises to the same height from which it is released if air resistance is neglected.



An object released from height h will rise to height h if friction is neglected.

Although Galileo's interest in mechanics was extensive throughout his life, his chief interest lay in challenging the Ptolemaic-Aristotelian world system and championing the Copernican. He probably did not realize, however, the type of Pandora's box he opened when he turned his telescope skyward in 1609. Everywhere that he looked he saw irrefutable evidence that the Aristotelian-Platonic concepts of celestial perfection were false. Turning first to the moon, he observed that:

> ...the surface of the moon is not smooth, uniform, and precisely spherical as a great number of philosophers believe it to be, but uneven, rough, and full of cavities and prominences, being not unlike the face of the earth, relieved by chains of mountains, and deep valleys.



His observations of the sun, planets, and stars showed him things were not as Aristotelians claimed. Observations of the great planet Jupiter, with its four circling moons, proved conclusively to Galileo that the earth was *not* the center of the universe about which everything revolved. In 1610 he published a book concerning his observations entitled *Sidereus Nuncius*, or *Message from the Stars*, which brought him both fame and suspicion. In 1616, opposition from his Aristotelian colleagues led to his admonishment not to teach the Copernican theory. While Galileo publicly taught the Ptolemaic system to his classes, he privately worked on his treatise, *Dialogues Concerning the Two Great Systems of the World*, which he published in 1632, with the approval of the Church. The book, a skillful argument in favor of the Copernican system, brought prompt protests from the Aristotelians, who succeeded in having him brought before the Inquisition in Rome on a charge of heresy. Galileo was placed under house arrest and persuaded to recant his Copernican belief.

Throughout history men of science and champions of truth have had to endure various types of persecution and ridicule in support of their ideas and beliefs. The counterpart of the medieval court of Inquisition rears its ugly head in every age to stamp out ideas it finds a threat to its dogmatism and prejudices. We can find numerous examples in our time: Freud, for his use of psychoanalysis in the treatment of the mentally disturbed and for his theories of infantile sexuality; Einstein, whose scientific theories were banned in Nazi Germany because of his Jewish ancestry; John Scopes, who was convicted of teaching Darwinian evolution in the public schools of Tennessee; and the denial of security clearance to Robert Oppenheimer because of his views on thermonuclear weapons development. Still more recently, we have observed the debate and criticism concerning the morality of heart and organ transplants. Science can only progress whenever there is freedom to question and explore the unknown. Perhaps this is why America is the undisputed scientific and technological leader in the world today.

NEWTON AND THE MECHANISTIC UNIVERSE Perhaps no person in the history of science is more revered than Sir Isaac Newton with the possible exception of Albert Einstein. Both men possessed that rare quality of genius that set them apart from their fellow men, and at the same time, their humility endeared them to generations to come. The parallel between the two men is strikingly similar: both were creators of universes and a revolution in man's concept of his world and nature that reached far beyond the field of science into the political, economic, and social arenas; both men made their great discoveries at approximately the same age; both were mathematicians par excellence; and both sought unity and mathematical order in Nature with a religious zeal and reverence. Perhaps George Bernard Shaw paid them the finest tribute when he said:

> "Napoleon and other great men of his type were makers of Empires. But there is an order of man who gets beyond that. They are makers of universes and as makers of universes their hands are unstained by the blood of any human being."

The contributions of Newton in different areas of physics are manifold, but his greatest contribution was in the synthesis of terrestrial and celestial mechanics. We have already seen some of the groundwork laid down by Kepler and Galileo toward this goal. Now, let us see how Newton fused these works into a unified theory of the universe.

Newton is perhaps best known for his three laws of motion and his Law of Universal Gravitation than for his other contributions, but these things did not come into beginning with him alone. We have seen how Galileo began to introduce the concept of inertia while studying the pendulum which Newton finalized in his first law of motion. Many of the concepts and ideas which Newton utilized were part of the common knowledge and philosophy of his time. The mechanistic attitude toward motion, the idea that the universe was a complex machine or clockwork, set into motion by the first act of creation, and therefore, requiring no Divine intervention for its continuation, but continues to function according to some undiscovered principle of motion, had its origin with Descartes, and was challenged by Robert Boyle and others who believed that God could intervene to invalidate the laws of Nature. Thus it was when Newton remarked,"If I have seen farther, it is because I have stood upon the shoulders of giants." But it was primarily due to Newton's keen insight and genius that he was able to see the interrelationship between Kepler's laws of planetary motion and Galileo's physics of falling bodies to develop his theory of gravitation. All of this work was accomplished in the brief span of about two years while Newton was vacationing at his home at Woolsthorpe due to the closing of Trinity College to prevent the spread of the bubonic plague which was sweeping Europe at this time.

As the story goes, Newton began to wonder whether the force of gravity which caused an object such as an apple to fall toward the earth might not be the same force that caused the moon to orbit the earth and kept the planets in their orbits around the sun. In 1666, he wrote:

> "I began to think of gravity extending to the orb of the moon, and thereby, compared the force requisite to keep the moon in her orb with the force of gravity at the surface of the earth."

Using Kepler's Law of Periods, Newton showed that the necessary force must be one that varied as the reciprocal of the square of the distance between the center of the two bodies. Stated in symbols:

$$F \propto \frac{1}{r^2}, \qquad (1.10)$$

where F is the force of gravity, and r is the distance between the centers of the bodies. In order to derive this law of force, Newton first had to invent the differential calculus and develop his laws of motion, two major scientific accomplishments in themselves.

Newton did not immediately publish his findings, for he was a shy, introspective person by nature who shunned criticism. He also hesitated publishing his results because he could not obtain a good experimental verification until he could show that the mass of a body could be considered concentrated at its center. This he accomplished in 1685. Urged by his close friend Halley, who discovered Halley's Comet, Newton finally agreed to publish his work in 1687. His treatise, *Philosophiae Naturalis Principia Mathematica*, or *Mathematical Principles of Natural Philosophy*, was a masterpiece of scientific thought and methodology, and provided the basis for the science of mechanics for almost 300 years.

In his *Principia*, Newton stated his Law of Universal Gravitation:

> *Every object in the universe attracts every other object with a force that is directly*

proportional to the product of their masses, and inversely proportional to the reciprocal of the square of the distance between them.

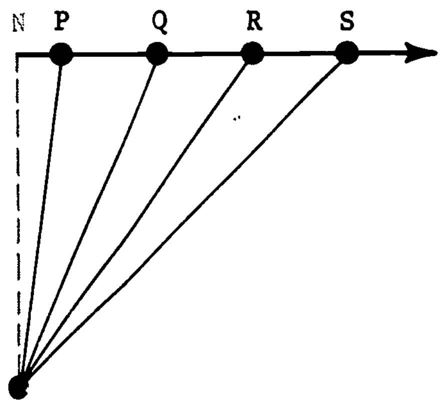
Stated in mathematical symbols:

$$F \propto \frac{m_1 m_2}{r^2} \quad ,$$

or,

$$F = G \frac{m_1 m_2}{r^2} \qquad (1.11)$$

where F is the gravitational force, m_1 is the mass of one body, m_2 is the mass of the second body, r is the distance between the center of the masses, and G is a constant of proportionality, called the Universal Gravitational Constant. Here in one mathematical statement is summarized the laws of Kepler and the cause of these laws, the mass of the bodies. Before we attempt an explanation of how this law applies to the motion of planets, we will need to look at Newton's three laws of motion themselves. Newton's first law of motion stated in the *Principia* can be summarized as:



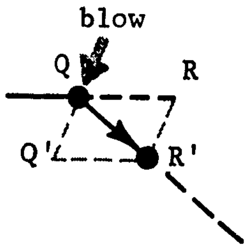
An object moving with uniform motion about point O.

> *A body continues in its state of rest or state of uniform motion in a straight line unless acted upon by an unbalanced force.*

This law is usually referred to as the Law of Inertia and had its origin with Galileo. It is the antithesis of the Aristotelian concept of motion.

Newton's second law of motion states:

> *A body will undergo an acceleration when acted upon by an unbalanced force, and its acceleration is directly proportional to the applied force in the same direction as the applied force, and inversely proportional to the mass.*



An object deflected from its path *QR* by a blow at *Q*.

Stated in terms of symbols:

$$a \propto \frac{F}{m}$$

or,

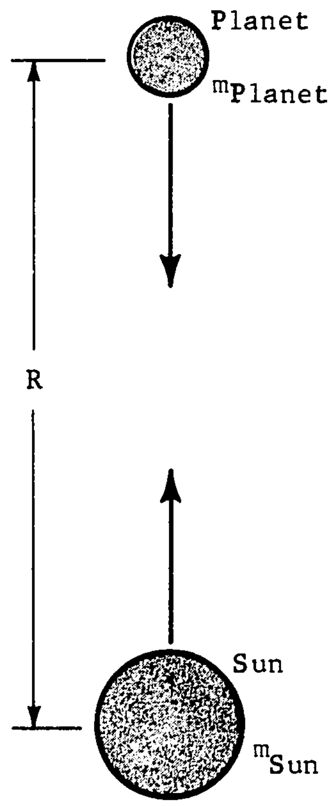
$$a = K \frac{F}{m} , \qquad (1.12)$$

where a is the acceleration of the object, F is the applied force and K is a proportionality constant. m is called the *mass* or *inertia* of the object. The inertia of a body is the resistence that it offers to motion, and is a measure of the amount of matter in the body.

Newton's third law of motion may be stated as:

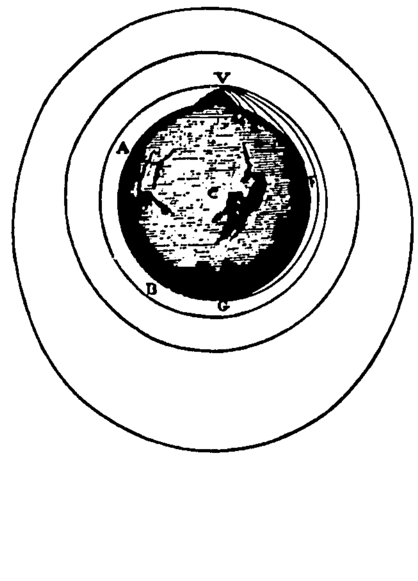
For every force (or action) there is always an equal and opposite force (or reaction).

This law means that if an object pulls or pushes on another object that the second object pushes or pulls back on the first object with an equal but oppositely directed force. A good example of this law is illustrated by a light-weight football player running into a blocking or tackling dummy. If the dummy is much more massive (has greater inertia) than the player, he may be knocked backward by the reaction force that the dummy exerts on him. Another example is a baseball striking a bat. The bat exerts a force on the ball, and in turn, the ball exerts an equal but opposite force on the bat which can be felt by the hitter.

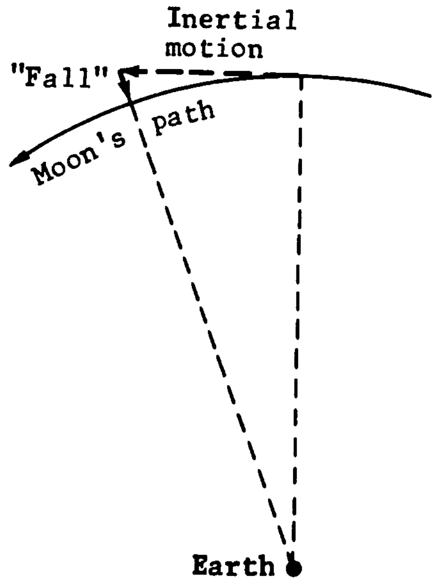


The sun exerts a force on the planet. The planet exerts an equal but opposite force on the sun.

We are now ready to explain how Newton explained the motion of the planets. Newton knew that gravity attracted a freely falling body toward the center of the earth, if it fell from rest, because of Galileo's experiments with falling bodies. He also knew that if a body is projected upward at some angle to the horizontal that it falls to earth along a curving path which Galileo showed was a parabola. The height and horizontal distance that the object travels is a function of the projection angle and its initial speed. Newton reasoned that if an object is projected from a high mountain top with increasingly greater speeds, it would travel increasingly greater distances before falling to earth. Newton next envisioned its speed being increased to the point where it traveled half-way around the earth before falling to the ground.

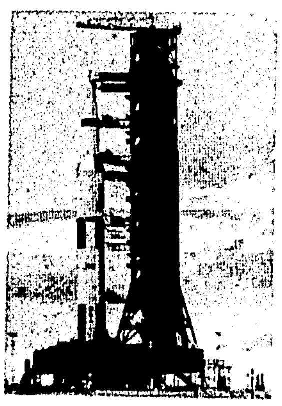


Finally, he considered the speed to be great enough that the object would travel completely around the earth and return to its starting point. He then concluded that the object would continue to circle the earth indefinitely if the effect of air resistance were neglected. Thus, Newton arrived at an explanation of satellite motion and the reason that the moon revolved around the earth. The moon is actually falling around the earth in a circular path! The reason that the moon is attracted toward the earth is the earth's gravitational attraction for the moon. If the gravitational force did not exist, then the moon would continue in a straight line in accordance with Newton's first law. Due to the continous application of the earth's gravitational force, the motion of the moon is changed, in accordance with Newton's second law, and it is accelerated in the direction of the earth. According to Newton's third law, the moon also exerts an equal and oppositely directed force on the earth, but due to the earth's greater inertia, its orbital path is not affected to a great



extent. The gravitational attraction of the moon on the earth is manifested, though, in the ocean tides, which result because the water is free to move under the moon's influence. The sun also exerts a tidal effect on the earth due to its gravitational force. Newton concluded that the planets revolve about the sun because of this attraction for the same reason that the moon revolves about the earth.

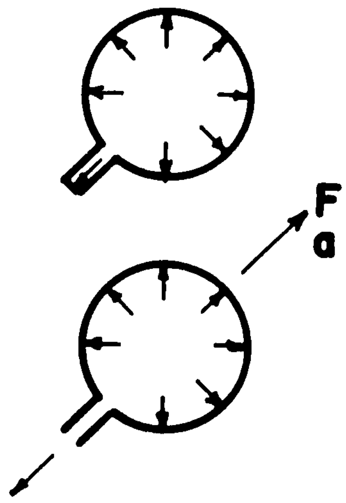
THE FLIGHT OF APPOLO 8 REVISITED Now let us go back and analyze the flight of Apollo 8 in terms of what we have learned about motion and gravity. The flight itself is an excellent example of Newton's three laws of motion. Before liftoff, the rocket, with its enormous weight, represents a great inertia at rest. According to Newton's first law, it will continue at rest until a force is applied to it. At ignition, the thrust of the engines supply the necessary force to overcome its static inertia and to start it accelerating upward against the downward pull of gravity. After the first stage burns out, it falls back to earth along a parabolic path, while the second stage engines fire to thrust the rocket into an earth orbit. Once the spacecraft has reached the necessary speed to allow it to "fall freely" around the earth, the engines are cut off.



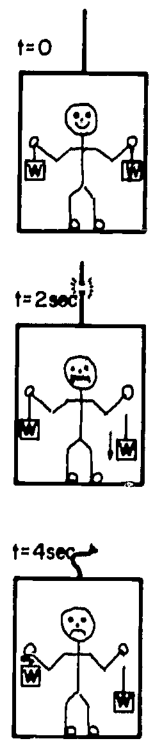
Now, since the spacecraft is above the atmosphere, there is no air resistance to slow it down, so it continues to free fall about the earth in a circular or elliptical orbit. Again, appealing to Newton's first law, we find that it will continue in orbit indefinitely unless a force is applied. In order to return to earth, the spacecraft must be slowed down by the firing of "retro" rockets. As its speed decreases below orbital speed, it begins to fall to earth like a ballistic projectile. Reentering the earth's atmosphere at these enormous speeds results in violent collision of air molecules with the heat shield of the spacecraft which heats up several thousands of degrees and partially burns away. At a certain altitude, parachutes are deployed to slow the spacecraft to a gentle landing in the ocean to be picked up.

If, instead of firing the "retro" rockets, the main propulsion rockets are fired, then the speed of the spacecraft increases relative to the earth, and if it is great enough, the spacecraft goes into a higher orbit to allow it to fall toward the moon. Because the spacecraft is still in the influence of the earth's gravity, it begins to decelerate as it travels toward the moon. At some point in its flight, the earth's gravitional field becomes equal to that of the moon, and the spacecraft enters the moon's sphere of influence to be accelerated toward its surface. If the trajectory is correct, the spacecraft passes close to the moon and swings back into space to return to earth. If a lunar orbit is desired, retro rockets are fired to slow it to the correct speed to enter lunar orbit. To return, the main propulsion engine increases its speed to "escape velocity" and it returns to earth.

The rocket engine is an excellent illustration of Newton's third law of motion. As the fuel is burned, it expands outward in the combustion chamber and some exhaust gases travel out of the exhaust nozzle. The escaping gases cause a reaction force in the opposite direction which propels the spacecraft in the opposite direction. You may have made a toy rocket as a child with a cardboard tube and a balloon. If the balloon is inflated and sealed, the pressure is the same in all directions and no movement occurs. If, however, the cardboard nozzle is opened, air escapes, creating a difference in pressure which causes an unbalanced force which propels the balloon rocket forward.



Now that we know how the rocket functions and how it places the spacecraft into orbit, let us investigate another phenomenon associated with spaceflight-that of "weightlessness." If you watched the Apollo 8 mission on television you probably saw the astronauts "floating" about the cabin of the spacecraft or releasing a pen or flashlight and observed it float suspended in air. How can we explain this phenomenon? A similar condition can be accomplished if an elevator suspended from the top of a very tall building suddenly began to fall to the earth. Suppose you are inside such an elevator. (Just suppose.) What would happen if you happened to drop a pen that you were holding? Would it fall to the floor of the elevator? If you answered no, you are correct. The reason that the pen does not fall to the floor is that it is falling at the same acceleration as the elevator (approximately 32 feet per second per second) and yourself. The pen would appear to "hover" or "float" in mid-air at the same position that you released it. If you did not know that the elevator was falling freely, you might assume that someone had "turned off" the earth's gravitational force inside the elevator, or that the pen is "weightless." The reason then that the astronauts appear to "float" about the cabin is the same reason that the pen in the freely falling elevator "floats." Everything in the observer's frame of reference is falling at the same rate. That is why it is erroneous to say that the pen is "weightless." It is not that there is no gravitational attraction acting on the pen, but rather that everything shares the same acceleration in free fall. This also



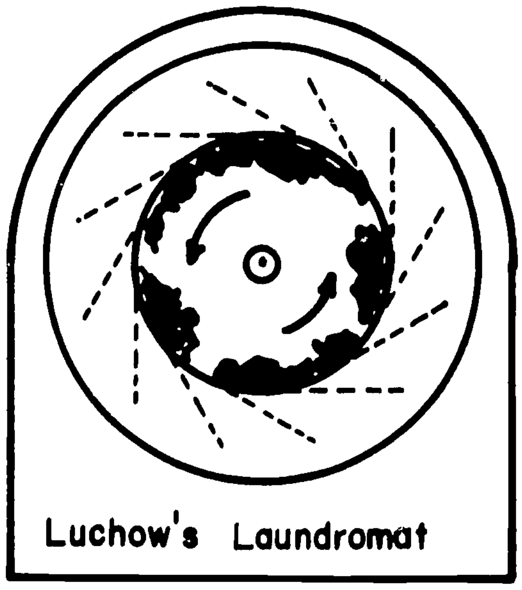
Astronaut in free fall orbit.

explains why an astronaut can step outside of his spacecraft in orbit and not be left behind. He is traveling just as fast as the spacecraft. There is no way that he could be left behind unless he slowed himself down intentionally with a small rocket propulsion unit.

This brings up the question, what is the difference between weight and mass? Mass, or inertia, we said earlier, is the resistence a body offers to motion. Even in free fall in space an object has mass, because it requires a force to change its direction of motion. One must fire retro rockets to change the direction and speed of the spacecraft, for instance. What is weight then? It is a *force* that an object possesses due to its presence in a gravitational field. Whenever an object is accelerated it acquires a force or weight. At the surface of the earth gravity pulls downward on objects with a certain acceleration, and we say that the object possesses "weight." If the gravitational acceleration changes on an object, then its weight changes. An object that weighs 6 pounds on earth would only weigh 1 pound on the moon because the moon's gravitational acceleration is only one-sixth that of earth's because of its smaller mass. But an object on the moon would still require the same force as it would on earth to set it into motion.

The general term applied to a force that keeps the moon and earth satellites moving in their orbits is *centripetal* or "center-seeking" force. In the case of the moon and satellites, the centripetal force is the force of gravity. However, not all centripetal forces are gravitational. You probably experience some kind of centripetal force each day and are not aware of it. The force that causes a car to go around a curve is one such example of a centripetal force. Here the centripetal force is supplied by the reaction force that the highway exerts upon the tires of the car - a frictional force. If this centripetal force were not supplied by friction, as in the case of an icy highway, then the car would continue on in a straight line, off the highway, in accordance with Newton's first law of motion. But wait, as you rounded the curve, didn't you press outward against the doors of the car? Where does this force which pushes you outward come from? This force is generally referred to as a *centrifugal* or "center-fleeing" force. Actually, this is not a real force at all, but merely the result of the fact that the car is being accelerated around the curve beneath you, while you tend to travel in a straight line due to your inertia. The car supplies the necessary *centripetal* force to your body to keep it moving around the curve. If you were sitting in an open convertible without any doors or seat belts, you might find yourself flying out of the car as it abruptly rounded a sharp curve. Note that the centrifugal force appears only in the *moving frame* of reference of the car. To someone viewing the scene from a helicopter overhead, the only force acting on the passengers in the car is the *centripetal* force causing them to move around the curve.

Centripetal force is actually a very useful force. We use it to spin-dry clothes and separate cream from milk. Chemists use it to separate suspended solids from liquids, and it is used to create "artificial gravity" or *"g"* forces to condition astronauts. Consider a clothes dryer for instance. As the perforated drum rotates, centripetal force causes the clothes to move in a circular path by providing an inward force from the walls pushing on the clothes. The water in the clothes is free to move, so it flies off tangentially to the circumference of the drum. Heat is also applied to aid in drying the clothes faster. Note again that an explanation depends upon your frame of reference. From the viewpoint of the water droplets inside the drum, a centrifugal force acts upon them to cause them to move outward.



EINSTEIN AND RELATIVITY The previous discussion brings up the importance of reference frames and how they affect our description of physical phenomena. Beginning with Aristotle and continuing down through Newton, we find the development of the idea of a "place" or a position for objects in the universe. Newton extended this to a concept of Absolute Space, or a frame of reference that exists of itself somewhere in the universe to which motion may be referred. This concept is implicit in Newton's formulation of mechanics. In order to describe motion one must be able to refer to some reference point to show that a change in position has taken place. Imagine how difficult it would be to give directions if the earth were a perfectly smooth, homogeneous sphere with no identifiable features such as mountains and rivers. How could you direct someone to a particular point on the sphere? You couldn't, unless you began to draw straight intersecting lines on its surface and labeled them to serve as reference points. Perhaps now we see Newton's problem. Not meaning to detract from Newton's greatness, suffice it to say, it was a slight oversight on his part to assume an Absolute Space as a reference point. If we gaze into space, everywhere that we look we see only one thing, motion. How are we going to set up a reference frame if every object is moving about? What if your house changed its position each day? Is there anything in the universe to which we could assign a permanent reference frame? Well, the "fixed" stars are such enormous distances away that their motion is not as easily discernible as is the motion of the moon or planets. However, based upon this notion of Absolute Space, Newton formulated his laws of mechanics which were used for almost 300 years (until 1905) as the "correct" means of describing motion. Today, we realize that they have only a limited range of validity. Newton, himself, probably realized that his concept of Absolute Space was only an approximation, but it served his purpose well enough then that he accepted it in spite of its drawbacks.

Around 1881, a team of American scientists, named Michelson and Morley, undertook a series of experiments to measure the speed of light relative to Newton's Absolute Space, which was then referred to as the ether, and which was believed to be a weightless, invisible substance that pervaded all space and served as a medium for the propagation of light and as a fixed frame of reference. Michelson and Morley attempted to measure the speed of light relative to this frame and from their measurements determine the relative speed of the earth through the ether. Their experiment was based upon the idea that speeds add and subtract like numbers. The situation is similar to determining the speed of a man relative to the ground, when he walks forward aboard a moving train which has some speed relative to the ground, and then measuring his speed as he walks backwards aboard the train relative to the ground.

One should obtain a difference in the two cases according to Newton's mechanics. Michelson and Morley performed a similar experiment using light instead of a man and performed this experiment at different times of the year to obtain a different direction in the earth's speed through the ether. When they performed the experiment, to their amazement, they could detect no difference! The speed of light was the same in all directions regardless of the direction of travel of the earth. This constancy of the speed of light for all observers (along with some other electromagnetic phenomena) led Einstein in 1905 to formulate his Special Theory of Relativity (or relative motion). Since the speed of light, c, is always a constant for all observers and we define speed to be $v = d/t = c$, or $d = ct$, then our concepts of distance and time must not be the same for all observers. This led Einstein to postulate that distance and time are interrelated. The consequence of this postulate is that now distance, or space, and time are no longer absolute quantities in themselves, but part of a greater unified concept, the concept of a space-time continuum. Newton also believed that there was a Universal Time for the universe and that all events appeared the same for all observers. Such could only be the case if the speed of light is infinite which it is not. Einstein says that only the speed of light appeared the same for all observers, and space and time are relative quantities for each observer. Because of this uniqueness and the difficulty of comprehension of the advanced mathematics used to derive these concepts, they have often been the object of considerable distortion by science-fiction writers who speak about the "fourth-dimension" as some dimension of space inaccessible to our senses in which weird things occur. This is not what the concept of space-time means. What it means is that one cannot adequately describe the motion of an object without giving three spatial coordinates or reference points and one coordinate of time.

For Einstein, there was no such thing as absolute motion, only relative motion. This was such a unique idea that conflicted with the commonsense notion of space and time that the world found it difficult to comprehend or accept at first. Still today, one finds some serious scientists who still do not accept it as true, even in the face of no experimental evidence contradicting it.

GENERAL RELATIVITY THEORY In 1915, Einstein further astonished the scientific community of the world by publishing his General Theory of Relativity which dealt with the description of motion in accelerated frames of reference instead of motion in uniformly moving frames as in the Special Theory. Here, in one monumental mathematical exposition, Einstein tied together the concepts of space-time, mass, and gravitation in a single unified theory. Briefly, it states that gravity is the result of the mass of a body, and that this mass affects the structure of space-time, or the geometry of the universe. If there is no matter present in the universe, then space-time would be "flat" or Euclidean in character. But because of the presence of matter, space-time is curved near massive bodies and this curvature of space-time accounts for gravitational effects. Objects travelling through space-time travel along the shortest path consistent with the geometry of space-time. The path is only straight in a Euclidean sense in a region of space-time where no other matter exists to distort its structure. Near matter the paths of objects are curved. This effect has been verified on numerous occasions for light passing near the sun. During solar eclipses, stars appear shifted when light from them passes near the solar disk.

Here, then, is a geometrical explanation for gravity in the Solar System. The earth and planets revolve about the sun, not because of its gravitational force that it exerts upon them, but because the sun distorts space-time and thereby causes the planets to follow the path that they do. Thus, Einstein can say with Kepler, "this is the reason...."

SUGGESTED OUTSIDE READINGS:

E.N. da C. Andrade, *Sir Isaac Newton - His Life and Work*, Doubleday (Anchor), Garden City, New York, 1958.

E.A. Burtt, *The Metaphysical Foundations of Modern Science*, Doubleday (Anchor), Garden City, New York, 1954.

B. Cohen, *The Birth of a New Physics*, Doubleday (Anchor), Garden City, New York, 1960.

G. Gamow, *Gravity*, Doubleday (Anchor), Garden City, New York, 1962.

A.R. Hall, *The Scientific Revolution 1500-1800*, Beacon Press, Boston, Massachusetts, 1954.

Gerald Holton, *Introduction to Concepts and Theories in Physical Science*, Addison-Wesley, Reading, Massachusetts.

Fred Hoyle, *Astronomy*, Doubleday, Garden City, New York, 1962.

R.M. Hutchins, Editor, *Great Books of the Western World*, Encyclopedia Britannica, Inc., Chicago, 1952.; Volumes 8, 16, 19, 28, and 34.

Arthur Koestler, *The Watershed: A Biography of Johannes Kepler*, Doubleday (Anchor), Garden City, New York, 1960.

Cornelius Lanczos, *Albert Einstein and the Cosmic World Order*, Interscience, New York, 1965.

Eric M. Rogers, *Physics for the Inquiring Mind*, Princeton University Press, New Jersey, 1960.

George Sarton, *A History of Science*, John Wiley, New York, 1965.

Raymond J. Seeger, *Men of Physics: Galileo Galilei, his life and his works*, Pergamon Press, London, 1966.

Stephen Winter, *The Physical Sciences*, Harper and Row, New York, 1967.

Study Questions:

1. What were some of the significant contributions of the Babylonians to the science of astronomy?

2. In what sense were the science and mathematics of the Babylonians and Egyptians of a strictly utilitarian nature?

3. Contrast the philosophical views of the Babylonians and the Greeks.

4. What Greek school of thought was responsible for advancing the notion that perfection in mathematics may carry over into the real world? What were some of its beliefs?

5. What was Plato's problem and why did it occur?

6. What was the Greek solution to Plato's problem?

7. Contrast the cosmological models of Aristotle and Aristarchus.

8. What was the chief advantage of Aristarchus' model over that of Aristotle?

9. Why was the Aristarchian model rejected by the Greeks?

10. What innovations did the astronomer Ptolemy make in the geocentric model?

11. Why did the Ptolemaic model survive for nearly 1500 years as the correct model of the universe?

12. Why was the Copernican model so vehemently attacked and rejected by the Scholastic school of thought?

13. Why was it difficult for unbiased thinkers to choose between the Ptolemaic and Copernican models?

14. What developments were necessary before a valid choice could be made between the Ptolemaic and Copernican models?

15. What was Tycho Brahe's major contribution to astronomy?

16. In what way was Kepler's work representative of Greek thought and in what way was it representative of modern scientific thought?

17. Cite some examples of Kepler's mysticism occurring in later literary works.

18. Why was it difficult for Kepler to determine the reason for his laws of planetary motion?

19. Why is Galileo considered to be the first modern scientist?

20. Discuss the characteristics of the medieval concept of the world.

21. Briefly describe the Aristotelian concept of motion.

22. What arguments did Galileo use to disprove the validity of the Aristotelian concept that a heavy body falls faster than a lighter body?

23. Why was Galileo unable to verify his definition of uniformly accelerated motion by direct measurements? How did he overcome this problem?

24. If you dropped a rock into a deep well and *saw* it strike the water surface 2 seconds later, what would be the depth of the well? (Hint: the rock accelerates at the rate of 32 feet per second each second) What would be the limiting factor in the accuracy of your determination of the depth of the well?

25. Suppose you dropped a rock into a dark, deep well and you *heard* the splash 4 seconds after you released the rock. What would be the depth of the well? (Hint: sound travels at the speed of approximately 1000 feet pre second in air).

26. What was the significance of Galileo's experiments with the inclined plane?

27. How was Galileo able to disprove the Aristotelian concept of "violent" motion and introduce the concept of inertia?

28. Why were Galileo's astronomical observations discounted by his opponents?

29. Why was the study of falling bodies so important? Discuss some innovations that Galileo brought to this study; how he went about it, and what major contributions he made in this area.

30. What did Newton mean when he remarked, "If I have seen farther, it is because I have stood upon the shoulders of giants."?

31. In what way is the work of Newton a synthesis of work of Kepler and Galileo?

32. State and discuss significant features of Newton's laws of motion. What do you understand the concept of force to mean?

33. Discuss some of Newton's contributions to society. Contrast the social, political, and religious climate in which Galileo worked to that of Newton.

34. Using the laws of Universal Gravitation and Newton's second law of motion, show that bodies fall with the same acceleration regardless of their mass. What assumptions have you made?

35. If an astronaut wishes to rotate his spacecraft through an angle of 90 degrees in orbit, he must fire two sets of rockets. Why?

36. If someone remarked that an astronaut was weightless because there is no gravity "up there", how would you answer him?

37. Alexander Pope wrote:

Nature and Nature's Laws lay hid
in night
God said "Let Newton be" and all
was light

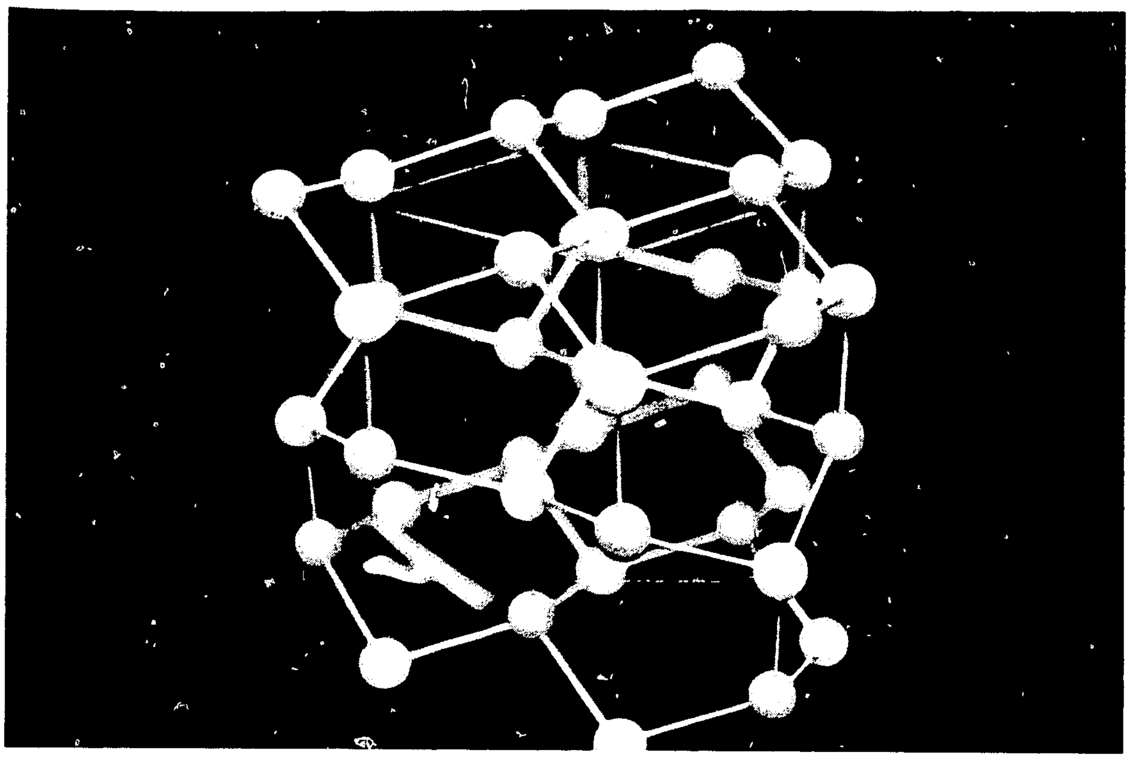
It did not last, the
Devil howling "Ho
Let Einstein be!" restored
the status quo.

(J.C. Squire)

Interpret the above quotes.

CHAPTER 2 CONCEPTS OF MATTER

PAGE



INTRODUCTION Our knowledge of matter has come a long way from the time of Aristotle. Aristotle taught his students that all matter was made from the four elements - fire, earth, air, and water. With the proper combination of these four basic elements, any and all types of matter could be created. You might be familiar with the story of the alchemists who attempted to create gold by obtaining the proper combination of the basic elements - fire, earth, air, and water.

Even though the alchemists did not create gold, they contributed much to man's fund of knowledge of matter. They were the forerunners of the science of chemistry, and it was through their endeavors that a concept of some of the basic elements that make up matter began to become accepted. A partial periodic table of the elements was formulated in 1869 that predicted the elements which had not yet been discovered, but it was not until the 1940's that the entire periodic table which included all of the elements found in nature, up through uranium, was compiled. Let us look at some of the contributions that led to this.

In 1896 Henri Becquerel of France discovered radioactivity. In 1911 Rutherford of England gave us our present theory of the atom having a small nucleus. To account for some of the mass in atoms the neutron had been theorized previously, but it was not until 1932 that it was confirmed by experiment, so you can see that our present theories of the atoms and molecules are relatively new. One generation ago it was taught by many teachers that the atom could not be subdivided. Since then we have found the atom can be split into many components. We have learned how to synthesize new elements not found in nature. We have learned that energy can be converted into matter and have performed this feat at an installation in California. Our concept of matter has had to be revised as we have seen that matter can be regarded as a form of energy when we utilize the equivalence principle as stated by Einstein in his famous equation, $E = mc^2$.

The atomistic theory of matter can be traced back over two thousand years. Credit for the origination of the atomic theory is given to Leucippus of ancient Greece. By setting a limit to the concept of divisibility of distance, he was led to the concept of the atom, the physically "indivisible." He was followed by Democritus of Abdera (about 460-370 B.C.) who wrote that the world consisted of empty space and an infinite number of indivisible, invisibly small atoms.

The thoughts of both of these men undoubtedly played a role in the conclusions arrived at by Aristotle who was born in 384 B.C. The Aristotelian concept of the four basic elements was founded upon the idea of indivisible particles. Aristotle defined an element as "one of those bodies into which itself is not capable of being divided into others." Aristotle did have to introduce a fifth, immaterial element, which he called *quintessence* to explain why there could be so many combinations of fire, earth, air, and water. This fifth element corresponds to the ether of the nineteenth century physicists.

Even though the atomistic theory of matter goes back very far in recorded history, it did not play any major role in scientific progress until after the discovery of the chemical law of multiple proportions by John Dalton. Roger Bacon of England had attempted to make alchemy a respectable science in his writings of around 1267, but he spent the last 15 years of his life (1277-1292) in prison for criticizing some of the famous Dominicans like Albert Magnus and Thomas Aquinas.

The thoughts of three more men concerning the atomistic nature of matter should be mentioned before introducing Dalton's concept since these three men did have an influence on scientific thought. The corpuscular theory of the Frenchman, Descartes, supposed that the properties of substances depended mainly on the shapes of their particles. He envisioned acids as having sharp spiky particles, which prick the tongue, and their salts formed sharp crystals. In precipitation reactions, the spikes of the acid particles break off in the pores of metal corpuscules, e.g., of silver, and are carried down in the precipitate. He explained the reaction of metals with acids by the points of the acid tearing apart the particles in the mass of the metal.

Robert Boyle (1627-1691) has been called the founder of modern chemistry for introducing a rigorous experimental method into chemistry and for proving, by experiment, that the four elements of Aristotle did not deserve to be called elements at all, since none of them could be extracted from bodies, e.g., metals. Boyle believed in the atomic thoery and gave a fairly good definition of an element, but he seems to have regarded different elements as being made up of some primary matter. In other words, he still believed in the possibility of alchemy.

Newton's concept of atoms becomes important since all of Newton's teachings later came to be held in such high esteem. For a time in history, Newton's concepts were given the same unquestioned respect that the Aristotelian concepts had been given previously. Newton viewed matter as being composed of absolutely hard, indestructible particles, and though all changes could be regarded as separations, associations, and motions of these permanent atoms. Although Newton depended more on mathematical interpretations of observations, he visualized atoms as being smaller elements of sensibly experienced objects. To quote from his *Principia:*

> "We no other way know the extension of bodies than by our senses, nor do these reach it in all bodies; but because we perceive extension in all bodies that are sensible, therefore we ascribe it universally to all others also. That abundance of bodies are hard, we learn by experience; and because the hardness of the whole arises from the hardness of the parts, we therefore justly infer the hardness of the undivided particles not only of the bodies we feel but of all others. That all bodies are impenetrable, we gather not from reason, but from sensation. The bodies which we handle we find impenetrable, and thence conclude impenetrability to be a universal property of all bodies whatsoever. That all bodies are movable, and endowed with certain powers (which we call the *vires inertiae*) of persevering in their motion, or in their rest, we only infer from the like properties observed in the bodies which we have seen. The extension, hardness, impenetrability, mobility, and *vis inertiae* of the whole, result from the extension, hardness, impenetrability, mobility, and *vires inertiae* of the parts, and thence we conclude the least particles of all bodies to be also all extended, and hard, and impenetrable, and movable, and endowed with their proper *vires inertiae*."

Newton could not accept Descartes' idea that by endless encounters some of the particles might have broken off and become "old worn particles" different from the rest. Of particular importance in the Newtonian concept of atoms was his belief in the infinite hardness of these particles.

John Dalton (1766-1844) was not such a brilliant person as Newton and many other outstanding scientific contributors; he modestly attributed his success to "perseverance." His great independence from contemporary scientists and from contemporary thought seems to have played a role in his contributions although the influence of Newton is apparent in Dalton's work. By plodding through problems which might have seemed trivial to some, the resolution of these problems led him to the generalization of his atomic theory which is in close agreement with our present concepts. Lothar Meyer said that Dalton's Atomic Theory is so simple that "at first sight it is not illuminating." It asserts that:

1. The chemical elements are composed of very minute indivisible particles of matter, called atoms, which preserve their individuality in all chemical changes.

2. All atoms of the same element are identical in all respects, particularly in weight. Different elements have atoms differing in weight. Each element is charaterized by the weight of its atom.

3. Chemical combination occurs by the union of the atoms of the elements in simple numerical ratios, e.g. 1 atom A + 1 atom B; 1 atom A + 2 atoms B; 2 atoms A + 1 atom B; 2 atoms A + 3 atoms B, etc.

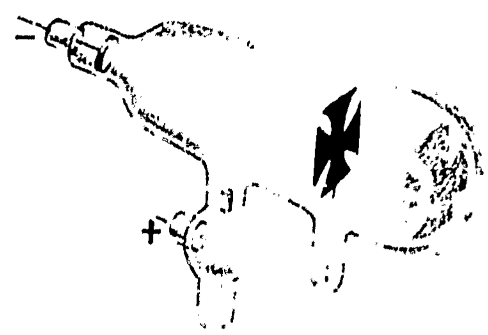
It should be realized that Dalton's Atomic Theory could not determine the relative weights of atoms from the combining proportion unless the number of atoms in the particle of the compound is known. It was this problem which caused Dalton's theory to remain in dispute for many years. Although Dalton assumed the correct ratio of elements for some compounds, he mistakenly assumed a molecule of water was composed of one atom of hydrogen and one atom of oxygen.

Dalton's theory has been modified as we have gained greater knowledge about elements, and it should also be realized that the present atomic and molecular theories are still theories and, in all probability, will be modified as we continue to increase our knowledge, but a better understanding of macroscopic behavior can be obtained if we obtain an understanding of the atomic and molecular theories. These theories are the result of hundreds of years of observation of macroscopic behavior of matter by men who devoted their lives to unlocking the secrets behind their observations. Each generation has been able to advance in knowledge because they have been able to utilize the results of the previous generations. With the breakthrough brought about by Dalton's atomic theory in the early part of the 19th century, our discoveries have been pyramiding at a fantastic rate.

By the last quarter of the nineteenth century chemists had accumulated overwhelming evidence in favor of Dalton's theory on the basis of chemical reactions and it had become firmly established. By 1875, Maxwell had formulated his theory of electromagnetism, thereby bringing together the fields of optics, electricity, and magnetism. Work by Kelvin, Clausius, and others promised to unite the fields of heat, kinetic theory of gases, and atomic theory into the science of thermodynamics. At last, scientists had a complete, comprehensive view of nature explainable by Newtonian mechanics, atomic theory, and electromagnetism. Many scientists felt that science held little challenge outside of performing more exact measurements to give better agreement between the theory and the experimental values.

Now, scientists, notably physicists, began to turn their attention to unlocking the secrets of the atom itself. Some of the research took the form of investigating the nature of electrical discharges in rarefied gases. Gases were sealed at low pressures in glass tubes with metal electrodes sealed in each end. An electrical discharge was initiated in the tube by passing a current from an induction coil or electrostatic machine through the tube. In 1869 Wilhelm Hittof showed that something passed through the tube capable of producing a shadow at the anode or positive end if a metal object was inserted in the tube. He called this phenomenon *cathode rays* as it appeared to travel from the cathode to the anode within the tube.

In 1877, William Crookes succeeded in producing discharge tubes of extremely low pressure and was able to show that the cathode rays were negatively charged particles capable of being deflected by a magnetic field. Many scientists became interested in this phenomenon and began investigating it with Crookes tubes. One of these was a German physicist named Wilhelm Roentgen. In 1895, as he experimented with his tube, he observed a faint greenish glow coming from a nearby table. Investigating further, he discovered the source of the glow to be a screen coated with a phosphorescent chemical. Within weeks, he was able to determine that the Crookes tube was emitting some unknown type of radiation, which he termed "x-rays" due to their unknown nature. Experimenting with his "x-rays," he soon discovered that they possessed the property to expose photographic film like light, and could do this even if the film was placed behind an obstacle. Realizing the importance of his discovery to medical science, he published his findings in the journal of the Wurzburg Physical Medical Society. Within days, news of his grand discovery spread, and physicians immediately began to utilize it to set fractures, locate foreign objects, and observe organs. Roentgen's fame spread far and wide, and so did many misconceptions regarding his "x-rays." Newspapers carried ads for "x-ray proof underclothing" while people object-



A Crookes tube. Electrons leave the cathode (-) and travel toward the anode(+). Those electrons passing the anode cast a shadow at the end of the tube.

ed that x-rays would invade their privacy. This poem by L. K. Russel appeared in a copy of *Life* magazine in 1896:

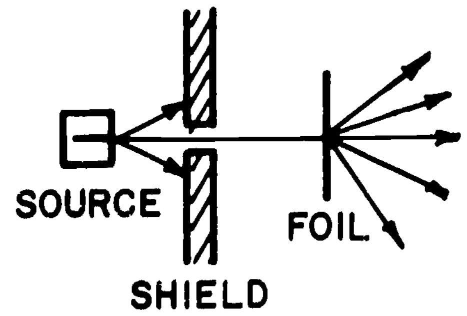
> She is so tall, so slender; and her bones -
> Are well produced by cathode rays sublime:
> By oscillations, amperes and by ohms,
> Her dorsal vertebrae are not concealed
> By epidermis, but are well revealed.

Such was some of the initial reaction, but Roentgen's discovery was to have still greater impact on the future development of physics than it did on medicine. It was probably one of those ironies of history that several people could have been the discoverer of x-rays other than Roentgen, for many experimenters prior to his publication had noticed that their photographic plates fogged up when stored near their equipment. Crookes, himself, returned film to the manufacturer, thinking it faulty. As Joseph Henry, the co-discoverer of Faraday's law of induction, remarked, "The seeds of great discoveries are constantly floating around us, but they only take root in minds well prepared to receive them." Often, great advances are made in science by what is termed the "happy accident."

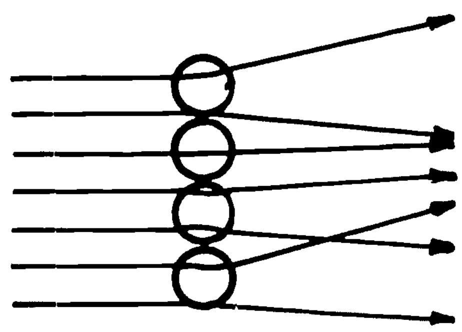
Discoveries came fast and furious following Roentgen's announcement. In 1896 Becquerel discovered the phenomenon of radioactivity accidently in a manner similar to Roentgen's discovery. Two years later, Pierre and Marie Curie isolated polonium and radium. In 1897, Sir J. J. Thomson discovered that cathode rays were actually streams of electrons and advanced a model of the atom consisting of electrons embedded in a sphere of positive charge, which became known as the "pudding model." There were numerous other contributors to the modern theory of the atom during this period, but we will only mention four others because they marked the turning point in the development of modern physics; these were Planck, Einstein, Rutherford, and Bohr.

Max Planck made his contribution to the development of modern physics in another field - in the study of blackbody radiation in thermodynamics. Prior to Planck's discovery, it was believed that radiation consisted of electromagnetic waves of various wavelengths, which were emitted continuously by atoms. Several people advanced theories to explain the type of spectra obtained by radiations being emitted by an ideal radiator - a blackbody. None of these theories agreed with the experimental data accurately enough. Planck discovered that he could obtain agreement between the experimental data and theory if he assumed that atoms emit radiation in a discontinuous manner. This was the beginning of the modern concept of quantization in physics. Planck, himself, balked at the idea that the radiation itself could be discontinuous in nature, and spent considerable time and effort trying to verify continuous absorption of continuous waves.

In 1905, Einstein proved energy was both emitted and absorbed in a discontinuous manner in units of energy that he called *quanta*. This discovery was a consequence of the photoelectric effect in metals. Whenever ultraviolet radiation was incident upon certain metals, they emitted electrons. Einstein was able to show these metals could only absorb energy in a discontinuous manner. Eight years later, Niels Bohr made use of Einstein's quanta in developing his atomic theory.

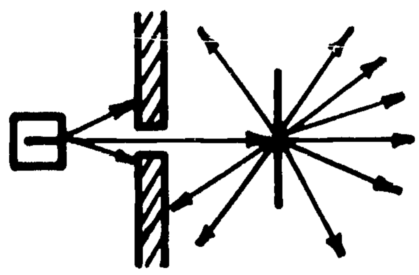


In 1908, Ernest Rutherford was awarded the Nobel Prize for Chemistry for his work concerning the chemistry of radioactive materials. Shortly afterwards, he made a more important discovery in physics. While studying the scattering of alpha particles by thin metal foils, he observed that some were deflected through very large angles. Prior to this the Thomson model of the atom was in vogue and on the basis of this model, one would expect very little scattering at large angles. This result perplexed Rutherford and on a later occasion he wrote, "It was almost as incredible as if you fired a 15-inch shell at a piece of tissue paper and it came back and hit you." Rutherford realized that the only way this result could have occurred was for most of the positive charge in the atom to be concentrated in a very small volume or nucleus. In 1911, he published a paper presenting his *nuclear* model of the atom, consisting of a small, positively charged nucleus surrounded by a sphere of negative charge.



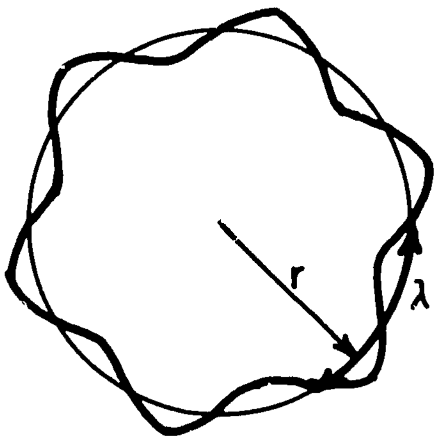
Thomson atoms scattering α - particles.

In 1925, Niels Bohr joined Rutherford's research group and began work on devising an atomic model based upon Rutherford's nuclear model. One of the perplexing problems during this period was to develop an explanation for discrete atomic spectra observed in gaseous discharges. Within a year, Bohr had the answer to this problem. His solution was that atoms were similar to a planetary system, consisting of a heavy, positive charged nucleus at the center, and circling the nucleus in diccrete orbits were negatively charged electrons. Bohr formulated his theory by bringing together the Rutherford nuclear model and Planck's and Einstein's ideas on energy quantization. Prior to this other scientists had proposed planetary models, but these models were un-



Nuclear atoms scattering α - particles.

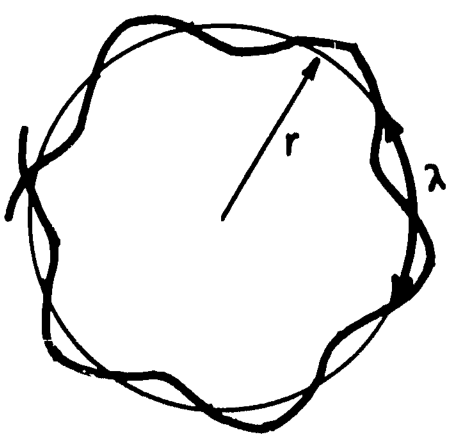
stable because classical electromagnetic theory postulated that in such a model electrons would radiate energy and spiral into the nucleus. Bohr overcame this difficulty by assuming that only discrete orbits were possible in which the electrons *did not* radiate. Bohr's theory met immediate success in explaining the discrete atomic spectrum of hydrogen; however, for more complicated atoms it deviated considerably from the experimental data. Numerous attempts were made in following years to revise Bohr's theory to give better agreement with observation but with little success.



Electron wave in orbit.
Constructive interference.

In 1925 Louis de Broglie published a paper in which he predicted that matter itself possessed wave properties in a manner analogous to light possessing both wave and particle (quantum) properties. In a short period of time afterwards, Schrödinger and Heisenberg developed quantized wave mechanical models of the atom. Although both men used different approaches, their theories were equivalent. These quantum mechanical theories were mathematical in nature and required a new abstract interpretation of the atom. Thus, we find that the two modern theories governing the behavior of the infinitesimally small and infinitely large are mathematical models that still leave unanswered the nature of the concrete real world if, indeed, such a thing exists.

In the remaining sections of this chapter we will look at some aspects of atomic and molecular phenomena in light of what we have discussed.



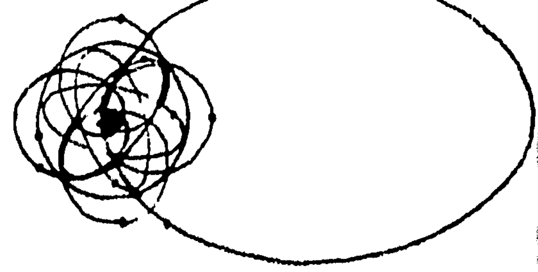
Electron wave in orbit.
Destructive interference.

ATOMIC THEORY Usually, the hydrogen atom is studied first when we study the atomic theory, but we need to realize the hydrogen atom is unique. All atoms of elements, with the exception of hydrogen, are made up of three components - positively charged protons, neutrally charged neutrons, and negatively charged electrons - but the vast majority of all hydrogen atoms are composed of only one proton and one electron. Most of an atom can be thought of as empty space, with the mass of the atom concentrated in what we call its nucleus. The nucleus of an atom consists of protons and neutrons and is located in the center of the atom occupying an extremely small amount of space occupied by the atom. The electrons orbit the nucleus in a complicated three dimensional manner that is still not completely understood.

Many different analogies have been given in an attempt to help visualize the atom but no analogy is completely successful. The Bohr analogy is to compare the nucleus to the sun with electrons compared to the orbiting planets. Some benefit can be obtained from this analogy, since we can begin to visualize that most of the atoms can be thought of as empty space. Also, electrons do spin while they are orbiting although not in the same manner as planets. The quantum mechanical model frequently given is to think of the nucleus as a small particle surrounded by a cloud of electrons smeared out in shells around the nucleus. These are only hazy concepts which we are attempting to get across and you must modify your visualizations of the atom as more information is given and discovered.

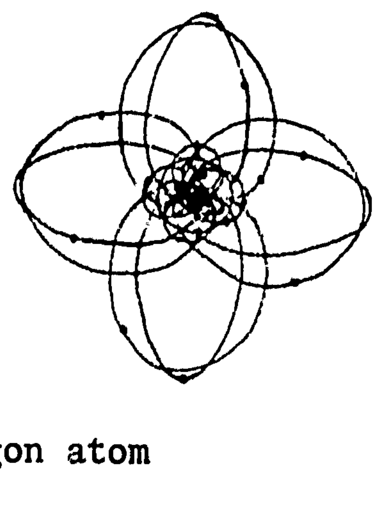
We have discovered that the electrons occupy more space than the protons and neutrons, but the mass of the electron is negligible when compared to the mass of a proton or a neutron. A proton or a neutron is approximately 1836 times more massive than an electron, although the electron occupies approximately 2.5 times the space as a proton or a neutron. The space occupied by the electron might be due in part to the wave characteristics we have come to associate with all matter.

There are presently 103 or 104 known elements. All elements above uranium (92) have been artificially made and are not found in nature, and there is reason to assume that more elements will be synthesized. We have numbered these elements in order from 1 through 103. We can associate these consecutive numbers, which we call the *atomic number* of the element, with the number of protons in the nucleus of the atom. The element hydrogen has an atomic number of 1; it has only one proton in or as its nucleus. Helium has an atomic number of 2; it has 2 protons in its nucleus, and so on.



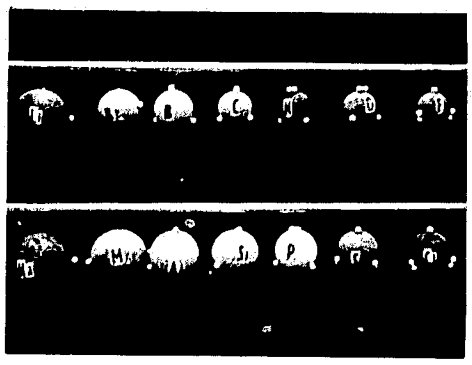
Sodium atom

The mass of an atom is essentially dependent upon how many protons and neutrons are contained in the nucleus. The number of neutrons in the nucleus can vary without affecting the chemical characteristics of the atom which are dependent upon how many protons are contained in the nucleus. We call these atoms of the same elements having different masses caused by different numbers of neutrons *isotopes*. For example, there are two isotopes of hydrogen *found in nature*. The vast majority of atoms of hydrogen have one proton and no neutrons, but there is a certain percentage of hydrogen atoms which have one proton and one neutron called deuterium. There is a third isotope of hydrogen called tritium which has been synthesized, having one proton and two neutrons.



Argon atom

Most helium atoms have two protons and two neutrons, although isotopes of helium can be found or can be made which have two protons and one neutron or two protons and three neutrons. Since the mass of a proton and the mass of a neutron are almost identical, you should be able to see that the percentage of isotopes found in samples of specific materials could cause a difference in weight of the samples even if there were the same number of atoms of the element in each sample.



Atomic models showing relative size.

The percentage of occurence of isotopes of elements is relatively stable in nature, and we do not notice this difference in weight unless we artificially separate the isotopes. Besides changing the mass of the atoms, neutrons also seem to serve a major purpose of somehow holding the nucleus together. If we get the right combination of protons and neutrons, the atom can become stable. The nucleus will not change unless it is bombarded with sufficient energy to cause it to split or to fuse with the bombarded forms of energy. But when there is an improper combination of protons and neutrons in the nucleus, the nucleus becomes *unstable*. It tends to give off forms of energy until it has a more stable arrangement of protons and neutrons. This emission of forms of energy we call *radioactivity*. There are four basic types of emission due to radioactivity, although we should not think radioactivity is restricted to these four types of emission. An unstable nucleus will frequently emit either α (alpha) particles, β (beta) particles, γ (gamma) rays, neutrons or a combination of these forms of energy. *Since the nature of these emissions was not originally known, the emitted forms of energy were named after the first letters of the Greek alphabet - α, β, and γ.*

Further investigation has revealed the nature of these particles. The α particles are identical to the nuclei of helium atoms and, once emitted, attract two electrons into orbit and become a stable helium atom. The β particles have the same physical properties as electrons, and there seems to be no reason not to think of them as electrons. The γ rays were found to be different from either of these in that they did not possess an electrical charge; they can be thought of as resembling x-rays although some gamma rays have more energy than x-rays. The emission of an α particle causes the nucleus of an atom to lose two protons and the atomic number of the element which is created drops two places on the periodic table. Since the α particles are composed of two protons and two neutrons, the mass of the newly created element would be four units less than the original atom. When a β particle is emitted, a proton is created where a neutron was before, and the atomic number of the atom increases by one unit due to one more proton in its nucleus. Such an atom would change its chemical characteristics and would become an atom of the next succeeding element on the periodic table. The emission of γ rays does not affect the number of protons and neutrons in the nucleus and, consequently, does not affect the atomic number or the atomic mass (to any noticeable extent). Frequently, γ radiation occurs concurrently with other forms of emission. The emission of neutrons does not change the chemical characteristics of the atom, so the emitting atom remains the same element although the mass would decrease by one unit. The nucleus *usually* is left more stable after any form of radioactive emission.

Group→ Period ↓	I	II											III	IV	V	VI	VII	0
1	1.0080 H 1																	4.0026 He 2
2	6.939 Li 3	9.012 Be 4											10.811 B 5	12.011 C 6	14.007 N 7	15.999 O 8	18.998 F 9	20.183 Ne 10
3	22.990 Na 11	24.31 Mg 12											26.98 Al 13	28.09 Si 14	30.97 P 15	32.06 S 16	35.45 Cl 17	39.95 Ar 18
4	39.10 K 19	40.08 Ca 20	44.96 Sc 21	47.90 Ti 22	50.94 V 23	52.00 Cr 24	54.94 Mn' 25	55.85 Fe 26	58.93 Co 27	58.71 Ni 28	63.54 Cu 29	65.37 Zn 30	69.72 Ga 31	72.59 Ge 32	74.92 As 33	78.96 Se 34	79.91 Br 35	83.80 Kr 36
5	85.47 Rb 37	87.62 Sr 38	88.91 Y 39	91.22 Zr 40	92.91 Nb 41	95.94 Mo 42	(99) Tc 43	101.07 Ru 44	102.91 Rh 45	106.4 Pd 46	107.87 Ag 47	112.40 Cd 48	114.82 In 49	118.69 Sn 50	121.75 Sb 51	127.60 Te 52	126.9 I 53	131.30 Xe 54
6	132.91 Cs 55	137.34 Ba 56	* 57–71	178.49 Hf 72	180.95 Ta 73	183.85 W 74	186.2 Re 75	190.2 Os 76	192.2 Ir 77	195.09 Pt 78	196.97 Au 79	200.59 Hg 80	204.37 Tl 81	207.19 Pb 82	208.98 Bi 83	210 Po 84	(210) At 85	222 Rn 86
7	(223) Fr 87	226.05 Ra 88	† 89															

*Rare-earth metals	138.91 La 57	140.12 Ce 58	140.91 Pr 59	144.27 Nd 60	(147) Pm 61	150.35 Sm 62	151.96 Eu 63	157.25 Gd 64	158.92 Tb 65	162.50 Dy 66	164.93 Ho 67	167.26 Er 68	168.93 Tm 69	173.04 Yb 70	174.97 Lu 71
† Actinide metals	227 Ac 89	232.04 Th 90	231 Pa 91	238.03 U 92	(237) Np 93	(242) Pu 94	(243) Am 95	(245) Cm 96	(249) Bk 97	(249) Cf 98	(253) E 99	(255) Fm 100	(256) Mv 101	(253) No 102	(257) Lw 103

The atomic mass that is listed on the Periodic Table above each element represents the average relative mass of all isotopes of the element as the element occurs in nature. These figures can be considered as relative to each other. Originally, it was thought that all atoms of oxygen were exactly alike; i.e., they had 8 protons and 8 neutrons and there were no other isotopes of oxygen found in nature. Oxygen was arbitrarily assigned an atomic mass of 16 exactly, and the atomic masses of the other elements were calculated in relation to how massive samples of equal numbers of atoms of these elements were when compared to oxygen. It has since been found that oxygen does have an isotope found in nature, so a new Periodic Table is now accepted which is based upon the relative number 12 being assigned to the most common isotope of carbon which has 6 protons and 6 neutrons in its nucleus. All of the atomic masses of the elements have been recalculated to compare with this particular isotope of carbon.

We frequently speak of radioactivity as the *decay* of an element or atom. As emission of forms of energy from nuclei occurs, the atoms decay into lower states of energy. Even though, in the case of β particle emission, an atom of the next higher element on the Periodic Table is formed, we can think of this as decay since the nucleus is giving up energy and becoming more stable. The decay of energy from the nuclei of atoms is a random occurrence. We could never be able to predict when random emission from a given nucleus would occur, but due to the large number of atoms found in any given sample we have found there is a definite probability that emissions will occur at a certain rate. Once emission occurs, the nucleus might become stable and never emit any more forms of energy, or it might emit a different type of energy at a different rate until eventually the nucleus becomes stable.

One way of classifying radioactive materials is to refer to what we have termed *half-life*. Due to the complete randomness of nuclear emission, some nuclei will emit energy immediately and some nuclei will not emit the excess energy which makes them unstable until years in the future. The half-life of radioactive materials is the time it takes for the rate of emission to become one-half of what it originally was. This is not predicted theoretically but has been determined by experimentation. The half-life of radioactive materials varies from fraction of seconds to tens of thousands of years.

There are three major means by which atoms of one element might be changed into atoms of other elements. One method is by radioactive decay which has just been discussed. Another means is by nuclear fission and the third major method is by nuclear fusion. Fission means to split apart and fusion means to join or fuse together. In nuclear fission the nucleus absorbs some forms of energy and becomes so unstable that it immediately splits apart. The nuclei of the newly formed atoms might be stable, although in many cases the newly formed nuclei are themselves unstable and will emit energy by decay until they become more stable. In nuclear fission the nucleus does not necessarily split into equal parts, and the elements created are not always the same elements. We have found by experimentation that we can predict the percentage of new elements created from the fission of a given element quite accurately. The bombarding of energy which causes nuclear fission might be in many forms. The neutron, because of its electrical neutrality, is not deflected when it gets close to the electrically charged nucleus

2

and therefore can penetrate the nucleus without too much difficulty.

The chance of collision between the neutron and the nucleus of an atom becomes small in ordinary situations due to the vast empty space that is part of the atom. You might compare it to the chances of a blindfolded man in a baseball stadium shooting at a baseball located on the pitcher's mound. Given enough chances, hits will occur but the number of hits will be extremely small when compared to the total number of shots.

To produce such a device as the fission bomb used at the end of World War II required establishing conditions such that neutrons could penetrate more nuclei. If we can establish the right conditions, we can set off what we call a chain reaction. Visualize a situation where the nucleus which absorbs a neutron splits into two smaller nuclei, and in the process also emits two separate neutrons. These neutrons can in turn penetrate other nuclei causing them to split apart and also emit two neutrons each. In small amounts of radioactive material the relatively few neutrons emitted, when compared to the vast amount of empty space that atoms occupy, eliminates the possibility of producing an immediate chain reaction. But as we increase the amount of a radioactive substance emitting neutrons, we increase the number of neutrons penetrating the substance and, consequently, increase the chance of nuclear reaction. We call the amount of a radioactive substance when the point is reached that an immediate chain reaction is produced its *critical mass*.

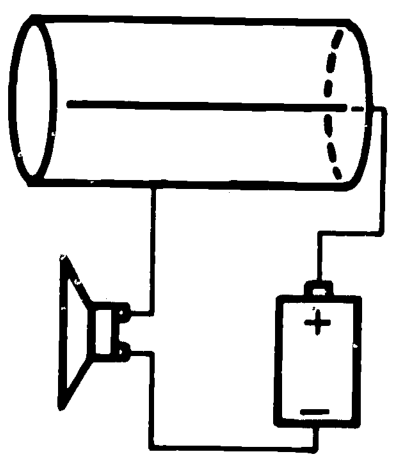
The amount of radioactive material in an atomic bomb must exceed its critical mass, but various portions of the amount would be separated from each other. When it was desired for the bomb to explode, conventional explosives would be used to drive the various portions of the radioactive material together. The nuclei would be pressed very close to each other, thereby increasing the chance of reaction between the neutrons and the nuclei and causing an immediate chain reaction releasing enormous quantities of energy in the form of heat and radiation.

Einstein was the first person to theorize that matter is a form of energy. He gave us his famous formula $E = mc^2$ which briefly summarized means that when there is a conversion of matter into energy we can calculate the energy released by multiplying the loss in mass by the speed of light squared. The energy released in such a manner is enormous. There is very little loss of mass in a nuclear explosion. Most of the mass can be accounted for in the smaller nuclei which are formed from splitting a larger nucleus, but the very little mass that is converted into pure energy is responsible for the awesome reaction we observe from a nuclear explosion.

Man has learned to control this energy to some extent in nuclear reactors. By creating a controlled chain reaction we are able to utilize this energy released to power submarines, ships, and for the generation of electricity.

A nuclear fusion reaction releases even more energy than a nuclear fission reaction because there is a greater amount of mass converted into energy. Through experimentation we have found when small nuclei are forced together they can fuse into larger nuclei, but there is still a loss of mass in the process. The methods we use include bombarding elements with forms of energy in a nuclear reactor to make their nuclei unstable and to form particular isotopes. The triggering device in a nuclear fusion bomb might utilize a fission bomb in order to create particular isotopes and to drive the particles close enough together so they will fuse.

We theorize that the sun is releasing energy due to the nuclear reactions occurring there. We believe that the sun is composed mostly of hydrogen and through a fusion process of hydrogen nuclei, helium is being created. The fusion process is not limited to the creation of helium from hydrogen but could possibly explain how all of the various elements found on the Earth originated in the untold billions of years which have preceded us in the creation of the universe. We theorize stars evolve or change from a primitive or young star composed mostly of hydrogen to much denser stars as atoms containing more protons and neutrons in their nuclei are created.



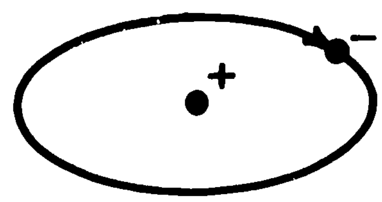
A simplified diagram of a Geiger counter. Radiation entering the tube causes ionization which is detected by an amplified sound.

In order for radioactive emissions to be detected, they must react with some material or other form of energy. There are two general ways these emissions can be detected. First, the radioactive emission may form an *ion* in the detecting material by knocking one or more electrons out of the atoms of the detecting material and the ions may be detected by either electrical or chemical methods. Second, the particle may raise the amount of energy in the electrons of the atoms without actually removing them from the atoms. When the electrons return to their previous states, they have to lose this energy, and they do this by emitting radiation such as light which can then be observed. This second way of detecting radioactive emission is called *scintillation* and requires special equipment since the amount of light or radiation emitted is so very small.

Because of the different natures of α particles, β particles, and γ rays, they may travel different distances through materials. α particles, because of their relatively large mass and electrical charge, are the easiest to stop. γ rays, because they have neither a charge nor a rest mass, are the most difficult to stop in a material. β particles generally fall in between these two other kinds of radioactive emissions. Neutrons, since they have no electrical charge, and do not react with the electrons, are not detectable unless they suddenly decay into other forms of energy or are captured by atomic nuclei which, in turn, release some other form of more detectable emission.

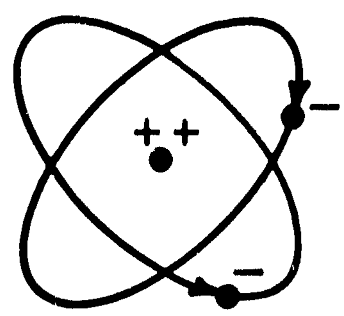
MOLECULAR THEORY All chemically stable matter is in molecular form. To understand molecules we must begin to understand the importance the electrons play in their role of orbiting the nuclei of atoms. As we learned in the previous section, atoms are composed of protons, neutrons, and electrons with the exception of hydrogen. The protons and electrons are electrically charged and the neutron is electrically neutral. It has been found that there are two types of electrical charge and we have arbitrarily named one type of charge *negative* and the other type of charge *positive*. It was subsequently found that the charge which was named positive is associated with the proton and the charge which was named negative is associated with the electron.

Like electrical charges repel and unlike electrical charges attract each other. It is the attractive force caused by unlike charges attracting each other which keeps the electrons in orbit around the positively charged nucleus. (The binding force of the nucleus must be very great to overcome the repulsive force between the protons caused by the protons having the same type of electrical charge.) To be electrically neutral each atom or group of atoms must have the same number of electrons in orbit as it has protons in its nucleus or nuclei, but this alone does not necessarily make the atom or group of atoms chemically stable. The electrons must orbit in certain conditions of paired symmetry for the atom or group of atoms to be chemically stable.



Hydrogen atom.

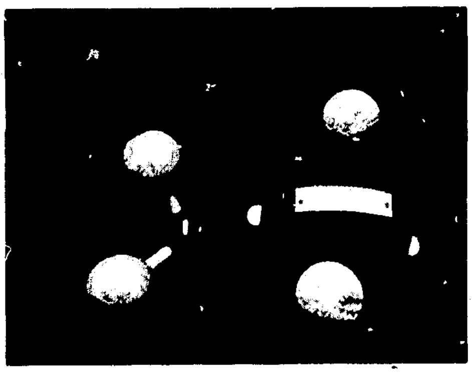
It is convenient to separate hydrogen and helium from all other elements when we discuss molecular formation. Since hydrogen only has one proton we can see it would only require one electron in orbit to make it electrically neutral, but we should see it would lack the condition of paired symmetrical orbits. To become chemically stable it must have another electron orbiting it, but if it is to remain electrically neutral it must associate with another atom or group of atoms in such a way that the positive charge of the nuclei will balance with the negative charge of the orbiting electrons. In the case of hydrogen and helium, one pair of electrons in orbit creates sufficient



Helium atom.

orbital symmetry. One hydrogen atom can combine with another hydrogen atom in such a manner that two electrons are orbiting both nuclei in symmetrical orbits while also producing electrical neutrality for the group. Two atoms of hydrogen combined in such a manner would be said to be a molecule of hydrogen.

Both conditions are met with one atom of helium, though, it has two electrons to balance the two protons in the nucleus and these two electrons can orbit in symmetry. An atom of helium could also be called a molecule of helium. Helium is so well satisfied with its condition that it will not react and it is not found combined with any other element in our soils or rock formations. We can separate the electrons from the helium nuclei in electrical tubes and cause the helium to emit green light as the electrons jump back into orbit but, when we get through, the gas in the tube will still be helium in molecular form.



Models of water molecules.

Hydrogen, on the other hand, is very reactive. It burns with the hottest flame we can produce by chemical reactions. It combines with many different elements and is frequently found combined with oxygen and carbon. Chemical bonds are sometimes compared to marriages and most hydrogen marriages can be thought of as double ring ceremonies. Usually, the hydrogen atom shares its electron orbit with the electron orbit of one of the electrons of the atom it is associating with. In a very common reaction two hydrogen atoms will associate with one oxygen atom and a very chemically stable molecule will be formed - water vapor. This type of chemical bonding, typical of hydrogen, where electrons are shared to make paired symmetry for both atoms involved, is called *covalent bonding*.

Another convenient method we use to explain chemical bonding is to speak of electron *shells*. The electrons whiz around the nucleus so fast that a more graphic analogy might be to think of them as forming an electron *cloud*, but we have found by experimentation that the electrons seek specific energy levels in their orbiting which we have named their *shells*. The closest energy level to the nucleus has been named the *K* shell, the next energy level, *L* shell, and so on in alphabetical order. The closest energy level, or K shell, is satisfied with one pair of electrons orbiting in symmetry. The L shell can have up to 4 pairs of electrons or a total of eight electrons in orbit. The next shell, the M shell, can have up to 4 pairs of electrons *if it is the outer shell*, but as the number of energy levels or shells increase, we find the *M* shell can hold up to 9 pairs of electrons when it is one of the inner energy levels of electrons.

From observation, we have concluded that the *outer shell*, after we get past the *K* shell, is limited to 4 pairs of electrons or 8 electrons. In chemical bonding necessary to form molecules, the electrons in the outer shell, in most cases, are responsible for the joining or close association of one atom with another. Occasionally, unpaired electrons in the energy level below the outer shell also play a part in chemical bonding. It should be realized that since the electrons are electrically charged particles, and since the electrons are responsible for chemical bonding or association, that chemical bonding or association is electrical in nature. In fact, many of our molecules are held together by electrical attraction caused by the atoms having opposite electrical charges. Such chemical association of atoms is called *ionic bonding*.

A rule of thumb we might utilize, not always true, is that when there are less than four electrons in the outer energy level or shell there is a tendency for atoms to give these electrons up in order to establish a paired symmetry of electronic configuration. When there are more than four electrons in the outer shell there is a tendency for atoms to borrow electrons or gain electrons so there will be symmetry of four pairs of electrons in orbit in the outer shell. When there are four electrons in the outer shell, or close to this number, there is a tendency for atoms to share electron orbits to make the four pairs of symmetrical orbits.

Lithium, for example, has a total of three electrons in orbit - one pair in its *K* shell and a single electron in its *L* shell. This condition makes lithium very reactive, and it is always found combined with another atom in nature. It tends to give up its single electron in order to establish a paired symmetry of electron orbits and is held closely to the atom it associates with by electrical attraction. Fluorine, on the other hand, has a total of nine electrons - one pair of electrons in its *K* shell and 7 electrons in its outer shell. Six of these outer electrons can be considered paired, but we need four pairs to cause the electron orbits to be balanced. Fluorine will borrow or gain one electron to satisfy this balance. Lithium and fluorine are frequently found together as a mineral. The mineral is called *lithium fluoride*. The individual atoms are held together by electrical attraction and the combination is electrically neutral. This is an example of ionic bonding.

It might be beneficial to think of most chemical bonds as being ionic to some extent because, even when there is covalent bonding between two atoms, it has been found that the area surrounding the free sides of the atoms involved have varying degrees of opposite electrical charge. We conclude that the electrons involved are attracted to a greater degree by one of the atoms or that due to the complexity of electron orbits that the electrons involved spend more time in orbit around one of the atoms. It is due to the slight imbalance of electrical charge around the molecule that molecules associate with other molecules in certain ways to form such things as liquids and crystalline solids.

Atoms have a tendency to give up electrons or gain electrons until the electronic symmetry in their outer level is satisfied. When such an event takes place, leaving the atom or group of atoms associated with the atom which does such a thing electrically charged, we say an ion has been formed. Ions can be charged positively or negatively. If an atom or group of atoms gives up an electron it will give up a negative charge. There will be one more proton or positive charge in the nucleus or nuclei of the atom or group of atoms than orbiting electrons and the resulting ion would be charged positively. When an electron is gained in order to establish electronic symmetry, the atom or group of atoms will gain one negative charge and the ion formed will have a resulting negative charge. For all practical purposes we can consider that ionic formation must precede chemical reactions or the recombination of atoms.

The elements whose atoms have a tendency to give up electrons, thus forming positive ions, have been defined as *metals*, and those elements whose atoms have a tendency to gain electrons forming negative ions have been defined as *non-metals*. There are certain elements where difficulty arises in thinking of them as metals or non-metals. Carbon and silicon are good examples of elements which do not really fit into either category. They can be thought of behaving as metals or non-metals depending upon which elements they are reacting with. Hydrogen is another element which is difficult to categorize. Most frequently we think of hydrogen as behaving like a metal by giving up an electron and becoming positively charged. Other times hydrogen is thought of as behaving like a non-metal. The most practical way to resolve the dilemma is to consider the other element involved in the combination. The other element will usually be more easily defined as a metal or non-metal.

It should be seen that atoms which have a tendency to give up electrons will combine with those atoms which tend to gain electrons and form molecules of electrical neutrality. In naming compounds formed by the combination or two or more elements it is standard practice to name the metal first and the non-metal last. When elements such as fluorine and lithium are combined we *do not* name the compound *fluorine lithide*. As mentioned earlier in the chapter it would be called lithium fluoride, and the formula for the compound would be written LiF and not FLi.

There are certain groups of atoms called *radicals* which are tightly bound together but which still form ions caused by one or more of the atoms involved giving up or gaining electrons. Most radicals form negatively charged ions and are considered as acting as non-metals. The only common radical that acts as a metal by giving up an electron and becoming positively charges is the ammonium radical made up of one atom of nitrogen and four atoms of hydrogen - NH_4^+ . The "+" signs denotes one unbalanced positive charge associated with the group of atoms. This radical could combine with some other atom or group of atoms which tend to gain electrons and form a molecule such as ammonium fluoride - NH_4F - or ammonium sulfate - $(NH_4)_2SO_4$. The sulfate radical forms an ion,, SO_4^{--} , having two unbalanced negative charges associated with it caused by gaining of two electrons in orbit. Since the ammonium radical only gives up one electron, we will have to have two ammonium radicals associated with one sulfate radical to obtain electrical neutrality.

From the observations of our senses of the macroscopic world of matter we are familiar with, it is difficult to conceive of these molecules we are attempting to visualize bouncing around and colliding with each other to the extent they do. The molecules which make up the page of this book are agitating quite vigorously. We do not notice this agitation when we place our finger on the page because the molecules in our skin are agitated to a similar relative degree and there will be an equal transfer of energy from your finger to the page as there is from the page to your finger. When you hold a cold glass in your hand it feels cold because your mind interprets there is a loss of energy from the molecules in your hand to the molecules in the glass.

There is a very easy method to demonstrate this if you have not done it already. You will need three containers of water; glasses or beakers will be fine. In one of the glasses run hot water (not scalding). In the second glass put in ice water, and in the third glass have room temperature water. Place the index finger of one hand into the glass of ice water and the index finger of the other hand into the glass of hot water and hold them there for a minute or two until the molecules in your skin adjust to about the same energy level as the molecules in the water. Then insert both fingers side by side into the container of room temperature water. Your mind will interpret that the water feels warm to one finger and cold to the other, but it is very clear to us that the temperature of the water is constant.

The degree of agitation of molecules is a relative thing and can be thought of as denoting the average energy of the molecules. Heavier molecules having the same average energy or temperature would move slower than lighter molecules. A golf ball would have to be moving much faster than a baseball to transfer the same amount of force upon impact.

Ordinarily, we do not notice the effect of individual collisions of molecules due to the extreme smallness of size of the molecules. It is the average impact of all molecules involved that is observed by our senses, but we can study the movement of small particles being agitated by the molecules of the air or gas as they are struck randomly by the gas molecules.

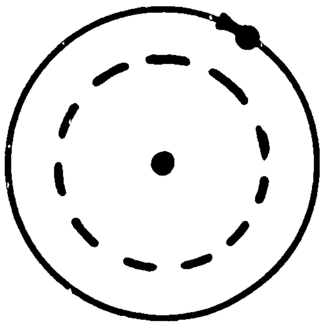
As these molecules collide with one another they can cause ionization to occur. We might say there are two conditions which must be met for these collisions to cause ionization. First, the colliding molecule must have enough energy to cause the atoms to break apart from each other, and, secondly, the colliding molecule must hit another molecule in the proper location, or the area where connected. To gain complete understanding of these collisions we must not think of something hard hitting something hard but of electron clouds from one atom coming close to electron clouds of other atoms. Since like charges repel, it is the electrical repulsive force which causes atoms to be deflected from their paths.

In liquids we have ideal conditions for chemical reactions to take place. The molecules are in close proximity to each other colliding continuously. Molecules in many solutions are continually forming ions and regrouping into molecules due to these collisions. The molecules are able to move rather freely and are not bound in a lattice-like structure which we have come to think of as related to crystalline solids. As these liquid molecules separate into ions they immediately seek electrical neutrality (they are attracted to oppositely charged ions) and regroup into either molecules of the same compound or molecules of a different compound depending upon which ions are close to each other. If the original molecules placed in liquid form are more difficult to separate by collision, the compounds are said not to react with each other. There will still be some reaction, usually, as ions are formed and regroup into molecules, but you should be able to see that equilibrium will be reached where the percentage of each type of molecule in the liquid should remain relatively stable.

If the ions regroup into molecules of a different combination which are held together more firmly making them more difficult to separate by collision, we say there has been a chemical reaction. Again, an equilibrium point is reached where the percentage of particular molecules found in the liquid remains relatively stable. The completeness of a chemical reaction depends to a great extent upon this equilibrium point. It will vary to a great degree for different compounds, although frequently, we have molecules so loosely held together that they recombine into molecules which are so tightly held together that we can virtually say that the chemical reaction went to completion or that as many of the new combinations were formed that could be formed. Liquids have the added advantage in chemical reactions of allowing free movement of molecules and ions throughout the liquid in order to shorten the time required for the equilibrium point to be reached.

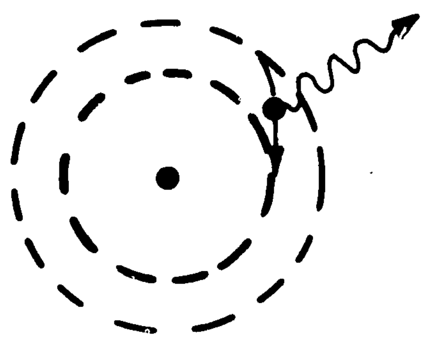
There is another major cause of ionization which must be mentioned, although it might be difficult for you to visualize since we have no satisfactory models to use for comparative purposes. Radiation from the sun is absorbed by the electrons orbiting the molecules in our atmosphere and on the surface of the earth. When this energy is absorbed by the electrons, the electrons gain potential energy and the radii of the electron orbits are increased. If sufficient energy is absorbed, an electron can escape from its orbit and an ion is formed. Whether the electrons which escape orbits are binding electrons is unimportant in some reactions since the binding electrons will frequently shift down to replace an inner electron given sufficient energy to escape orbit. The result is an ion which must either gain an electron or undergo a chemical change to become electrically neutral.

Plants are able to utilize the ionization caused by the absorption of the sun's radiation to build tremendously complicated molecules from carbon dioxide (CO_2), water (H_2O), and minerals containing such elements as nitrogen, calcium, sodium, phosphorus, potassium, and iron. Since the electrons have gained in potential energy by absorbing the radiation from the sun, animals can utilize the energy released as molecules are decomposed in their bodies to maintain a body temperature. As electrons shift down into closer orbits they give off energy in the form of radiation. Electrons are selective in the absorption and emission of radiation dependent upon the energy level at which the electrons are orbiting for absorption, and how far they shift down into closer orbits for emission.



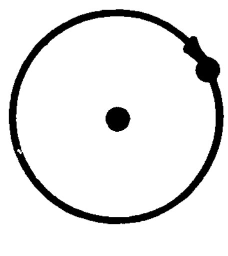
An electron in an excited energy state.

It is possible to calculate the energy level of binding electrons in a specific location of a molecule, and by beaming radiation of a certain energy level through a material, to cause these bonding electrons to escape orbit. By having the proper chemicals in a solution, it is possible to cause certain other ions to join at these specific locations. Through such a slow painstaking process, our scientists have synthesized tremendously complicated molecules and, in the process, have gained insight into the arrangement of atoms in many of our larger molecules such as DNA.



Electron transition emits radiation.

Possibly, one of the most difficult concepts of the physical sciences that we must comprehend is the concept of these invisible atoms which we cannot see being held together in combination by electrical forces. The molecules that make up this page and the molecules that make up the ink forming the print on this page are complex structures composed of many atoms joined together by electrical forces. What is this electrical force we speak of? It is not a substance. We cannot isolate it and carry it from place to place. The only way we can study it is to observe how substances are affected by it, and even though we now know much about it by observing the effects it causes, there are still many questions that remain unanswered.



Electron in lower energy state.

It is quite different from gravitational force. The only gravitational force we have discovered is an attractive force, but we have discovered there are two types of electrical charge. We have discovered that like charges repel and unlike charges attract, and it is the complexity of these forces which explain why atoms joined together in molecular form are kept apart while still being held together.

We have also discovered that molecules can be found in three different arrangements which determine whether the material could be considered a solid, a liquid, or a gas. There are two factors we must consider if we are to understand why molecules can change their form of appearance on a macroscopic scale. First, we need to visualize these molecules moving to some extent whether we think of solids, liquids, or gases. Secondly, we need to realize these molecules have varying degrees of attraction to other molecules caused by the imbalanced electrostatic charge developed around molecules. This imbalance of electrostatic charge around molecules can be due to several causes. The chemical bonding of the molecules involved might be ionic in nature; thus causing a slight electrostatic charge to be developed around the atoms making up the molecule.

Even where there is a covalent bonding of atoms within a molecule, there can be a slight electrostatic charge developed around the various atoms within the molecules due to elliptical orbits of the electrons, particularly the bonding electrons. In their elliptical orbits, the electrons will spend more time in certain portions of their orbits. This will cause varying degrees of electrostatic attraction and repulsion to other molecules at specified locations.

If the movement of the molecules is not sufficient to overcome the cohesive attraction of molecule to molecule, then the molecules will be found as either solid, or liquid. If the movement of the molecules is small enough to allow the cohesive attraction of molecule to molecule to bind the molecules into a lattice-like structure, then the material would be called a solid. In fact, the definition of a solid is dependent upon the formation of molecules into lattice like structures which we call crystals. Even though the molecules in a crystal of solid are bound into a lattice-like structure, we may still think of these molecules as agitating or moving within a certain space and of a wide range of movement permissable.

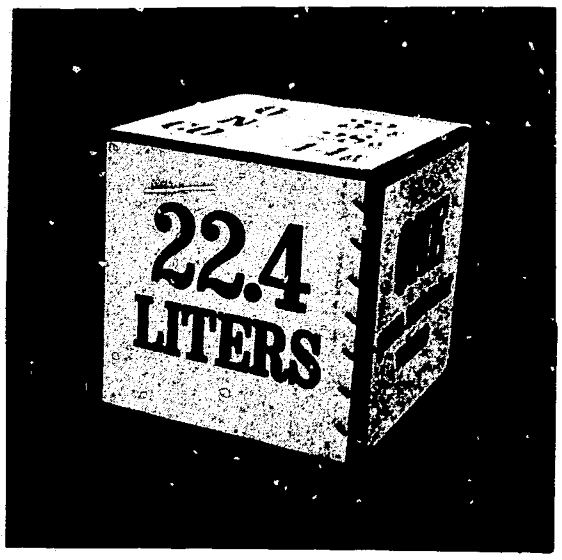
Crystal model.

If the cohesive attraction of the molecules is sufficient to hold them together but the movement of the molecules is sufficient to keep them from binding into crystals, then the material will be in a liquid state. Some materials, such as glass, which we normally think of as solids must really be categorized as very viscous liquids with this differentation between solids and liquids. The molecules in glass are not crystalline structure and thus cannot be solids by our definition of solids. To distinguish these materials, such as glass and sulfur, we call them amorphous solids to distinguish them from crystalline solids.

When the movement of molecules becomes sufficiently great to overcome the electrostatic attraction of molecule to molecule then the molecules can escape as individual molecules. When this occurs, a gas is formed. A gas is nothing more than a quantity of molecules moving around freely and independently of each other.

In all of the three states of matter, there will be molecules moving at different velocities. Some molecules will be moving slower than the average velocity of movement while others are moving faster. As these molecules collide with each other, a transfer of energy can occur. Sometimes this transfer of energy is sufficient to change their state.

For, example, in the process of evaporation some molecules are given sufficient energy through collisions with other molecules to escape from the surface of the water and go off as individual molecules of water vapor.



At normal atmospheric pressure, one gram molecular volume of a gas is 22.4 liters.

In materials such as iodine, the molecules go directly from a solid state to a gaseous state, thus by-passing the liquid state. Such materials going from a solid state to a gaseous state are said to *sublime* and the process is called sublimation. It is interesting to note that gaseous iodine and gaseous carbon dioxide can be liquified by submitting the gases to sufficient pressure. This introduces another factor affecting the three states of matter which must be considered. As you might realize, water boils at different temperatures dependent upon the altitude. At high elevations such as on the top of tall mountains the atmospheric pressure is so low that mountain climbers cannot make hot coffee since the water boils, or changes to a gaseous state, when it is still only luke warm.

THE PAST, PRESENT, AND FUTURE These, then, are some of the things that we have learned about the structure of matter since the time of Democritus. What use we make of this knowledge has, and will continue, to influence our lives and those of generations yet unborn. In the last quarter century man has learned to unleash the awesome energy stored within the nucleus for both good and evil. What is done with this energy in the future will depend upon your generation. The discoveries of science are innocuous in themselves. It is the ends to which men apply them that determine whether or not they will be for the benefit or detriment of mankind. Nuclear energy may be used to destroy life and civilization, or it may be used to lift civilization to heights undreamed of at its discovery. The sea could be converted into fresh water by the use of nuclear-powered desalinization plants which could produce electricity as a by-product and reclaim valuable minerals from the sea. The water could provide irrigation of deserts to produce an abundant food supply for the world's growing population. Today, much of the world is underdeveloped because it lacks the natural energy resources necessary for an industrial and technological society. Cheap nuclear power could supply adequate energy to develop these countries, much like TVA helped to develop the central southeastern United States in the 1930's. Today, due to peaceful nuclear research we have developed methods of curing cancer by irradiation of cancerous cells, provided the physician with tools for diagnosis of thyroid deficiency diseases, blood volume determinations, and the like.


A cobalt - 60 unit for treatment of cancer.

Industry utilizes radiation to preserve foods by destroying harmful bacteria by irradiation. It creates new and useful plastics by radiation induced polymerization. It detects flaws in materials by use of neutron and gamma radiation fluoroscopy to provide better quality control and products. Nuclear energy has probably played some part in many products and services that you use everyday and are not aware of it. To condemn science for giving us the secrets of the atom is comparable to condemning General Motors or Ford for the accidents on our highways. It is not the source that is responsible. It is the ends to which we apply our knowledge.

SUGGESTED OUTSIDE READINGS:

E.N. da C. Andrade, *Rutherford and the Nature of the Atom*, Doubleday (Anchor), Garden City, New York, 1964.

A.R. Beich, *The Story of X-Rays from Röntgen to Isotopes*, Dover Publications, New York, 1960.

Louis de Broglie, *The Revolution in Physics*, Noonday Press, New York, 1953.

G. Gamow, *Thirty Years That Shook Physics: The Story of Quantum Theory*, Doubleday (Anchor), Garden City, New York, 1966.

R.T. Lagemann, *Physical Science: Origins and Principles*, Little, Brown and Company, Boston, 1963.

Project Physics, *An Introduction to Physics: Models of the Atom*, Holt, Rinehart and Winston, New York, 1968.

Project Physics, *An Introduction to Physics: The Nucleus*, Holt, Rinehart and Winston, New York, 1968.

A. Romer, *The Restless Atom*, Doubleday (Anchor), Garden City, New York, 1960.

G. Thomson, *J.J. Thomson: Discoverer of the Electron*, Doubleday (Anchor), Garden City, New York, 1966.

ARTICLES FROM *Scientific American:*

T.K. Fowler and R.F. Port, "Progress Toward Fusion Power" (December 1966).

R.B. Leachman, "Nuclear Fission" (August 1965).

G.T. Seaborg and A.R. Futsch, "The Synthetic Elements: III", (April 1963).

G.M. Woodwell, "The Ecological Effects of Radiation" (June 1963).

STUDY QUESTIONS

1. How does radioactivity differ from the emission of radio waves?

2. How are isotopes of carbon similar and how do they differ?

3. What is the difference in atomic structure between an atom of nitrogen and an atom of oxygen?

4. Give some type of evidence to show that an atom has a nucleus.

5. Describe a method by which radioactivity has been used to measure age.

6. Discuss Bohr's theory of the atom. What were its successes and failures?

7. What conditions must be met for single atoms to be molecules?

8. What are radicals in chemistry?

9. Explain the difference between ionic and covalent bonding.

10. Explain two ways ionization can occur.

11. What is the difference between metals and non-metals?

12. What two conditions must be met for matter to be in molecular form?

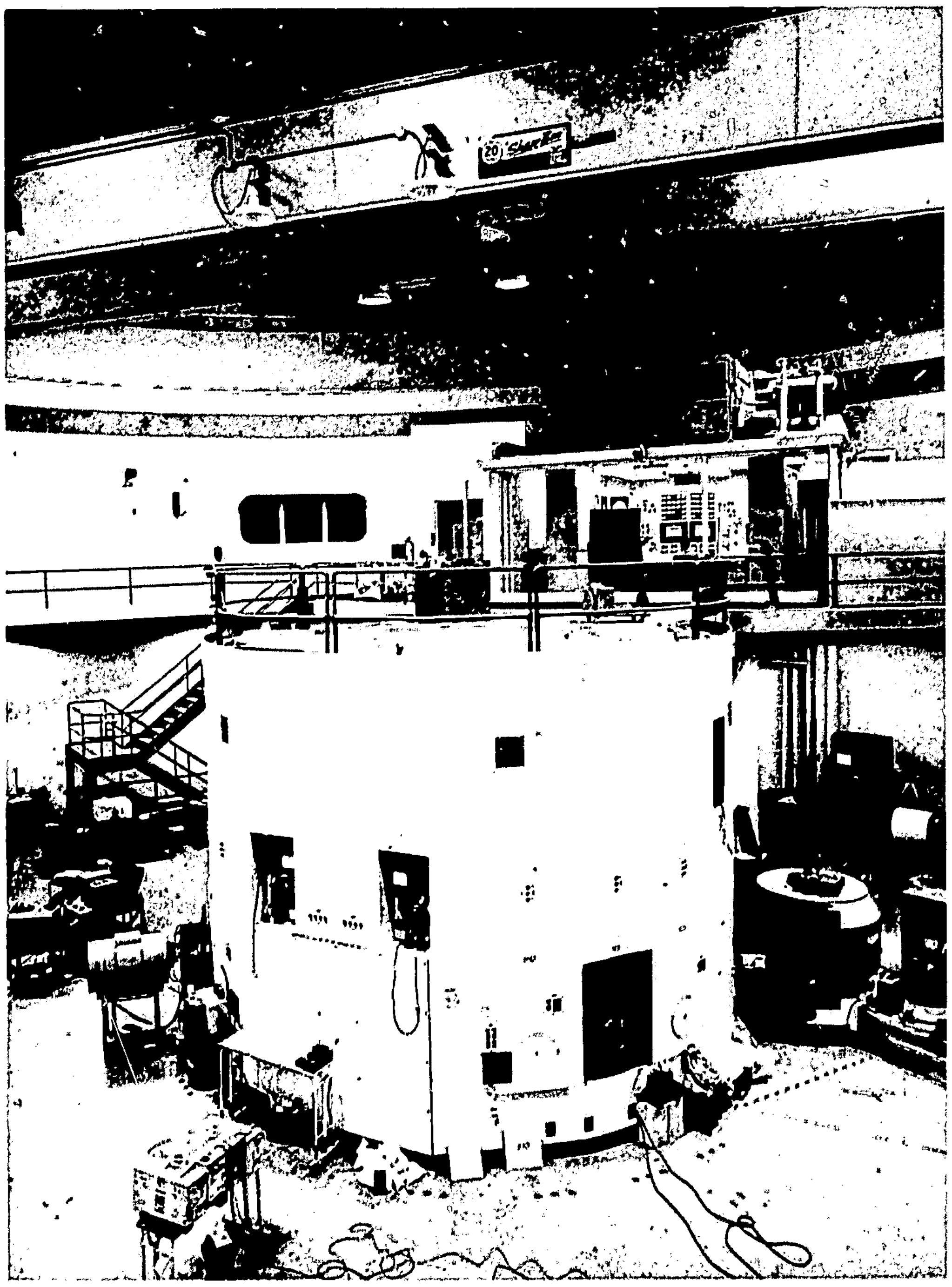
13. List three major premises of Dalton's atomic theory.

14. Give some evidence to support Dalton's theory.

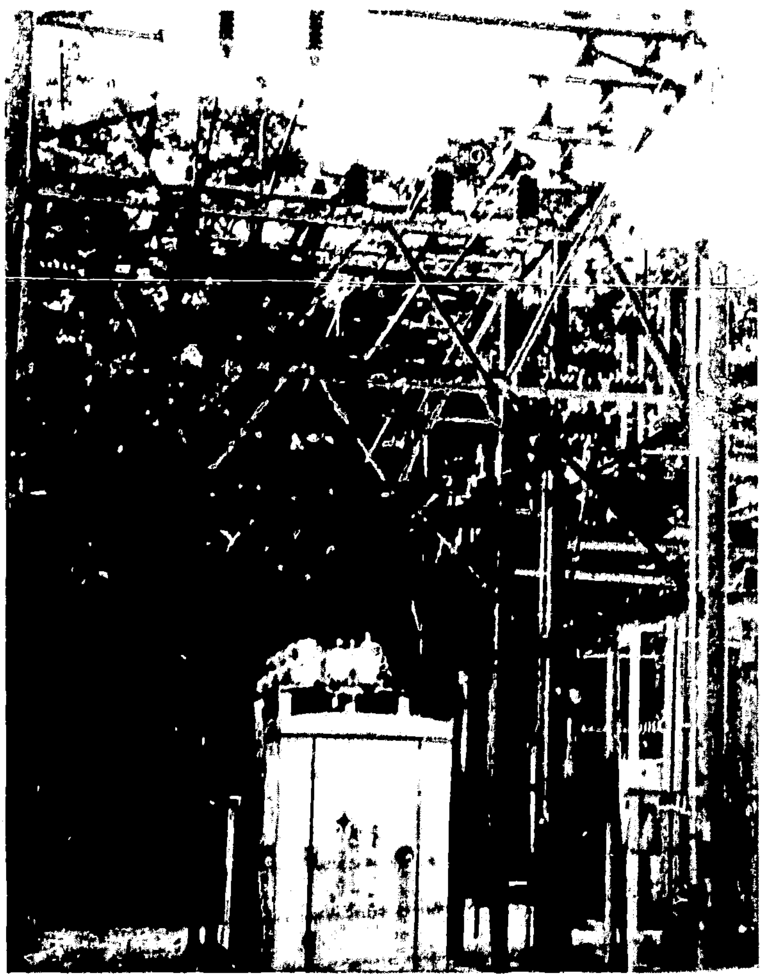
15. How do solids differ from liquids (from a molecular viewpoint)?

16. What is meant by sublimation?

The Georgia Tech Research Reactor



CHAPTER 3 CONCEPTS OF ENERGY



THE GREAT BLACKOUT It was a hot, humid summer day in 1965 while the population of the nation's largest city was going about its business that it happened. It did not occur everywhere simultaneously, but one by one, sections of the great city slowly began to die. Up and down Wall Street ticker tapes, carrying the latest financial news, sputtered to abrupt stops. Elevators, crammed with workers, suddenly paused between floors, casting their human cargo into inky darkness. Beneath the city, subway cars crowded with homeward-bound travelers rolled to a gradual stop, to freeze upon their rails like giant prehistoric creatures. Radio and television programming ceased until emergency power could be restored. Above the city, airliners began to stack up awaiting clearance to land. In thousands of offices throughout the city, typewriters, dictating machines, computers, and air conditioning ceased to function. People began to wonder. Is it a nuclear attack? An invasion from outer space? No, it was neither of these things, but its effect was almost as devastating.

Somewhere on the great power grid interlinking the cities and communities of the great metropolitan Northeast an overload occurred of such magnitude that it blacked out New York City and its neighbors. Before it was over, it had affected several states and millions of people. To these people, modern civilization as they knew it ceased. Will it happen again? Perhaps.

One generally does not stop to realize the extent to which he is dependent upon electrical energy until something like the Great Blackout occurs. We have become so accustomed to this modern genie providing us with heat, light, cooling, transportation, entertainment, and a multitude of other conveniences that we tend to take it for granted until a power failure occurs.

Today, our civilization is so dependent upon electrical energy that it is almost inconceivable that we could ever do without it. Yet we have only had this obedient servant to cater to our needs and pleasures for less than a century. Our great-grandparents did not have this servant and they survived. Still, they did not require its services to the extent we do today, because their time was less complicated and not as industrialized. In this section we will look at energy in its many forms and see how we have developed this dependence, how much we will depend upon it in the future, and some of the problems and possible solutions that an increasing demand for more energy will pose for our life in the future.

WHAT MAKES THE WORLD GO 'ROUND? From the dawn of human history to the present, man has depended upon the sun as a source of energy in some form. When he burned wood for fire, he was utilizing solar energy stored up in plants through the process of photosynthesis. As man progressed and learned the use of fire, he became less dependent upon the sun as a direct source of energy. With the development of agricultural societies, man acquired a surplus of food energy and this in turn allowed other members of the society to devote their time to pursuits other than daily survival. As time passed civilizations began to flourish, built by the surplus energy that man acquired. Often these sources of energy took the form of domesticated animals and at other times the form of slaves. Through the use of human slave labor the Egyptians built the great pyramids and temples of Egypt; the Romans, their empire; the Chinese, the Great Wall of China; and Indians, the Taj Mahal.

The human condition changed little until the beginnings of the Industrial Revolution following the invention of the steam engine in the eighteenth century. With this development, man turned from the use of human slaves to the use of mechanical slaves to do his work. Today we are witnesses to a new revolution, the computer revolution, which promises to liberate man in a greater degree than did the Industrial Revolution. For this revolution man will require even more energy to keep the automated factories and cities of the future functioning. If he does not obtain this energy, then his civilization will begin to decay until his demands balance his energy resources. At the present rate of utilization of our stored fossil fuel reserves, your generation may very well see the last of these consumed. Demand for these fuels will double by 1980, based on current rates of consumption. Before these are gone, we must find substitutes or our civilization will be faced with an energy shortage. Where will we get the energy we need? Nuclear energy appears to be one answer to this problem at the present. Greater use could be made of hydroelectric power but this requires the use of land and water and the world of the future will need more of these commodities also.

An alternative to nuclear energy may be the harnessing of solar energy to a greater degree. As new materials and technology come into being, ways may be found to convert solar energy directly into electrical energy inexpensively. Already, the space program has made considerable contributions in this field. Solar batteries are utilized to provide power for most space probes and lunar robot landing vehicles. Before such devices can be used to provide energy for our homes or cities considerably more research and development is needed.

Some scientists have proposed that gigantic orbital solar batteries be developed which would convert the sun's energy into microwave energy and beam this energy to receiving antennas on the earth for distribution. Such a system would require several such solar energy converters several miles in diameter in synchronous orbit to stay locked on their receiving antennas. Until costs of placing men and material into orbit can be drastically reduced such ideas are only dreams.

HOW THE SUN OBTAINS ITS ENERGY It is estimated that our sun has been radiating its energy into space at its present rate for at least 4.5 billion years (the estimated age of the earth) and will continue to do so for as many more. How is it able to produce this energy without being consumed in the process? This question intrigued many scientists in the past and several theories were advanced to explain it. Based upon calculations of the sun's mass, it was found that a chemical process, such as burning, would only provide energy for about 2000 years before the sun's mass was destroyed. In the nineteenth century some scientists thought the sun may produce its energy by contracting under the force of gravity and releasing gravitational energy in the form of heat. But this was discounted later because such a process would have resulted in a noticeable decrease in the sun's size over the centuries.

It was not until the discovery of nuclear reactions in the twentieth century that the answer was obtained. According to Einstein's mass-energy relation, $E = mc^2$, the sun produces its enormous quantity of energy through the conversion of mass into energy. The sun consists primarily of hydrogen and it is through the fusion of two deuterium atoms into one helium atom that this great energy is released. Each second the sun is radiating more energy into space than the human race has used since its life span on earth. Each day the earth receives approximately 8×10^{17} watts of energy from the sun. Over half of this is reflected back into space and the remainder goes into heating the land, oceans, and atmosphere. Man utilizes only a small fraction of this through the consumption of plants and animals, hydroelectric power, fossil fuels, and wind power.

ENERGY IN WIND Until the advent of the steam engine, most of man's energy was supplied by the action of falling water or the winds. Water wheels were used from antiquity to supply water to irrigation canals, but considerable time passed before man learned to harness the energy of falling water to *drive* water wheels to perform useful tasks such as grinding grain.

Wind was used extensively until the 1800's to power ships, and was used to power windmills to grind grain and corn from the late twelfth century A.D. until the nineteenth century. The earliest references to windmills is found in Persian documents dating from the seventh century A.D. These early Persian windmills were used primarily for irrigation and their blades rotated in a horizontal plane about a vertical axis. The familar design, using vertical blades rotating about a horizontal axis, was used by the Romans to power water mills, This design was used also by the Dutch to pump water from low areas to reclaim the land for farming. Variations of this mill were used in early America to pump well water to the surface. This design still survives today in some areas of rural America and in the Australian outback. Besides its use in grinding grains and pumping water the windmill has also been used to generate electrical energy with some success in certain areas.

Even wind power, however, can be considered a form of solar energy, for it is the sun which is responsible for the production of winds in our atmosphere. The heating of the air at the equatorial zones produces movement of air into the upper atmosphere, or creates convention currents.

ACTION FROM ATOMS Little did a small group of scientists realize that they might hold the key to the solution of the energy requirements of the twentieth century when they gathered around a cubic structure of carbon under Stagg Field at the University of Chicago on December 2, 1942, and initiated the world's first controlled nuclear chain reaction. From that modest beginning has come today's growing nuclear power industry. Today as our natural sources of coal, natural gas, and petroleum are decreasing, this industry is increasing in its ability and efficiency to provide inexpensive electrical energy to the world's ever increasing population. It has been predicted by the year 2000 A.D., that the electrical energy requirements in the United States will increase nine-fold. Much of the increase will be met by nuclear power, followed by fossil fuels and hydroelectric power. So, we can see the magnitude of the problem facing this developing industry.

Nuclear energy is utilized today in the form of heat energy produced by the fission of uranium, plutonium, or other fissionable material within a controlled environment. The fissionable material is placed within a shielded, metal container, generally called a reactor core, and allowed to undergo fission under controlled conditions. The heat energy produced by the conversion of mass into energy in the process is used to heat water or some other material flowing through the core to cool it, and this heated water is passed through a heat exchanger to heat more water, turning it to steam to be used to operate conventional steam-turbine electrical generators. The only difference between a nuclear power plant and a conventional power plant is the heat source. Fission provides the heat energy instead of fossil fuels.

Like all power-generating plants, nuclear power plants also represent pollution problems to the environment, but theirs is of a different nature - radioactive contamination and waste disposal. Despite these problems, the nuclear power industry possesses one of the best safety records of any industry and will continue to do so. Because radioactive contamination is somewhat easy to detect and because of the general public's psychological fear of it, elaborate safety precautions have been developed in the nuclear power industry that do not exist in the conventional power industry. With the increasing problem of air and water pollution we can probably expect to see such safeguards applied to them in the near future. In the past, certain groups have organized to prevent the erection of nuclear power plants in certain areas out of unfounded fears that they could explode like a nuclear bomb. Such is not the case, and the few accidents that have occurred in the past have posed no real threat to the population.

Nuclear power, because of its portability, offers the developing countries of the world a better opportunity to become industrialized and members of today's technological age. It is ideally suited as a power source for inaccessible regions such as the Antarctic which lacks natural resources of energy. Whenever a permanent base is established on the moon, it will probably be nuclear-powered.

POWER FROM PLASMA   Closely related to nuclear fission power is nuclear fusion power, but this is still to be realized as an important means of producing electrical energy. In a fusion process, two atoms of hydrogen or deuterium are fused together to form one atom of helium or helium and tritium respectively. Some mass is converted into energy in the process and this energy could be used to generate electrical energy. Work on this method dates from the 1950's after the development of the thermonuclear or "hydrogen" bomb. Thermonuclear reactions of this type are believed to account for the production of energy by the sun.

Research in thermonuclear energy development is along the lines of producing a method to confine a state of matter called a *plasma*, which is a "gas" of nuclei and electrons. In order for fusion reactions to occur, nuclei of deuterium must be brought close enough together to fuse into helium and tritium nuclei. This requires extremely high temperatures, hotter than the interior of the sun, to overcome repulsive electrical forces between the nuclei. Other requirements are low densities for production of the plasma and plasma confinement. These have so far proved to be insoluble.

The advantages of controlled fusion over controlled fission are several. The first and greatest is the almost unlimited source of energy - the oceans. These contain enough deuterium to supply energy at our present rate for some twenty billion years! Unlike nuclear reactors or power plants, fusion reactors would produce little radioactive byproducts and they could produce electrical energy directly without the use of a heat exchanger and the steam turbine generator.

The actual utilization of this form of energy is probably several decades away as much theoretical and experimental work must be done before it is a reality, barring some lucky discovery.

ENERGY, WORK, AND POWER So far we have used the terms energy, work, and power without defining what we mean by them. Everyone has a vague concept of what is implied by these terms. To scientists, these terms have very precise meanings. An accountant or typist may think that they have done a great deal of "work" sitting at their desks and working with their machines all day, but a scientist would not think so. To him work is done whenever a force acts through a distance. He defines work to be the product of a force and the distance through which this force acts. Stated in terms of symbols:

$$W \text{ (work)} = F \text{ (force)} \times D \text{ (Distance)}$$

Suppose we lift a box weighing twenty pounds a vertical distance of two feet. How much work have we done? Using our definition above:

$$W = FxD = (20 \text{ lbs}) \times (2\text{ft}) = 40 \text{ ft-lbs}$$

So, we have expended 40 ft-lbs of work in lifting this box. Where is this work now? To answer this question we must define what we mean by energy.

Energy represents the ability or capacity to perform work. Energy may exist in many forms and still possess this ability. In the previous example, your body supplied the energy, or ability to perform the work, in lifting the box. This energy was stored in your body in the form of chemical energy which you acquired when you ate foods containing energy that was stored up in them through the process of photosynthesis. This energy the plants absorbed from the sunlight that traveled to them through millions of miles of space in the form of radiant energy.

Scientists have developed several terms for distinguishing the origin or nature of various types of energy. Potential energy is *stored* energy. It may appear in various forms. Potential energy is stored in coal, or other fossil fuels burned in your car or kitchen range., Their potential energy is released in the form of *thermal* energy, or the random motion of molecules. This thermal energy may be transformed into energy of motion in your car, causing it to move. This energy of motion is called *kinetic* energy. Kinetic energy, like work, has a very precise definition. It is defined as the product of the square of the speed of an object and one-half of its mass. Stated in terms of symbols:

$$KE = \tfrac{1}{2} m v^2,$$

where KE is the kinetic energy, m is the mass, and v is the speed.

Almost all energy can be considered to be either potential or kinetic energy. Thermal energy is energy due to movement of molecules, hence, it is a form of kinetic energy. The box that you expended work in lifting, stored up that work in its new position. If you released the box, it would fall under the influence of the force of gravity through some distance and this potential energy, or work, stored in the box's position would be transformed into kinetic energy as it fell.

Energy may also take the form of electrical, magnetic, nuclear, or gravitational energy. These energies differ in their origins. Electrical energy is the energy due to electrical charges in motion, or the forces between two electrical charges; magnetic energy is due to the interaction of magnetic fields; nuclear energy is due to the nuclear force that holds the nuclear particles together; and gravitational energy is due to the force acting between two massive bodies. One property that all of these energies have in common is that they can be transformed from one form into another. This transformation occurs between different forms of energy without any of the energy being lost in the process. This fact was first verified by James Joule in the 1840's and has been verified on countless occasions since. This result is known as the *law of conservation of energy* and states: *Energy can neither be created nor destroyed, but only transformed from one form to another.* This law has become one of the cornerstones of modern physics and its validity, although questioned at times, has never been disproved.

As an illustrative example of energy transformations, consider the production of electrical energy for your home. If the origin is hydroelectric power, the original source of this energy is the sun which stores up gravitational potential energy in the water behind the dam. This gravitational potential energy is transformed into kinetic energy by allowing the water to fall some distance where it is diverted into a water turbine. Here the kinetic energy of the water is transformed into rotary mechanical kinetic energy in the turbine. The rotating turbine shaft turns an electric generator armature in a magnetic field. The magnetic field forces electrons in the armature windings to move and produce an electric current. Considerable amounts of energy are "lost" in the form of heat energy in each transformation, and hence, the electrical energy produced does not equal the original potential energy. No energy has disappeared, but some of it is merely in a form that cannot be easily utilized.

Power is a term closely related to energy. Power is a measure of how fast energy is being utilized or transformed. It is defined as the time rate at which energy is changing, or

$$Power = \frac{Energy\ transformed}{Unit\ time} .$$

Electrical power is measured in *kilowatts* and electrical energy in *kilowatt-hours*. Mechanical energy is generally expressed in terms of *horsepower*, or the amount of work a horse can perform in one second. James Watt is the originator of this unit of measurement. He found that a horse could lift approximately 550 lbs a distance of one foot in one second. Hence, one horsepower is equivalent to 550 ft-lbs per second.

THE CHALLENGE AND THE GREAT POTENTIAL Today, while enjoying the highest standard of living in history, man is faced with some of the most perplexing problems in history. Many of these are due to his mismanagement in the use of energy. Our streams and rivers are clogged with the residue of waste products from our factories and cities; our air is polluted with smoke and poisonous gases generated by automobiles and factories; and wildlife is threatened by pollution of the environment by indiscriminate use of pesticides.

How are we going to correct these problems? Until we begin to realize that we cannot continue to produce products and services indiscriminately, and pollute our environment as a byproduct of modern civilization, we will continue to have these problems. What is needed is a planned coordinated approach to these problems. The sociologist must work with the scientists, technologists and architectural planners to improve our cities In order to do this he needs to understand the scope of the problem. One does not possess the solution to one problem if it creates ten more in its place. We must stop implementing the first easy solution that comes along because quite often it is this type of attitude that creates our bigger problems. We must be willing to devote time and effort to determine the basic causes of our problems and once we understand all of the implications, then we should act. Our natural resources are not unlimited, and someday soon we may have to pay the penalty of centuries of neglect of these problems. Unless science and society understand each other's goals and work together to obtain these goals and correct the past problems, then the world of the future will be no better than it is today.

Today we are beginning to see a move in the right direction in some areas. In Europe and Great Britain, regional planning commissions are surveying plans and requirements for cities and are acting to limit their size to prevent them from choking themselves to death. This is being done by building satellite cities at fixed distances from the central city with wide spaces in between. Some communities in the United States are experimenting with this concept of regional planning, but more should be doing so. The location of airports, for example, should not be decided merely on the basis of costs, but rather, how it will affect the overall growth, transportation, and quality of life of our cities. Surely, fast air transportation is important, but if we spend more time in reaching the airport than in our air travel, something is wrong. The citizens of the future should learn from the mistakes of the past and not repeat these, or life will become worse instead of better. Therefore, our citizens must understand what science has accomplished in the past, what it is doing at present, and how it can be used in the future for man's benefit. Science cannot solve our problems if the citizens who vote on priorities do not understand what these priorities are.

What does all this mean to you? It means that you need to prepare yourself for a changing world situation. The world of twenty years ago is not the world of today. Neither will be the world of 1980. Although a college education can help you in this preparation, you must continue to study and prepare yourself for whatever changes occur. Our world is still a Darwinian world, and only those prepared for change will be able to make the most of it.

SUGGESTED OUTSIDE READINGS:

H. Rau, *Solar Energy*, MacMillan, New York, 1964.

Eric M. Rogers, *Physics for the Inquiring Mind*, Princeton University Press, Princeton, New Jersey, 1960.

ARTICLES FROM *Scientific American:*

A. Briggs, "Technology and Economic Development" (September 1963).

K. Davis, "The Urbanization of the Human Population" (September 1965).

K. Lynch "The City as Environment" (September 1965).

E.S. Mason, "The Planning of Development" (September 1963).

S.H. Schurr, "Energy" (September 1963).

G. Sjoberg, "The Origin and Evolution of Cities" (September 1965).

SCIENTIFIC AMERICAN READER:

L.R. Hafstad, "Reactors".

R.E. Marshak, "The Energy of Stars".

L. Svirsky, "Atomic Power".

STUDY QUESTIONS:

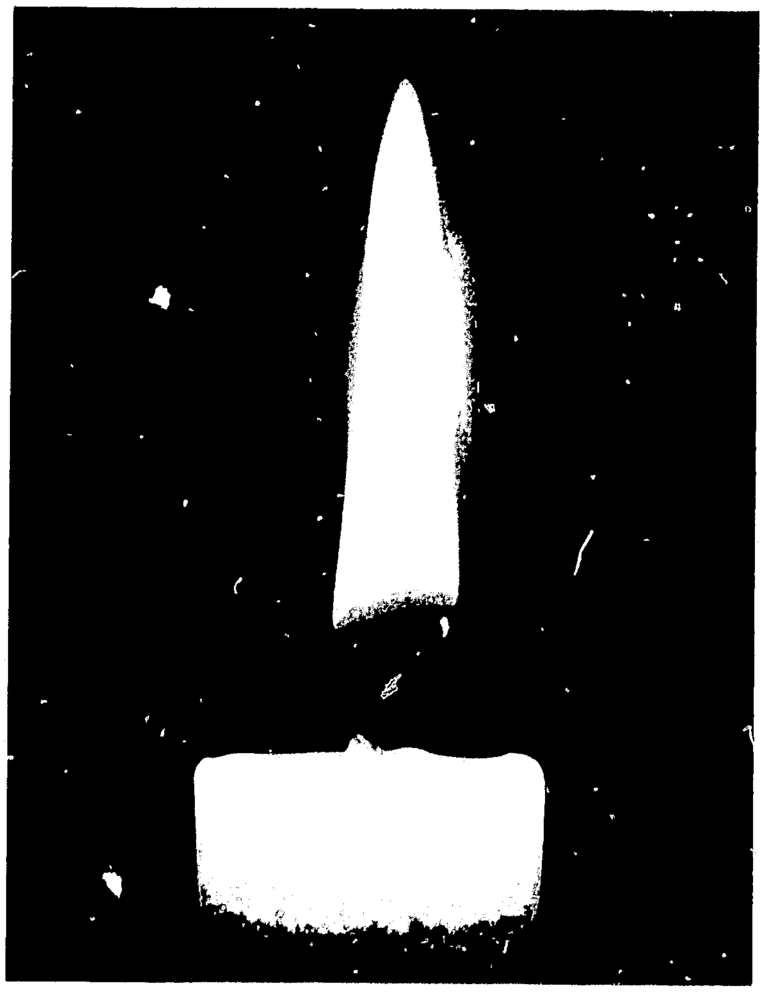
1. Discuss the origin of the principle of conservation of energy.

2. Why is it that we must use energy to survive in present day technology? Discuss the various types of energy that you use everyday in your life.

3. Write a short essay on the availability of energy *now* and in the future in different countries of the world.

4. Describe the energy interchanges that may occur as and after an automobile is started.

CHAPTER 4 ATOMS IN MOTION



INTRODUCTION Today when we think of *heat*, we may envision many different things: a crackling fire in winter, a summer day at the beach, or warm food cooked in a modern oven. These are all things that connote a feeling of "warmth," but have you ever stopped to think what "heat" really is and how man came to understand and use it to warm his home, cook his food, and to provide energy for the many industrial uses of modern civilization? To get a clearer understanding of all these things, we need to begin our inquiry in antiquity.

MAN DISCOVERS FIRE - THE GIFT OF THE GODS Man's probable first encounter with heat was the earth's sun which provided energy to warm his body and light to see. As prehistoric man slowly advanced, he discovered that skins and fur of animals helped to provide protection against the elements when the sun was absent. Fire was probably discovered accidently near a volcano or in a forest fire started by a bolt of lightning. Many ancient fables were devised to explain the origin of fire - you may recall the story of Prometheus, in Greek mythology, who is credited with stealing fire from the gods on Mount Olympus and bringing it to earth, thereby making man independent of the gods who controlled the sun. The mythologies of many other ancient cultures have similar stories to account for the origin of fire. Prehistoric man probably discovered by accident that certain foods tasted better and were better digested when heated in a fire and hence the art of cooking was invented. One Chinese legend attributes the discovery of cooking to an unfortunate peasant whose house burned down, roasting a pig trapped inside. As man advanced from the Stone Age toward the Iron Age, he probably discovered the ability to ignite certain substances with sparks produced by striking stones together, thereby giving him the ability to produce fire at will. The Iron Age found man refining metals by placing metallic ores in fires. He found that heat allowed him to work metals into tools, weapons, and jewelry. Heat was also used to fire pottery made of clays into durable, useful forms and to produce bricks for construction of houses. The crude tools produced by these early blacksmiths enabled man to cultivate fields, and agriculture developed which in turn gave rise to cities and the first permanent civilizations. In order for any civilization to develop, man must be free to do other things than seek food and shelter.

The ancient Romans were able to steam heat their houses by piping in heat from volcanic fields through lead pipes. This was the first application of steam heating. As time progressed, man discovered that heat would reform materials and enable him to make numerous other useful articles.

During the Middle Ages alchemists attempted to utilize heat to convert baser metals such as lead into gold. From their experiments and empirical observations of the effects of heat, a theory of heat gradually evolved. Aristotle held the belief that all matter was composed of four basic elements - earth, water, air, and fire. The alchemists believed that fire could be combined in various proportions with other elements to produce gold and other precious metals, as well as an Elixir of Life to extend life indefinitely. Indeed, the noted alchemist and founder of medical chemistry, Paracelsus (1493-1541), remarked, "What is accomplished by fire is alchemy; whether in the furnace or in the kitchen stove." These alchemists, through destructive distillation of various materials, manufactured many chemicals useful in refining metals, making soaps, and medicinal applications.

The alchemists sought ways to arrange their chemical reactions into Aristotelian concepts of matter and in doing so developed a theory of heat known as the Phlogiston Theory. The German alchemist Stahl (1660-1734) produced a system of chemistry based on water and three types of earths, one of which he called *phlogiston*. Chemical reactions were explained in terms of these four quantities. For example, sulfuric acid + phlogiston = sulfur, and metal oxides + phlogiston = metals. Phlogiston was believed to be a weightless, invisible substance which was emitted by substances during combustion. Reactions were explained in terms of the presence or absence of phlogiston.

The Phlogiston theory of heat dominated chemistry until about 1765, when Henry Cavendish (1731-1810) began a detailed study of gases. The discovery of the pure gases such as oxygen by Joseph Priestly in 1774, and the subsequent experiments by Lavoisier in 1778, showing that oxygen and metals combined to form metal oxides, gradually led to the decline of the phlogiston theory. Oxidation processes replaced the phlogiston concept in chemical reactions.

In 1789 Lavoisier published a textbook based on the chemistry of oxidation entitled *Traite' Elementaire de Chimie* in which he introduced a new system of chemical nomenclature which is still used in a modified form by chemists today. In his book, he listed chemical elements and included among them Light and Heat or *Caloric* as he called it. Caloric was a fluid substance present in fire and capable of flowing into matter. Lavoisier believed that caloric was not necessarily a fluid substance but rather that it was the *cause* of various heat phenomena. Caloric was considered to be related to Light, in that they both produced similar effects in certain processes such as photosynthesis in plants.

Lavoisier attributed the different states of matter to different quantities of caloric present in a substance. Solids contained little caloric while liquids contained more, and gases still more. He envisioned caloric as a substance that penetrated between the particles of matter and caused their separation. The greater the caloric content, the greater the separation of particles and the material expanded.

The sensation of warmth was attributed to an accumulation of caloric, whereas the sensation of coldness was due to the absence of caloric. If one touched a "hot" object caloric flowed from the object into the hand. If a "cold" object was touched, caloric flowed from the hand into the object. Touching an object of the same temperature as the hand resulted in no transfer of caloric.

To explain the fact that some substances melted while others did not when both were placed in the same fire together, Lavoisier introduced the concept of heat capacity of different materials to contain caloric. Fires contained "free caloric" which could enter the material when it was placed in the fire. The reason that one substance accumulated more caloric and melted was explained by assuming that different substances were composed of spherical particles while others were composed of cubes, tetrahedrons, etc. Because of the different packing arrangements, some substances had more empty space between particles and could contain more caloric.

Lavoisier also defined a *specific heat capacity*, or the quantity of caloric required to change the temperature of bodies of the same weight an equal number of degrees. These terms are still used today in a modified form. Lavoisier considered heat to be a sensation produced by the *movement* of caloric, a concept not far removed from the present day.

The caloric theory did not die as easily as the phlogiston theory. Many experimenters and experiments were necessary to disprove its validity. One of the earlier experimenters to undertake this task was Benjamin Thompson, Count Rumford. A Tory sympathizer who fled America during the Revolutionary War, Rumford rose to be a Count of the Holy Roman Empire and the Inspector General of Artillery of the Bavarian Army. During the course of his duties as supervisor of the boring of the cannon for the Army, Rumford became impressed with the huge quantity of heat produced by the boring process. According to the caloric theory, friction would gradually force caloric from the cannon. To test its validity, Rumford had the cannon bored under water and attempted to measure the quantity of caloric liberated by measuring the time required for the water to boil. He was able to show that continuous boring produced heat continuously with no decrease in caloric. He concluded that the heat was generated by the friction of the drill in the boring process and that *motion* was necessary for the production of heat. Rumford attempted some quantitative measurements of the heat produced in the drilling process as a function of the amount of work done by the horses used to turn the lathe. While he did not reach any definite quantitative conclusions, he did remark that it would be more efficient to boil the water by burning fodder rather than feeding the horses and having them heat the water by friction.

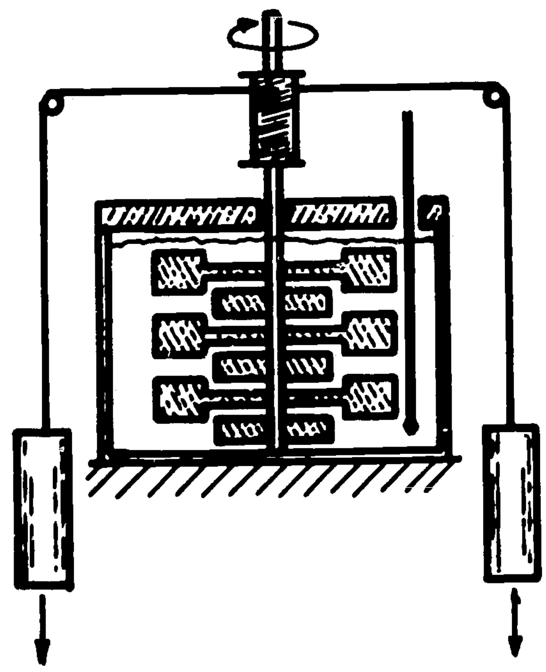


Diagram of Joule's experiment. The falling weights cause the paddle-wheels to rotate, heating the water.

About 1840 James Joule began some accurate quantitative experiments concerning the production of heat by mechanical means. His experimental apparatus consisted of a vessel containing a set of paddle-wheels and a quantity of water. The paddle-wheels were turned by means of strings attached to falling weights. As the weights fell, they caused the paddle-wheels to rotate which heated the water by friction. By measuring the change in temperature of a given quantity of water and by knowing the distance through which the weights fell, Joule was able to relate the mechanical work done by the weights to the quantity of heat produced in the water.

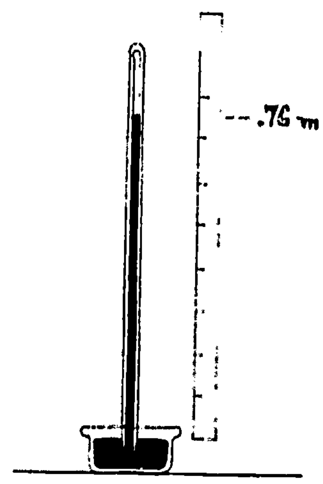
The unit of heat that Joule used to measure the heat produced was the *calorie*, which is the amount of heat required to change the temperature of 1 gram of water by one degree on the Celsius temperature scale. The mechanical work is calculated by multiplying the weight by the distance it falls. Its units are newton-meters or foot-pounds. After these experiments, the unit, the newton-meter, was renamed the *joule* in honor of James Joule. The relationship that Joule determined between mechanical work and heat was that 4.186 joules (newton-meter) of work produced 1 calorie of heat. This relationship today is known as the Joule Mechanical Equivalent of Heat.

One of the important results of this work was that energy in the form of mechanical work could be *transformed* into heat. This conclusion led Joule, Mayer, and others to postulate a law of conservation of energy:

Energy is neither created nor destroyed in physical processes, but merely transformed from one form to another.

This law has since become a cornerstone of modern physics and its validity proven in every type of physical process envolving every type of energy.

HEAT ENGINES Around 1700 devices were designed and built by several people which operated on steam and atmospheric pressures. Pressure, you may recall, is a force per unit area. Galileo and his student Torricelli are credited with the invention of the first instrument for measuring atmospheric pressure, the barometer. Their simple barometers indicated fluctations of the earth's atmospheric pressure by indicating how high a column of mercury is supported in a closed glass tube inverted in a dish of mercury by the weight of air above the open dish.



A barometer. Atmospheric pressure will support a column of mercury to a height of 76 cm.

The early steam engines were cylinders fitted with a moveable piston which moved back and forth when pressures changed on different sides of the piston. Usually steam was admitted into one side of the cylinder at pressures greater than atmospheric, and the piston was pushed up. If the cylinder was then cooled, the steam inside condensed into water, forming a partial vacuum, and the atmospheric pressure, being greater, would push the piston down. This process could be repeated and useful work could be obtained from the engine to pump water from mines and run machinery such as looms.

The inventors of these devices possessed little theoretical knowledge concerning the nature of heat itself, but were able to design workable engines based entirely on the expansion of gases and atmospheric pressure. In fact, analysis of these engines led scientists to a more complete understanding of heat. Basic research does not always precede applications or technology. The inventions of Thomas Edison are another example of this truth.

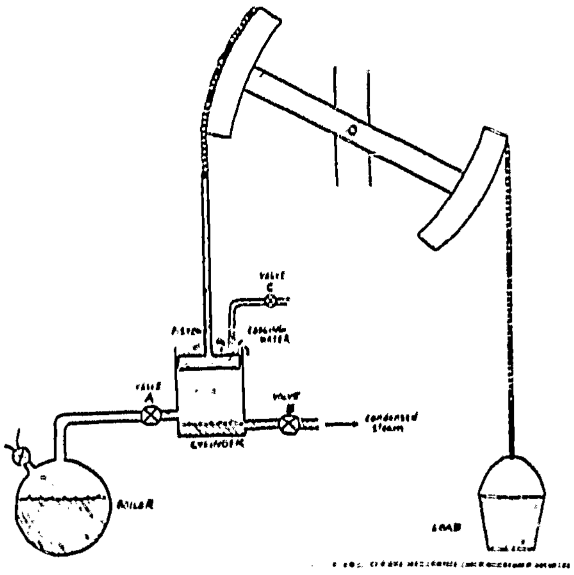


Diagram of an early steam engine.

The analysis of the operation and efficiencies of these steam or heat engines by William Thomson (Lord Kelvin) Carnot, and Clausius led to a theoretical understanding of heat and its relationship to work and energy. This theoretical knowledge is today summarized by the branch of physics called *Thermodynamics*. The practical applications of this knowledge of thermodynamic principles led the European countries and America of the 1800's into an era of Industrial Revolution with repercussions in the social and political arenas. More will be said concerning these effects after we discuss the modern theory of heat.

THE NATURE OF HEAT With the introduction of the atomic theory of matter by John Dalton in 1808, the nature of heat began to be placed on a more sound theoretical foundation. All matter was considered to consist of atoms and molecules bound together in a definite arrangement. In solids the atoms were closely packed and exerted strong attractive forces on each other. Liquids consisted of atoms or molecules loosely bound together and able to move about, interacting only weakly with their neighbors. In gases, atoms were separated by relatively large distances and only interacted through collisions or contact. Because the forces determining their motion were simpler to understand, gases became the chief center of interest among scientists who sought to gain a deeper insight into the nature of heat in the nineteenth century.

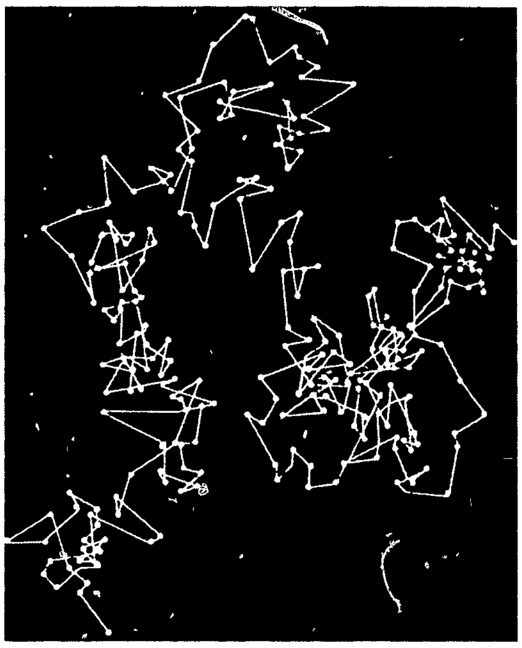


Diagram of a particle in Brownian motion.

In 1828 the English botanist Robert Brown observed the behavior of microscopic particles suspended in a liquid. Viewed through the microscope, the particles seemed to move erratically in all directions, changing directions violently. This phenomenon became known as Brownian movement. Similar results can be observed with smoke particles which were colliding with invisible molecules. The random motion of the molecules produced the erratic motion of the particles. Heating the solution caused the molecules to increase their violent motion. This effect and others led Joule and other scientists to believe that heat was really due to the random *motion* of atoms, a conclusion that Rumford suspected earlier. Since the motion or kinetic energy of molecules was considered the reason for the expansion or contraction of gases, this area of investigation became known as the *Kinetic theory of gases*. By applying the laws of Newtonian mechanics to these systems of atoms (the gas), the scientists were able to explain phenomena discovered earlier and accurately predict new phenomena.

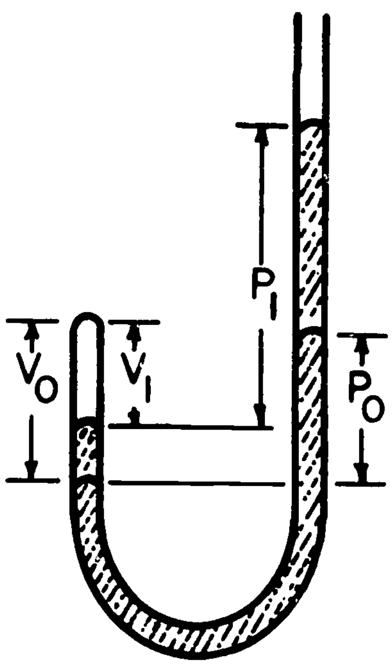


Diagram of Boyle's experiment. The volume is V_0 when the pressure is P_0. Increasing the pressure to P_1 decreases the volume to V_1.

THE GAS LAWS In 1660 Robert Boyle discovered that the pressure, *P*, and volume, *V*, of a gas varies inversely if the temperature, *T*, was held constant. Stated in symbols:

$$P \propto \frac{1}{V}, \text{ if } T = \text{constant, or } PV = \text{constant } (K).$$

The kinetic theory of gases was able to explain this result by assuming that atoms behaved like small, hard spheres which transferred momentum and kinetic energy on contact. By increasing the motion of particles one could increase the number of collisions of particles with themselves or the walls of their container. If the volume of the gas is fixed and heat energy is added to the gas, the pressure in the container would rise. Cooling the container lowered the pressure. Pressure is considered to be the *average force* exerted by the gas particles *per unit area* of the container. Increasing their kinetic energy increases the number of collisions of particles with the walls, or the force transferred to the walls. If the walls are *elastic*, increasing the pressure results in an increase in volume if the temperature is constant. One sees examples of this in inflating a toy balloon or an auto tire.

If one raises the temperature of a confined gas at constant pressure, an increase in volume also occurs. This result is known as Charles' Law and was discovered about 1660 by the French scientists Jacques Charles and Gay-Lussac. Stated in terms of symbols:

$$V \propto T, \text{ if } P = \text{constant, or } V = KT,\ P = \text{constant.}$$

Combining these two laws resulted in a relationship that is known as the General Gas Law:

$$PV = kT, \text{ where } k \text{ is constant.}$$

This law holds for most gases at low pressures.

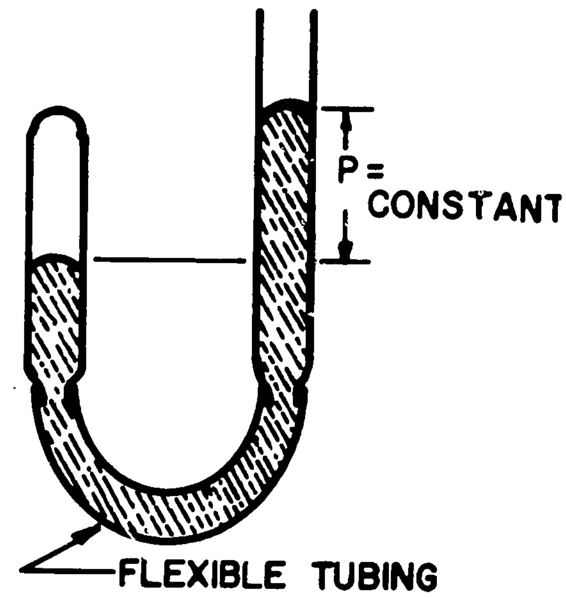
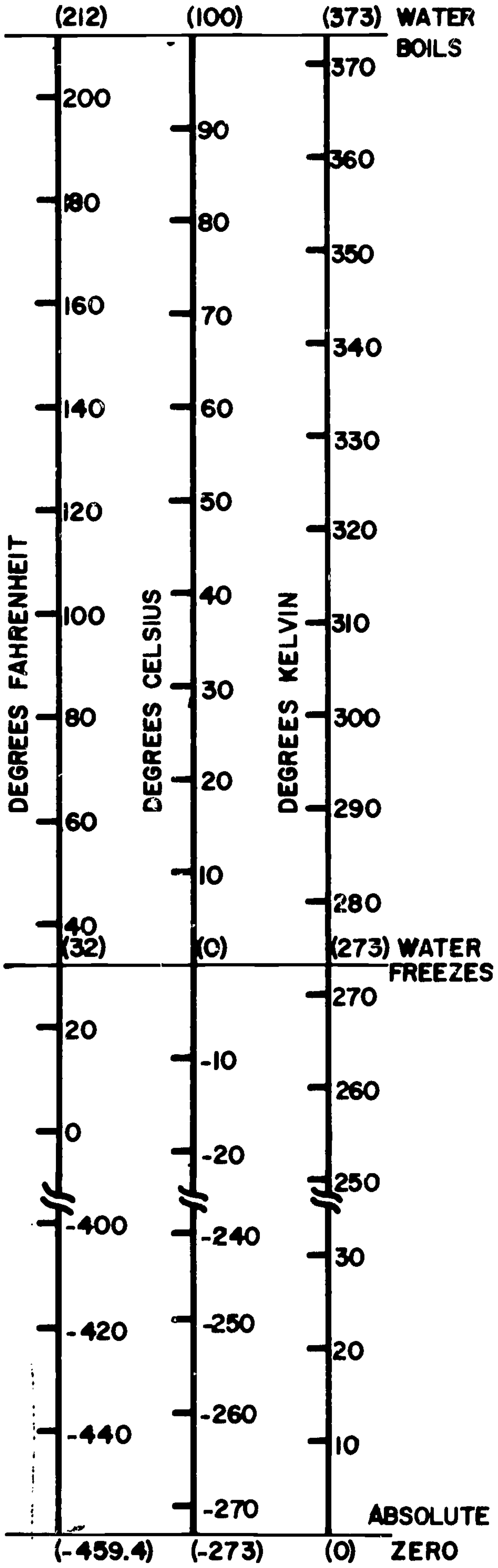


Diagram of Charle's experiment. A volume of air is trapped above the mercury in the left arm. The applied pressure is held constant by moving the right arm up or down as the volume varies with temperature.

TEMPERATURE AND TEMPERATURE SCALES Closely related to the concept of heat is the concept of *temperature*. To the scientist, temperature is more than a physiological sensation; it is a measure of the *average kinetic energy* of the molecules. Whenever a body absorbs heat energy, its molecules move faster and its temperature is said to rise. The molecules collide more frequently and tend to take up more volume, hence the material expands. This fact was used by Galileo to build a thermoscope, a device to indicate changes in temperature. The thermoscope consisted of a glass bulb and tube containing water or mercury. As the energetic air molecules collided with the container kinetic energy was transferred to the container and then to the liquid. As the average kinetic energy of the liquid increased, the liquid expanded and rose to a new level in the tube. This device was open to the atmosphere so it was actually a thermo-baroscope.

The first practical thermometer or device to *measure* changes in temperature was probably invented by some of Galileo's successors in the scientific society Accademia del Cimento which Galileo founded. A glass bulb and tube containing wine was sealed to the atmosphere and an arbitrary scale attached. Each person who constructed thermometers in this period usually assigned his own numbers to the scale. As a result, much confusion arose when one attempted to correlate various findings in experiments.

The German instrument maker Fahrenheit (1686-1736) was one of the first to standardize his themometer scales so that they could be easily reproduced. He chose as reproducible fixed points the temperature of an ice-brine mixture and the temperature of the human body. He arbitrar-



$$°F = \frac{9}{5}(X°C) + 32°$$

$$°C = \frac{5}{9}(X°F - 32°)$$

$$°K = X°C + 273°$$

ily assigned the number 32 to the position that the alcohol column stood when the thermometer was placed in the mixture and left awhile. The second fixed point was determined by the distance that the alcohol rose when the thermometer was held in the mouth. Fahrenheit assigned the number 96 to this position. By dividing the distance between 32 and 96 into equal parts, he arrived at the Fahrenheit degree. It just happened that the column rose to 212 when placed in boiling water, a third fixed point used today.

A second type of scale was devised by Andreas Celsius in 1742. He chose the melting point of ice (or the freezing point of water) as 100 and the boiling point of water (steam point) as 0. Later, in 1750, a colleague of Celsius, Mårten Strömer, inverted the Celsius scale and assigned 0 to the ice point and 100 to the steam point, but we still refer to this scale as the Celsius scale today.

A third type of temperature scale, the Absolute or Kelvin temperature scale, was developed by William Thomson (Lord Kelvin) in 1868. This scale was developed after studies by Carnot and Kelvin concerning the efficiencies of heat engines operating between two temperatures, t_2 *and* t_1. Carnot discovered that the efficiency of a heat engine was given by the equation:

$$\text{Efficiency} = \frac{\text{Energy Output}}{\text{Energy Input}} = \frac{t_2 - t_1}{t_2},$$

where t_2 is the highest temperature of the engine and t_1 the exhaust temperature. Kelvin reasoned that there must be a lower limit to the exhaust temperature t_1 if the Carnot efficiency cannot exceed 100% (if it did this it would mean that you get more energy out than you put into the machine and this violated the conservation of energy principle). Kelvin set $t_1 = 0°$ as the Absolute Zero of temperature. On his scale Kelvin set $0°C = 273°K$ and $100°C = 373°K$; therefore $0°K = -273°C$ or about $-459°F$. Today, low temperature physicists have approached temperatures as low as $0.000006°K$ in the laboratory. At this temperature all matter exists only in a solid form because the atoms are in their lowest possible energy states. Temperatures this low are approached on the outer planets of the solar system and in deep space. Any substance that is a gas on earth would be frozen solid there.

STATES OF MATTER Matter may exist in a solid, liquid, or gaseous state depending upon the energy content of its atoms. In solids, strong electrical forces hold atoms rigidly in position so that solids have a definite, fixed form. By adding heat energy to the substance, one can cause these atoms to vibrate more vigorously about their equilibrium or resting positions. Continuing to add energy will cause such rapid movement that the electrical forces cannot constrain the atoms and they will move farther apart. At this energy content, the solid is said to "melt" and it forms a liquid. Reversing this procedure causes the atomic or molecular motion to decrease to the point where the electrical forces can arrange the atoms into some definite order and the substance is said to "freeze." Each substance made up of different atoms has its own unique melting or freezing point temperature because the electrical forces between atoms depends upon the types of atoms composing the substance. For example, a given quantity of heat will cause one gram of ice at 0°C to melt; but it will not melt one gram of iron at 0°C. The forces bonding the iron atoms together are stronger than those bonding water molecules together and require more heat energy to break. The ability of materials to absorb heat energy for a given mass of material and for a given change in temperature is called its *specific heat capacity*. One gram of ice, for example, requires about 80 calories of heat energy at 0°C to break down the molecular bonds and change it to water at 0°C. This change of state is called a *phase change*. To change the temperature of one gram of *water* from 0°C to 1°C requires the addition of 1 calorie/gram for each degree of temperature up to the steam point. Water heated to 100°C (at sea level) will not boil at this temperature until an additional amount of heat is added. The reason for this is that the remaining bonds between the water molecules must be broken. An additional amount of heat equal to about 540 calories/gram must be absorbed by the water at 100°C in order to change it into a gas (steam). This phase change is referred to as "boiling." The temperature at which the liquid becomes a gas is known as the "boiling point." Each substance has its own unique boiling point temperature so it is possible to identify substances by measuring their boiling points. Chemists use this method to identify and separate two liquids; the substance which has the lower boiling point will evaporate first leaving the substance with the higher boiling point behind. By condensing the gas (removing the heat), one can recover the first liquid. Such a process is known as *fractional distillation* and is used to refine crude petroleum into numerous useful components such as gasoline, kerosene, and mineral spirits.

According to the General Gas Law, the temperature at which a substance boils is also dependent upon the pressure. Water is said to boil at a temperature of *100°C* at sea level atmospheric pressure. At a higher elevation, water would boil at a lower temperature because of the reduced atmospheric pressure. There are more air molecules per unit area at sea level to collide with the surface of the liquid and they carry away some of the heat energy, hence the substance requires a higher temperature to boil. This is why one can cook foods faster with a pressure cooker than in an open pot. By increasing the pressure in the cooker, the water requires a much higher temperature than *100°C* to boil; hence the steam produced is superheated and the foods can cook much faster.

For similar reasons, astronauts and pilots who fly at extremely high altitudes must wear pressurized suits. At these altitudes or in the vacuum of space, the heat of their bodies is sufficient to cause their blood to boil due to reduced pressure on their bodies. One can show that water will boil in the presence of ice by placing a beaker of ice and water in a vacuum chamber and evacuating the air inside. When the pressure is low enough, the water will boil at *32°F* or *0°C*.

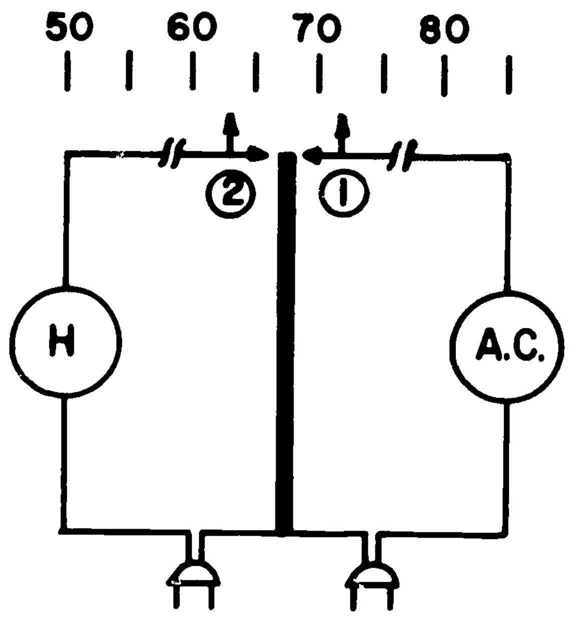


Diagram of automatic thermostat control. The contact (1) may be set for a given temperature for the air conditioning circuit. The contact (2) controls the heating circuit. If the bimetallic strip expands, it makes contact with (1), turning on the air conditioning. If it contracts, it makes contact with (2), turning on the heating circuit.

EXPANSION OF SOLIDS Almost all substances tend to expand upon being heated due to the increased motion of their atoms. One can use this phenomenon in the manufacture of many useful products. We have already seen how the expansion of liquids can be used to make thermometers. Another useful application of this property of matter is in the *control* of temperatures in buildings and homes. A device designed for this purpose is called a *thermostat*. Many buildings and homes today are equipped with automatic heating and air conditioning systems. How is the temperature of these systems regulated? A very simple method consists of alloying two different metal strips together to form a device known as a bimetallic strip. Consider a strip consisting of copper and silver alloyed together. Silver is a much better conductor of heat than copper and expands at a faster rate. Since both metals are mechanically joined together, the result is that the silver pushes the copper over. Upon cooling, the silver loses heat faster than the copper so it contracts faster and pulls the copper over in the opposite direction. By using the bimetallic strip as a contact in an electrical circuit, one can control the operation of the heating or air conditioning system.

By using the copper side of the bimetallic strip as the contact for the air conditioning and the silver side as the contact for the heating one can keep the room temperature within a certain range. Two contacts on each side of the strips are movable and allow one to set the temperature range. If the room is warm, the silver side expands and causes the copper side to make contact with the air conditioning circuit. The air conditioner cools the room and the strip contracts, breaking contact and turning the air conditioning off. As the room warms back up the above procedure is repeated.

If the room is cold, the silver side contracts and it makes contact with the heating circuit. The heater heats the room and the strip expands and breaks contact. As the room cools off the above procedure is repeated.

Bimetallic strips shaped like a spiral spring can also be used as a thermometer. An indicator needle is attached to one end and the other end is rigidly fastened. As the strip expands or contracts the needle can move across the scale and denotes the temperature. This type of thermometer is often found in the home to measure the temperature of roasts and meats.

TRANSFER OF HEAT ENERGY Heat energy is transferred from one place to another by three basic processes - conduction, convection, and radiation. Conduction is the transfer of heat by contact of more energetic particles with less energetic particles. The average kinetic energy of the molecules increases and the temperature of the substance rises. Heat is conducted from one end of a metal rod to the other end if one end is held in a flame. The molecules of the rod in contact with the flame gain kinetic energy by collision, vibrate faster, and their energy is transferred from one molecule to another until the entire rod is heated. Cooking on an electric or gas range utilizes this method of heat transfer. The pot absorbs heat by conduction and transfers this to the food by conduction and convection.

Convection is the transfer of heat energy by the *movement* of energetic molecules in a liquid or gas. The more energetic molecules, through collisions with less energetic molecules, lose heat energy and the less energetic molecules gain heat energy. If cold water is poured into hot water, the mixture will gradually reach a common temperature due to convection currents. This phenomenon was first successfully explained by Count Rumford who used the concept of convection currents to design more efficient fireplaces, stoves, and clothing. Today most homes are heated and cooled by convection currents. Air in contact with a hot surface gains energy and collides with colder air, thereby heating the room. Steam heaters and forced air heaters utilize this method.

Convection currents in the atmosphere also account for a great deal of our weather. As the sun's rays heat the equatorial regions of the earth, the warm air rises in convection currents into the colder regions of the atmosphere. These convection currents form much of our winds and sometimes develop into tornadoes and hurricanes.

Count Rumford also discovered the third method of heat transfer-radiation. During the experiments on conductivity of heat, he found that if he placed a thermometer in a vacuum chamber and heated the chamber that the thermometer showed an increase in temperature. Here we have an example of heat being transferred by radiation - without the presence of molecules or convection currents. Radiation accounts for the process by which the sun's rays reach the earth through millions of miles of empty space.

Heat energy is actually a form of electromagnetic radiation like radio waves called infra-red radiation which require no medium for propagation. A common form of an infra-red radiation emitter is the heat lamp. All living things also emit infra-red radiation.

One common means of preventing heat loss or absorption is by means of a vacuum bottle, or thermos. This is a double-walled bottle with a vacuum between the walls. Heat transfer by conduction and convection is minimized by the vacuum. Some heat is transferred by radiation, so the walls of the vacuum bottle are silvered like a mirror to reflect the radiation.

During the day heat is absorbed by the earth and at night it is radiated back into the atmosphere. If the atmosphere is cloudy, the cloud cover at night helps to reflect and trap the radiation and the night is not so cold. On clear nights the heat is radiated away into space and, usually, cooler temperatures result.

Air is a rather poor conductor of heat if trapped. Fur-bearing animals utilize trapped air in their fur to help keep in body heat. Humans require clothing to perform the same function. Skin divers wear rubber suits which they fill with water before diving into the cold water. The trapped water is warmed by the diver's body and acts as an insulator against the cold of the depths. Likewise, houses are constructed of double walls with a dead space of air between them to help insulate the interior of the house against extremes in temperature outside.

TECHNOLOGICAL DEVELOPMENTS AND THEIR IMPACTS ON SOCIETY

The invention of the steam engine gave man a cheap source of power which freed him from many back-breaking tasks and presented him with many new problems. The rise of steam power and the industrialization of Europe and America in the eighteenth and nineteenth centuries brought both promise for a better future and unpleasant working conditions in the rising factories. As factories replaced small shops and homes as the sources of products, exploitation of the working class by unscrupulous factory owners and managers began. This was a period of great social unrest in England as testified to by many writers of the period, such as Charles Dickens, Karl Marx, and the poets, Blake and Keats.

Eventually, these conditions gave way as workers united to protest their grievances and labor unions emerged to protect their rights. As sources of power became cheaper through the development of more efficient engines such as the steam turbine, the standard of living rose and the cost of products declined.

During this period other types of heat engines were developed. An English engineer named Charles Parsons invented the steam turbine in 1884. This engine stimulated the development of great ocean liners such as the *Titanic*, and produced cheaper electrical power. Today this machine produces most of the electrical power consumed in the United States, whether the heat source is petroleum, coal, nuclear, or water. A variation of the steam turbine, the gas turbine, propels giant airliners at speeds approaching the speed of sound and may someday power your car.

In mentioning cars, we must also mention another class of heat engine - the reciprocating internal combustion engine which is used to power these vehicles. This engine was invented by H. Otto, a German inventor, in 1876. The operation of this engine is referred to as the Otto cycle. It consists of four parts or strokes. The first is the *intake stroke* which admits an air-gas mixture at atmospheric pressure into a combustion cylinder by the downward motion of a piston which results in a decrease in pressure inside the cylinder. The second stroke is the *compression stroke* in which the mixture is compressed into a small volume, raising its temperature, by the upward movement of the piston. At the top of the piston's travel, an electric spark ignites the mixture which explodes, forcing the piston downward with greater force. This part of the cycle is referred to as the *power stroke*. As the piston rises again it forces the spent gases out of the cylinder in the *exhaust stroke*.

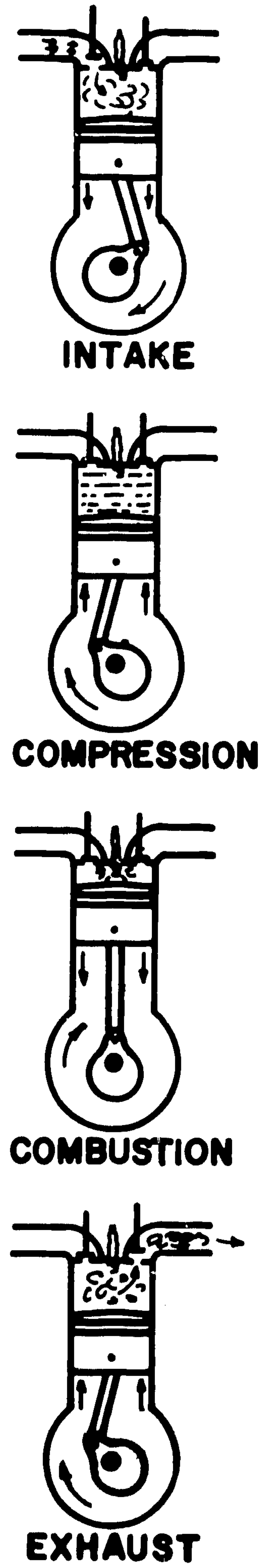


Diagram of the Otto cycle.

SUGGESTED OUTSIDE READINGS:

Sanborn C. Brown, *Count Rumford: Physicist Extraordinary*, Doubleday (Anchor), Garden City, New York, 1962.

F. Cajori, *A History of Physics*, Dover, New York, 1962.

A.R. Hall, *The Scientific Revolution 1500-1800*, Dover, New York, 1962.

K.B. Krauskopf and A. Beiser, *The Physical Universe*, McGraw-Hill, New York, 1967.

R.T. Lagemann, *Physical Science: Origins and Principles*. Little, Brown and Company, Boston, 1963.

D.K.C. MacDonald, *Near Zero: The Physics of Low Temperatures*, Doubleday (Anchor), Garden City, New York, 1961.

Project Physics, *An Introduction to Physics: The Triumph of Mechanics*, Holt, Rinehart and Winston, New York, 1968.

J.F. Sandfort, *Heat Engines: Thermodynamics in Theory and Practice*, Doubleday, (Anchor), Garden City, New York, 1962.

ARTICLES FROM *Scientific American:*

J.R. Clark, "Thermal Pollution and Aquatic Life", (March 1969).

E.S. Ferguson, "The Origin of the Steam Engine", (January 1964).

STUDY QUESTIONS

1. What is meant by heat capacity? Specific heat capacity?

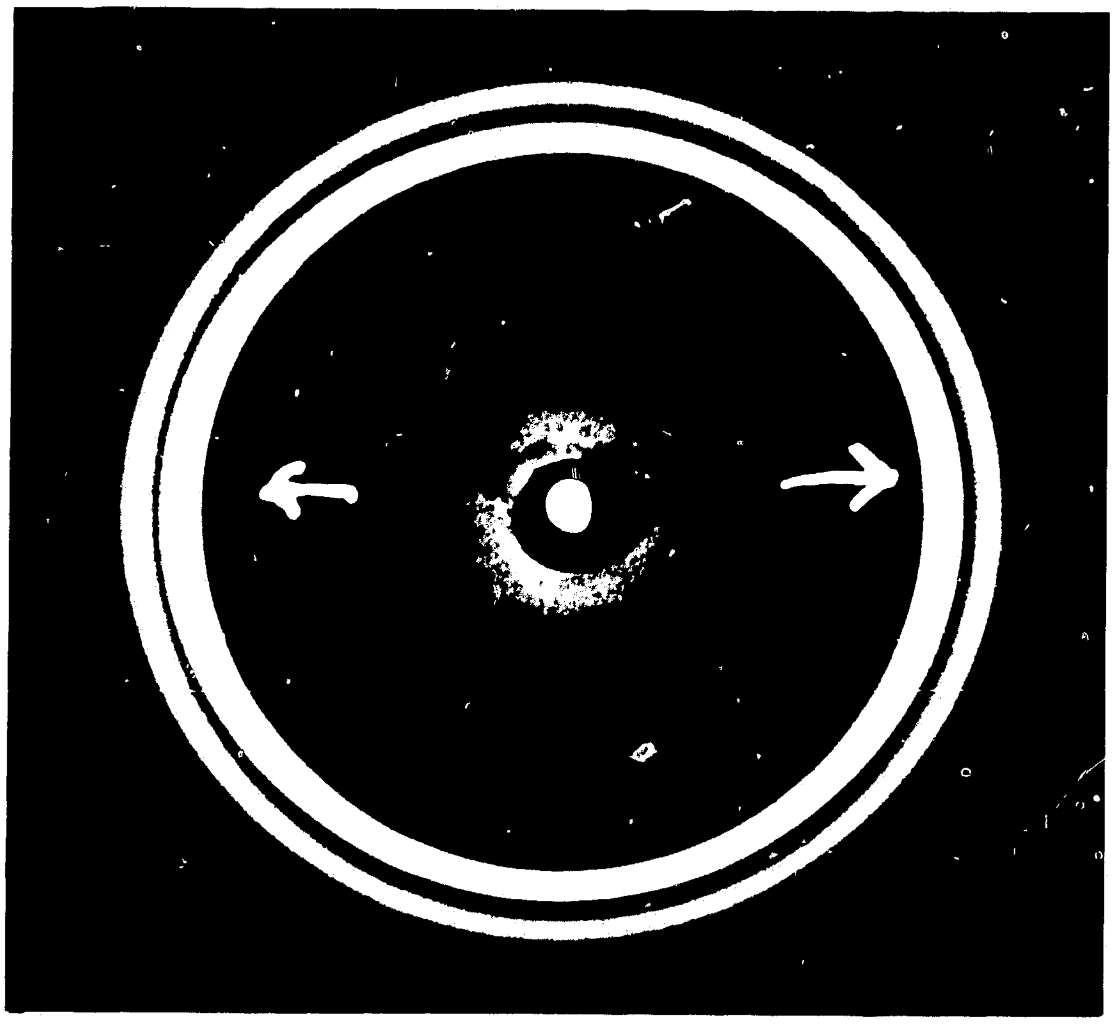
2. How did Rumford's experiments with cannon-boring disprove the caloric theory of heat?

3. Discuss the modern theory of heat.

4. What is temperature?

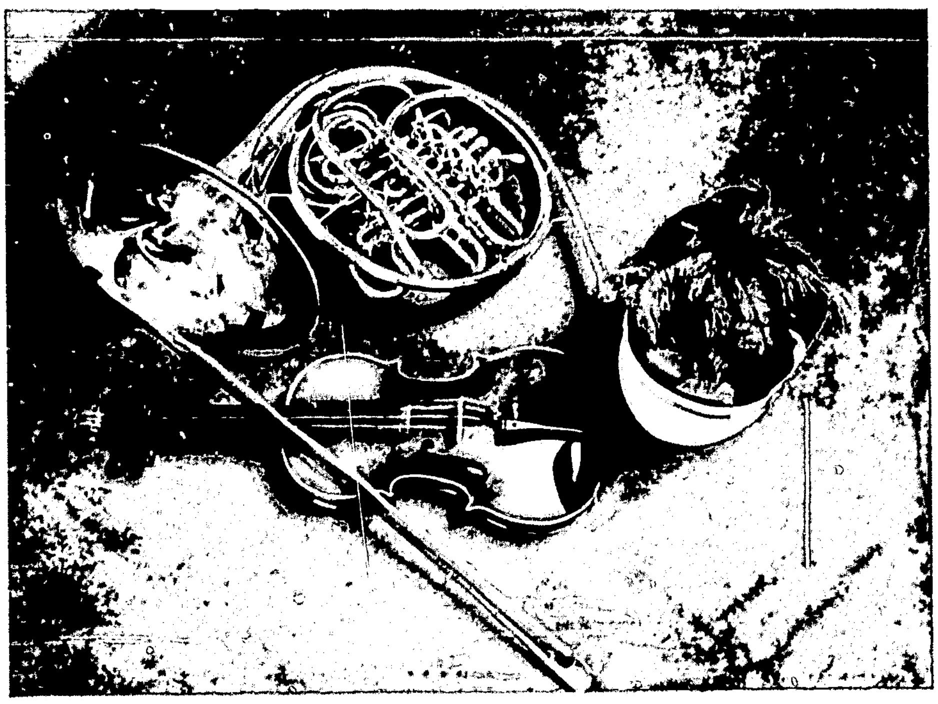
5. How did the invention of the steam engine advance theoretical work in the study of heat phenomena?

6. How is heat transferred?



X-ray diffraction pattern of atoms in a dielectric material.

CHAPTER 5 SOUND



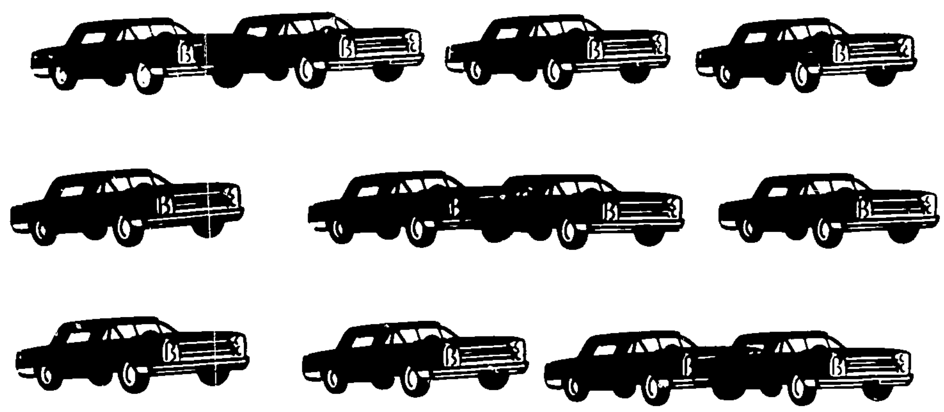
COMMUNICATIONS AND SCIENCE Today, some of our greatest problems in the world are the result of lack of communication between peoples. Some refer to it as a "generation gap," but it is primarily a "language gap"... the result of the versatility of the English language. Some words have come to denote one meaning to one group and something entirely different to another. Each group has its own special "in" words which outsiders cannot understand. Often this leads to conflict of ideas and misunderstanding. This is becoming an important problem between scientists and society in general. As Sir Charles P. Snow, the noted British novelist and scientist, describes it in his book, *The Two Cultures:*

> "Greenwich Village talks precisely the same language as Chelsea, and both have about as much communication with M.I.T. as though the scientists spoke nothing but Tibetan."

Science, like any specialized branch of knowledge, has its own vocabulary, with special meanings for various words. In order to become one of the "in crowd," one must learn a little about the "language." The change in the meaning of words, like science, is an evolutionary process and will continue to be so. In this chapter, we will not be concerned with semantics, however, but with the nature of sound itself and how it is used for the benefit of man.

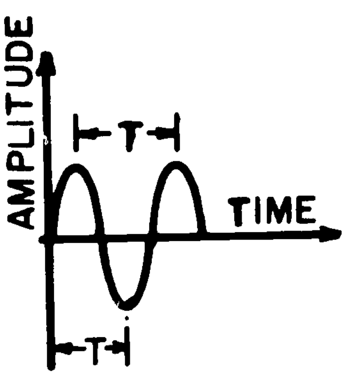
A baby learns to produce sounds immediately upon birth, and his future development and use of it can and does affect his remaining life. Through sounds, he soon learns that he can obtain food and other physical comforts. Later, he learns that he can communicate his wishes to others through the reproduction of certain definite sounds, or words, which are verbal symbols of objects or concepts. Without this development, it is doubtful that man could have ever progressed above the other primates and secured his dominant place in nature. While it is true that an understanding of sound is not necessary in order to utilize it, a knowledge of the nature of sound will deepen our appreciation of it. Consider how drab the world would be without musical sounds and poetry!

THE NATURE OF SOUND Sounds can be produced in many different ways, but each method has one thing in common--it requires the vibration of molecules in solids, liquids, and gases. Consider for a moment the sounds produced by musical instruments; they are the result of vibrating strings, reeds, drumheads, air columns, and other metallic instruments. The key word here is *vibrating*. To illustrate how sounds are produced by vibrations, consider a line of cars, bumper to bumper, on the freeway. Someone stops suddenly and a chain-reaction collision occurs along the string of cars. We refer to such a distur-



bance that propagates along the line of cars as a "pulse." If this collision is repeated numerous times in succession, we refer to the resulting disturbances as a "train of pulses" or "waves." Sound waves are produced in a similar manner by molecules of air colliding with each other and transferring energy.

Consider for a moment a plucked string on a musical instrument such as a guitar. The string vibrates back and forth, pushing the air molecules together in front of it, like the cars in the illustration above. Successive compressions of molecules give rise to sound waves.

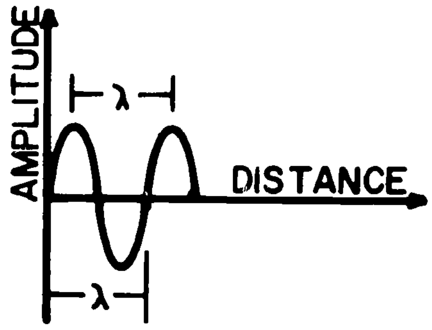


The number of times this sequence is repeated per second determines the type of sound or musical note produced. This is usually referred to as the "frequency" of vibration. For example, a musical note of middle "C" corresponds to a frequency of 256 vibrations per second. The *time* required to complete one vibration is called the "period" of vibration. The period for the musical note, middle "C", is 1/256 second.

The sounds that the human ear is capable of hearing generally lie in a frequency range between 20 vibrations per second and 20,000 vibrations per second. Sounds above 20,000 vibrations per second are referred to as *ultrasonic*. They have uses in fields such as sonar, dentistry, and sterilization of medical instruments and utensils because they produce violent vibrations. Some animals, such as dogs, are capable of hearing such high frequencies. Bats use ultrasonics to locate objects much the same way airplanes use radar and submarines use sonar.

We hear sounds when the molecules of air vibrate against our eardrums and cause them to vibrate at the same frequency. The vibrations produced in this membrane are converted into electrical pulses which travel to the brain via audiotory nerves and there they are interpreted as particular sounds.

SPEED OF SOUND The speed with which sounds are transmitted is dependent upon several factors...the density and temperature of the air, and the type of medium, being the most important ones. The distance that a wave moves before it is reproduced is called a *wavelength*. We will denote this distance in the following discussion by the Greek letter lambda, λ. You may recall that the speed of an object is the distance it travels, divided by the time required for the object to move this distance, or symbolically: $v = d/t$.

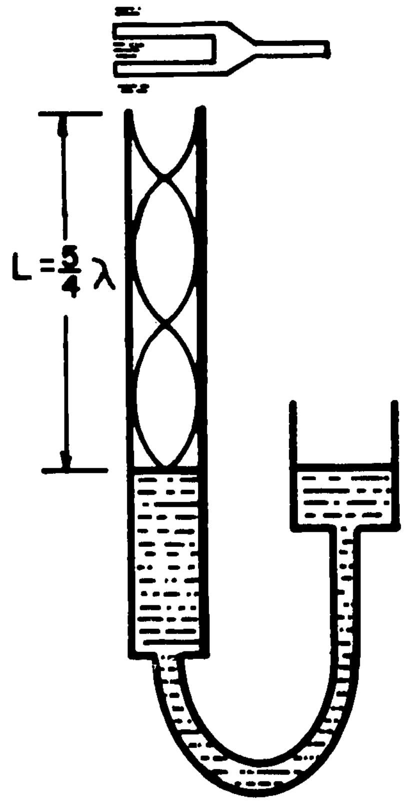


For a wave, $v = \lambda/t$, where t is the period of vibration. But t is related to the frequency, f, which is the number of vibrations per second. Therefore, $v = \lambda f$.

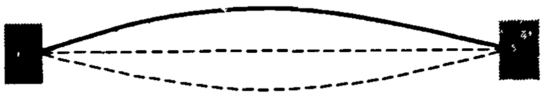
What this equation means is that one can determine the speed of sound waves if the frequency and wavelength are known. It is an experimental fact that sound waves of different frequencies travel with the same speed in air. Consider for a moment a concert orchestra playing a symphony. All of the notes that are produced at the same instant arrive at your ears at the same time, regardless of their frequencies.

One can measure the speed of sound in air in a very simply way. Consider a long tube containing water whose level can be raised or lowered. For a given frequency of sound, only a certain number of wavelengths will fit inside the tube or air column, whose lower level is the water surface. As a sound source we will employ a tuning fork designed to vibrate at one particular frequency. The vibration of the fork sets up waves in the air column which travel down the column with the speed of sound in air. When the waves meet the water surface, they are reflected back up the tube. If the length of the air column is such that $1/4\lambda$, $3/4\lambda$, or $n/4\lambda$, (n=odd integer) of a wavelength fits inside the tube, then a special condition arises. The reflected wave arrives at the tuning fork as a new wave is starting out and an increase in loudness is observed. This phenomenon is called *resonance*. One locates these resonance positions, and then knowing the frequency of the fork, one can use $v = \lambda f$ to calculate the speed of sound.

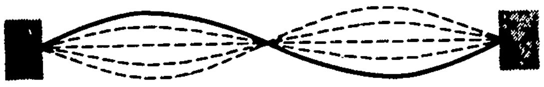
TUNING FORK



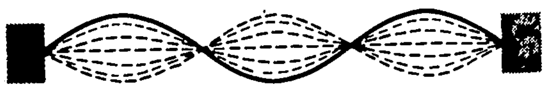
RESONANCE PHENOMENA The phenomenon of resonance accounts for the quality of sound produced by a particular musical instrument and gives it its own characteristic



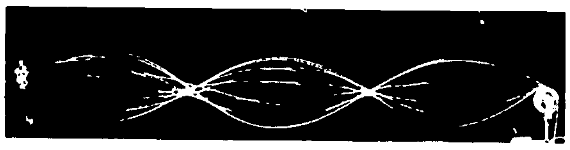
A fundamental frequency.



First overtone.



Second overtone.



A strobe photograph of a string vibrating in its second overtone.

sound or timbre. This is why a bass viol sounds different from a violin, and why peoples' voices are different. The sinuses and air passages of the head act as resonators to produce tones that are multiples of the fundamental frequency produced by the vocal chords. These *overtones* add the rich coloration to voices and musical instruments that help to distinguish them from one another. Musicians and scientists have tried unsuccessfully to reproduce violins of the same quality as those of Stradivarius. The factors that distinguish his violins from others is similar to the factors that distinguish one human voice from another.

Scientists believe that voices are unique as fingerprints. In fact, some courts have already admitted as evidence "voiceprints" of individuals as proof of their involvement in crimes. In the future, voiceprints may replace fingerprints as a means of identification. Someday you may have a car or home whose doors will open only upon your verbal command because its locks are controlled by a small computer coded to respond to your voiceprints and no other. You may be able to order products and services by merely telephoning your order to a central computerized distributer. The computer will analyze your voiceprint, bill your account, and ship the products to your home.

SUGGESTED OUTSIDE READINGS:

A.H. Benade , *Horns, Strings, and Harmony*, Doubleday (Anchor), Garden City, New York, 1960.

W.A. von Bergeijk, et al; *Waves and the Ear*, Doubleday (Anchor), Garden City, New York, 1960.

D.R. Griffin, *Echoes of Bats and Men*, Doubleday (Anchor), Garden City, New York, 1959.

W.E. Kock, *Sound Waves and Light Waves: The Fundamentals of Wave Motion*, Doubleday (Anchor), Garden City, New York, 1965.

J.R. Pierce, *Waves and Messages*, Doubleday (Anchor) Garden City, New York, 1960.

Eric M. Rogers, *Physics for the Inquiring Mind*, Princeton University Press, Princeton, New Jersey, 1960.

C.P. Snow, *The Two Cultures: And A Second Look*, The New American Library, New York, 1964.

S.S. Winter, *The Physical Sciences: An Introduction*, Harper and Row, New York, 1967.

ARTICLES FROM *Scientific American:*

L.L. Beranck, "Noise" (December 1966).

E.D. Blackham, "The Physics of the Piano" (December 1965).

E.E. Helm, "The Vibrating String of the Pythagoreans" (December 1967).

STUDY QUESTIONS:

1. How is sound produced?

2. What is the frequency of a sound wave? Its period?

3. What frequencies of sound is the average person capable of hearing?

4. Upon what factors does the speed of sound depend?

5. What is resonance? Give some examples.

6. Write a brief essay on the importance of sound in your daily life.

7. How is sound a benefit to man? What are some of its dangers?

CHAPTER 6 LIGHT - MESSENGER OF THE UNIVERSE

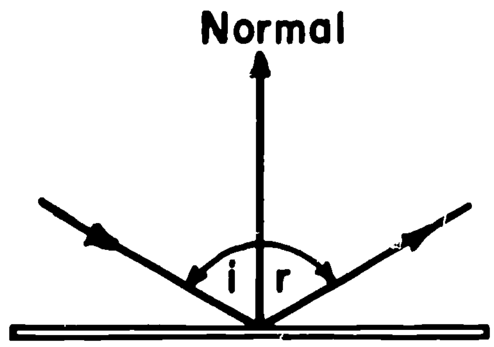
INTRODUCTION There is an old Hindu story about three blind men being asked to describe an elephant. The first felt the elephant's tail and said, "It's like a rope." The second felt a leg of the elephant and replied, "It's like a tree trunk." Finally, the last blind man felt one of the elephant's ears and said, "It's like a fan." Each man could only get a partial impression of the whole elephant from his sense of touch and each, therefore, came to different conclusions. A sighted person watching the scene would probably be amused by the three blind men's answers because he could see the elephant in its entirety and realize the absurdity of their answers. Most of us probably take our sense of vision for granted and rarely stop to think how different our world would be without it, or without the light which enables us to see. Without light, vision as we know it would be impossible. Indeed, life itself is dependent upon the light that the earth receives from the sun. Light is the chief catalyst in the biological process of photosynthesis.

Since light is so important to our very existence, perhaps we should strive to learn more about the nature of light. The questions concerning the propagation of light, how it interacts with matter, and how it may be used to benefit man will be the subject of this chapter.

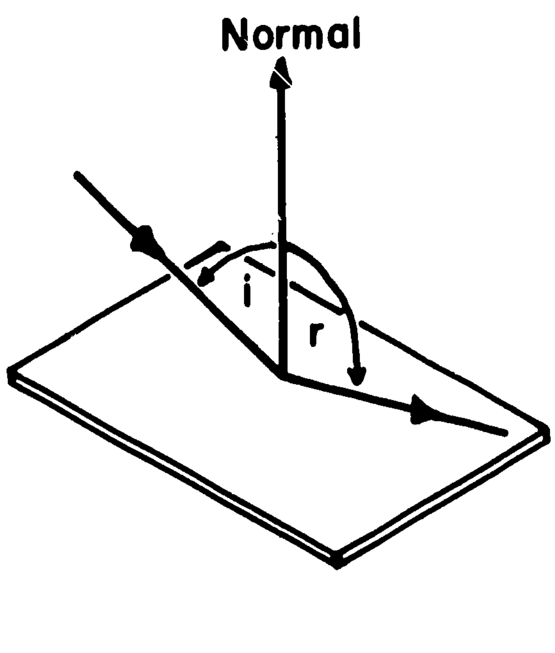
SOURCES OF LIGHT Some of the answers to the last question may be obvious to those of us living in an age of television, motion pictures, and electric lighting. Some of the answers to the other questions will require deeper investigation. We shall begin our inquiry by discussing sources of light. Early man's first light sources were the earth's sun, the stars, and the earth's moon. Later, these were supplemented by the discovery of fire. These sources served man from prehistory until the latter part of the nineteenth century when a new source of light was invented, the electric arc light. These first crude lights probably did as much to advance man's way of life as the discovery of fire. He no longer had to gear his working day to the sun, but could devote part of the night to study, recreation, and entertainment. Today we rarely realize our dependence on electricity until a power failure like the one that struck New York in 1965 comes along. How did we come to the point we are now in the understanding and utilization of light? We will seek the answers to this and other questions in the remaining sections.

THE NATURE OF LIGHT Some ancient Greeks thought that light consisted of particles that emanated from objects and entered the eye enabling the viewer to see. Others thought that the eye itself sent out "feelers" or tentacles which enabled the viewer to perceive the object it contacted. Little progress was made in understanding the biological and physiological processes of vision until man began to understand more about the nature of light through its behavior with matter.

REFLECTION AND REFRACTION The great Greek geometer Euclid performed some simple experiments with light and discovered several fundamental principles governing its behavior with matter. These principles are the laws of reflection of light from plane mirror surfaces, and the refraction of light by transparent materials such as glass. After man began to refine and work with metals he discovered that he could produce a mirror by polishing flat pieces of metal to a high luster. The images produced by these mirrors fascinated him and he attempted to learn more about their nature. Perhaps the reader has been amused sometimes in the past by a young animal, perhaps a parakeet or kitten, playfully toying with its own image in a mirror, not realizing that it is seeing itself reflected in the mirror. Early man probably was just as fascinated and intrigued as the kitten.



In determining the laws of reflection of light by mirrors, Euclid made a basic assumption that light travels in straight lines before striking a mirror and after striking a mirror. There are numerous examples in nature that suggest this is true: opaque objects cast sharp shadows; light traveling through a small hole forms a shaft of light; and one cannot see around a corner. With this assumption and some simple equipment it is easy to verify the law of reflection for a plane mirror: *The angle that light makes with the mirror surface when it is incident upon the surface is equal to the angle it makes when it is reflected from the surface.* Stated in terms of symbols: $\angle i = \angle r$. We shall say more about the *reason* for this result when we know more about the nature of light itself.



Using his geometry, Euclid was able to show that the image appeared the same distance behind the mirror as the object was in front. He also showed that the normal to the mirror and the incident and reflected rays all lay in the same plane.

Euclid also performed experiments concerning the refraction of light. Refraction is the bending of a light beam as it travels from one transparent material into another, such as from air into glass, or vice-versa. His conclusions on refractions are somewhat vague. Later, the Greek astronomer, Ptolemy, also studied refraction of light and came to some definite conclusions. He found that the ratio of the incident angle (the angle that the light path makes with the glass surface as it enters the glass) to the refracted angle (the angle that the light path makes with the glass surface after it enters the glass) was almost a constant value for any angle of incidence. Further experimentation in the sixteenth century by Snellius gave us the correct law governing refraction. This law is known as Snell's Law and it states: *The ratio of the sine of the incident angle to the sine of the refracted angle is a constant.* Stated in symbols:

$$\frac{\sin \angle i}{\sin \angle r} = n,$$

where n is a constant for a particular transparent material. This constant, n, is called the index of refraction, and is a measure of the speed of light in a medium.

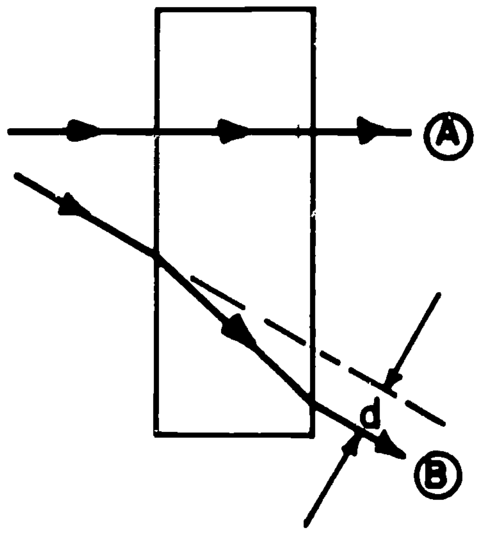


Fig. 6.1 Light entering a plane parallel plate at some angle is deviated a distance d from its original path.

Refraction occurs frequently in nature. Mirages which occur over hot deserts or over oceans are due to refraction of light by different densities of air. The same effect causes one to see "pools of water" on highways on warm days. One may actually view the sun after it is already below the horizon because light is refracted by the atmosphere. Refraction also causes stars to twinkle or appear to change positions.

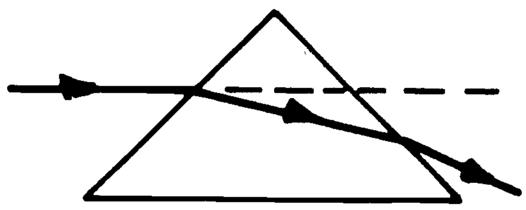


Fig. 6.2

The fact that the path of light is bent or refracted as it travels from one transparent material into another leads to several important consequences. Consider a rectangular piece of glass as shown in Fig. 6.1. Light striking the surface perpendicularly goes through undeviated along path A. Light striking the surface at an angle other than 90° is refracted upon entering the glass and refracted upon leaving the glass along path B. Its path after leaving the glass is displaced some distance, d. By using a triangular piece of glass, the deviation can be increased as shown in Fig. 6.2.

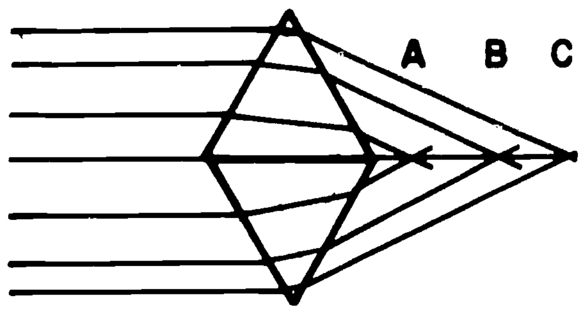


Fig. 6.3 (a)

Suppose we place two triangular pieces of glass with their bases together as in Fig. 6.3. The paths of light intersect at points A,B and C behind the glass. If we place the pieces of glass so that their vectices are in contact as in Fig. 6.4, the paths of light separate after leaving the glass. Suppose we now grind the plane surfaces until they are curved like a portion of a sphere. We have now produced two different types of *lenses*. The lens in Fig. 6.5 is a *convex* lens which brings parallel paths of light together at a common point called the *focal point, f*. The lens in Fig. 6.6 is a *concave* lens which separates or diverges the paths of light. If we

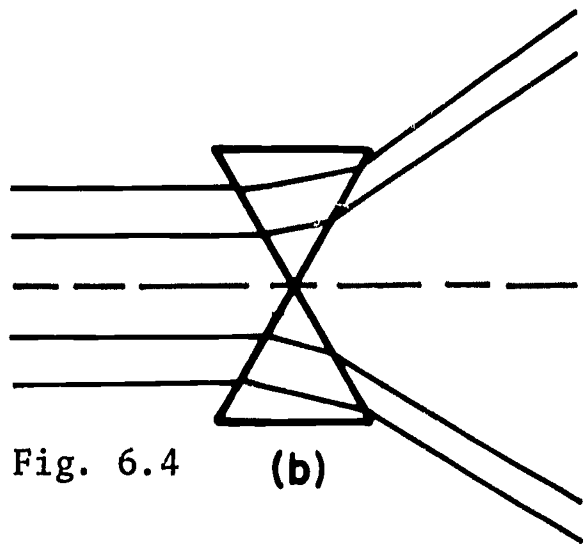


Fig. 6.4 (b)

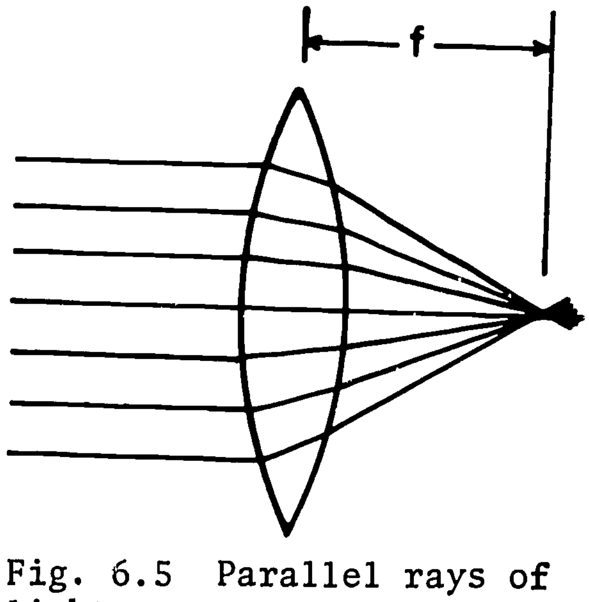


Fig. 6.5 Parallel rays of light converge to a common point, *f*, after emerging from the lens.

trace the paths through the lens, they appear to come from a common point, *f*, the focal point of this lens.

The convex lens is capable of producing what is referred to as a *real image*. Such an image can be projected upon a screen. You may have used a magnifying glass (which is a convex lens) as a child to ignite bits of paper or cotton by focusing the light from the sun on the paper. The bright spot that the lens forms on the paper is a real image of the sun. The concave lens cannot produce such an image. Instead, it produces an image like the image produced by a plane mirror. Such an image is called a "virtual image" because it appears behind the mirror or lens and cannot be projected.

Men learned that such devices as convex and concave lenses could help correct certain visual defects such as far-sightedness and near-sightedness. The Chinese were using these lenses as eyeglasses as early as the 1300's. Eyeglasses were introduced into Europe around the 1400's after trade routes were opened up to the middle and far East.

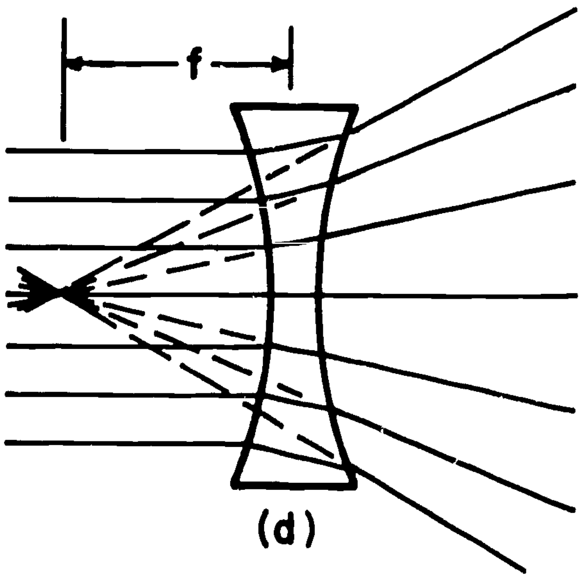


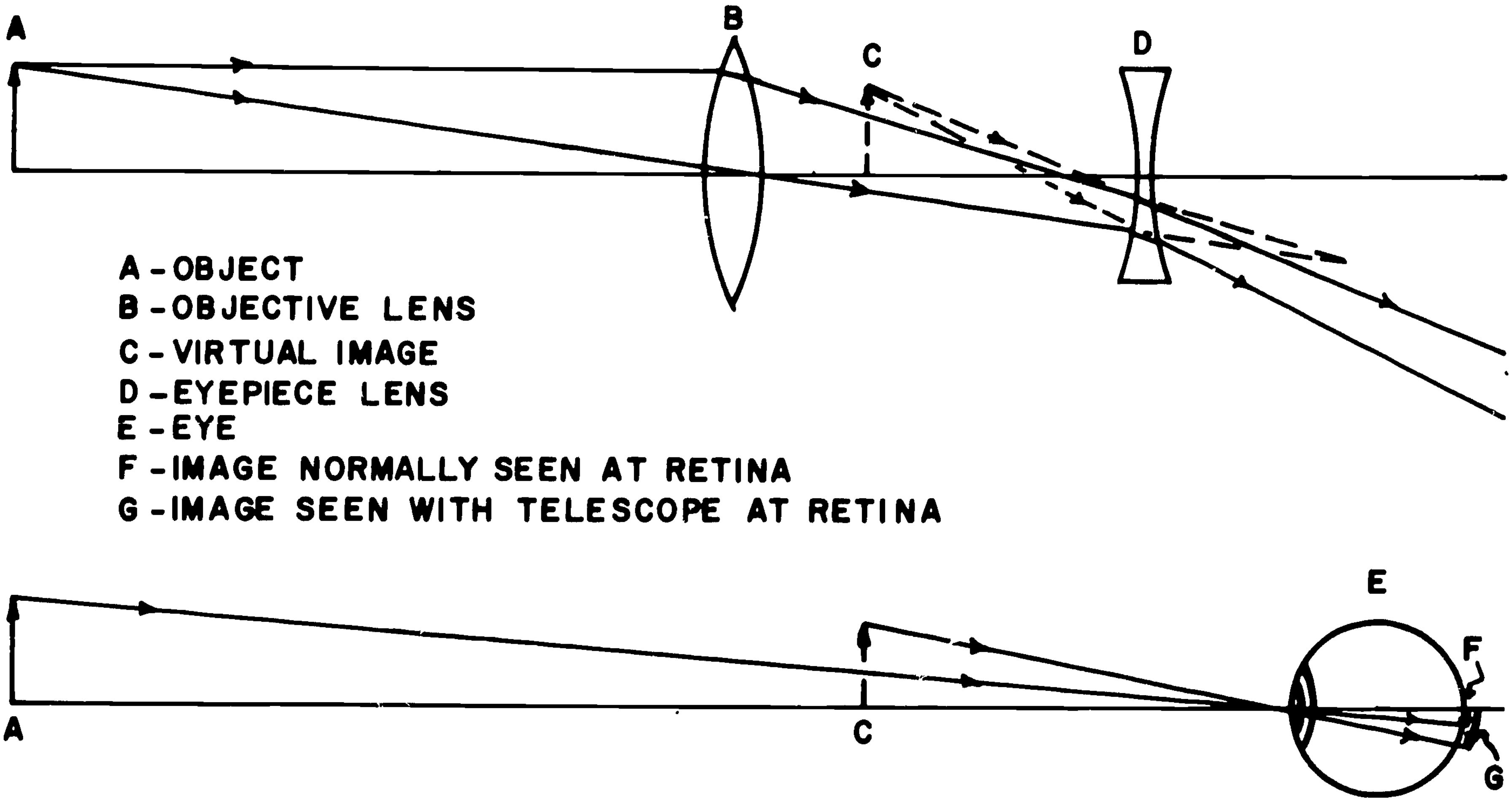
Fig. 6.6 Parallel rays of light appear to come from a common point, *f*, after emerging from the lens.

Almost two centuries passed before the next major development in optical instruments occurred. About 1608 a Dutch lensmaker named Hans Lippershey accidently discovered that two lens used in combination could be used to magnify distant objects and thereby produced the first telescope. About the same time another Dutchman named Joannides is credited with the invention of the microscope. You are already familiar with the story of what happened when Galileo turned his telescope skyward in 1609 and brought the science of telescopic astronomy into being. The telescope of Galileo was crude compared to the sophisticated optical telescope of today but the basic principles governing their operation are still the same. Galileo's first telescope utilized a convex lens and a concave lens and this combination is referred to as a Galilean telescope.

The convex lens is used as the objective lens, or the lens nearer the object to be viewed. It forms a real image of the object by focusing the light reflected from the object at approximately the lens' focal point. By placing a concave lens in front of the position where the real image is formed, one can cause the light path to diverge. By looking into the concave lens or eyepiece, one sees a virtual image of the object.

The image formed is in the same orientation as the object (erect) and magnified. A Galilean telescope is diagrammed in Fig. 6.7.

Fig. 6.7 Galilean telescope.



The reason that a telescope gives an enlarged image is that the image formed on the retina of the eye is larger. The image produced by the object without the telescope is smaller than the image produced by the telescope.

Galilean telescopes were first used as terrestrial telescopes because they produced erect or upright images. A popular version of the Galilean telescope is the compact opera glass of today.

A second type of telescope is produced if the concave eyepiece of the Galilean telescope is replaced by a convex eyepiece. Such a telescope is known as an *astronomical* telescope because it is best suited for this purpose as it *inverts* the final image. The astronomical telescope is constructed similarly to the Galilean telescope, except that the convex eyepiece is placed *behind* the real image formed by the objective lens. A convex lens is capable of producing a virtual image as well as a real image if certain conditions are fulfilled. In order to produce a virtual image, one merely places the object to be viewed *within* the focal length of the lens. This is what one does when the convex lens is used as a magnifying glass, or reading glass. The print to be magnified is placed within the focal length of the lens, and a magnified virtual image is produced in the same orientation as the object.

If the inverted real image produced by the objective lens is allowed to fall within the focal length of the convex eyepiece lens, then the eyepiece acts as a magnifying glass and gives an inverted, virtual, enlarged image of the real image produced by the objective lens. A typical astronomical telescope is diagrammed in Fig. 6.8. Since the final image is inverted, this telescope is of little use for terrestrial viewing.

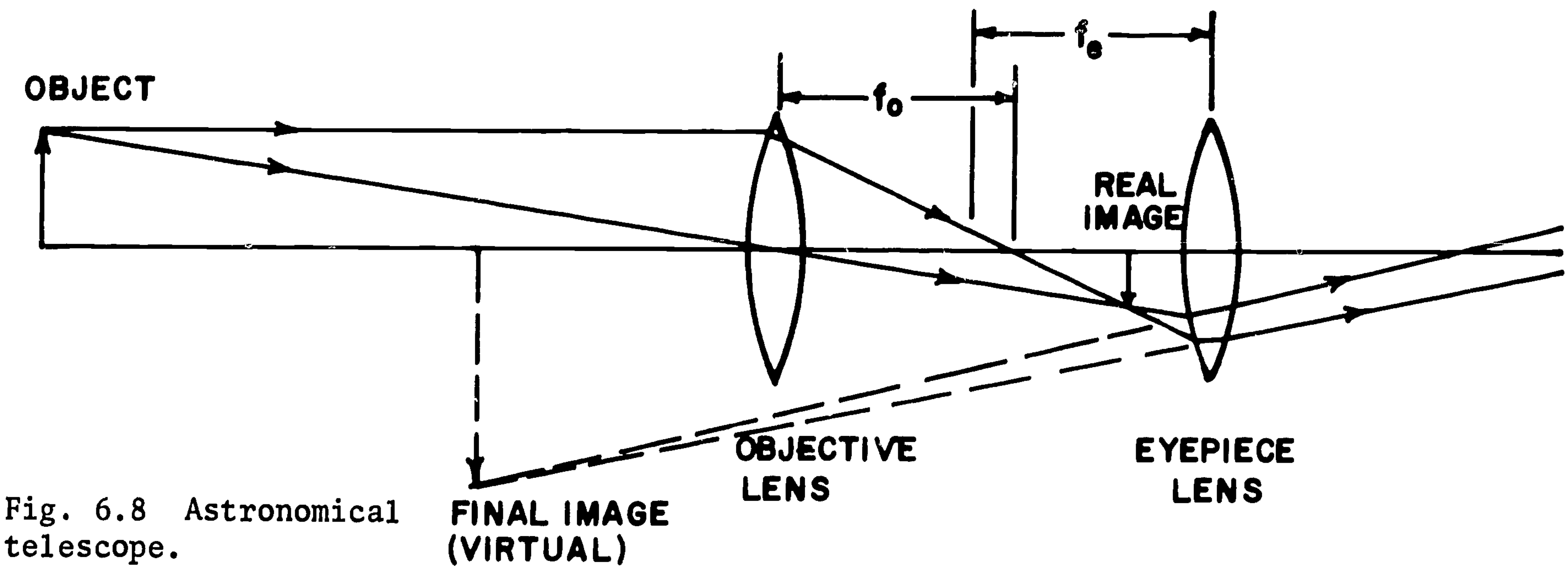


Fig. 6.8 Astronomical telescope.

A third type of telescope employs a single convex lens for an eyepiece and a mirror for the objective lens. This type of telescope is called a *reflector*. Sir Isaac Newton is credited with its invention. Its operation is basically the same as the astronomical telescope, except that a parabolic mirror is used to focus the light to a common point, forming a real image, and the convex eyepiece lens is used to magnify the real image. Consider again the basic law of reflection for a plane mirror and apply it to a concave mirror as shown in Fig. 6.9.

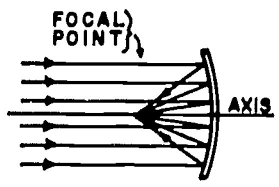


Fig. 6.9

One can verify that such a mirror also is governed by the law of reflection, if tangents are constructed at the surface where each ray of light strikes the surface. The reflection of light intersects at a common point, the focal point for the mirror, and forms a real image there. By placing such a mirror of long focal length at one end of a tube, one can reflect light back up the axis of the tube. One can reflect the light from the mirror so that it strikes the tube perpendicularly by inserting a small plane mirror or prism so that it makes an angle of 45° with the axis of the tube. See Fig. 6.10. By placing this mirror a short distance in front of the position where the real image is formed by the concave mirror, one can then install a convex lens eyepiece in the side of the tube to magnify the real image. This type of eyepiece mounting is known as a Newtonian mounting. Various

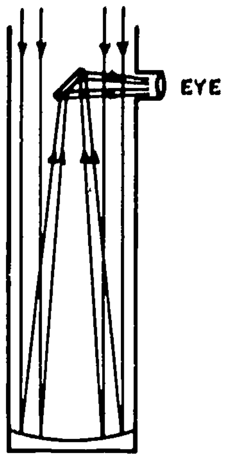


Fig. 6.10 Reflecting telescope (Newtonian mounting).

other arrangements are also possible. If the small plane mirror is placed with its surface perpendicular to the axis of the tube, the light from the concave mirror is reflected back upon itself. By using a concave mirror of longer focal length, with a hole cut in its center, one can reflect the real image to some point *behind* the objective mirror. The image is viewed by installing an eyepiece behind the mirror. This type of mounting is called a Cassegrainian mounting and is illustrated in Fig. 6.11.

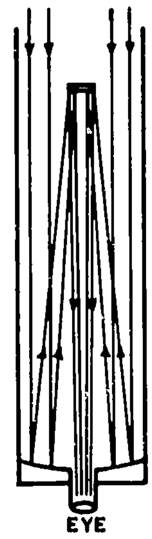


Fig. 6.11 Reflecting telescope (Cassegainian mounting).

This mounting has the advantage that the observer is at the bottom of the telescope when it is aimed skyward rather than perched precariously near its top on a ladder or stand.

The larger astronomical telescopes in use today are of the reflecting type as it is easier to produce large mirrors to close tolerances than it is to cast and grind lenses of large diameter that are free of air bubbles and other distortions. The largest such reflecting telescope is the 200 inch diameter reflector located on Mount Palomar. It is so huge that an observer can sit inside it without obscuring the light. By contrast, the largest refracting telescope is the 40 inch diameter telescope located at Yerkes Observatory. The range of the 200 inch reflector is approximately 10 billion light-years.* The light that this telescope collects started its journey through space billions of years before the solar system even existed! To look into the heavens with these instruments is to see not only stars and planets, but to look back into the dawn of time and creation itself!

LIGHT - PARTICLE OR WAVE? So far we have seen that light travels in straight lines, is reflected, and refracted by matter. We have assumed for these phenomena that light behaves as a material particle. Numerous examples exist to indicate that light behaves like a particle, or a stream of particles - for example, sharp shadows are cast by opaque objects when light shines upon them, indicating that some of the particles are absorbed by the opaque object, much like an object in the rain casts a "shadow" or dry spot when it absorbs drops of rain.

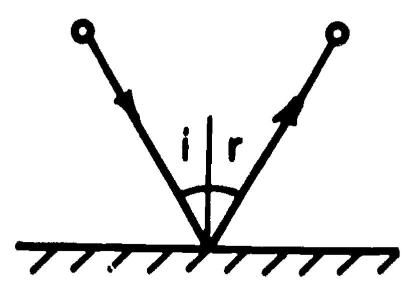
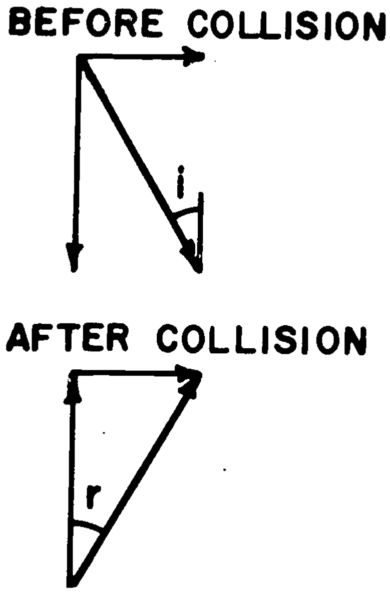


Fig. 6.12

Reflection of light from a mirror surface can be likened to a rubber ball bouncing against a smooth sur-

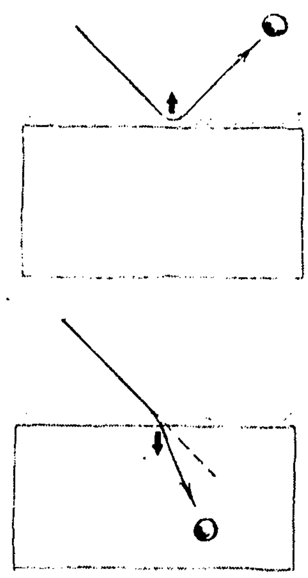
* One light year is the *distance* light travels in one year, traveling at a velocity of 186,272 miles per second, or approximately 5.880×10^{12} miles.



The downward component of velocity is reversed by the reaction force. The component directed to the right is unchanged.

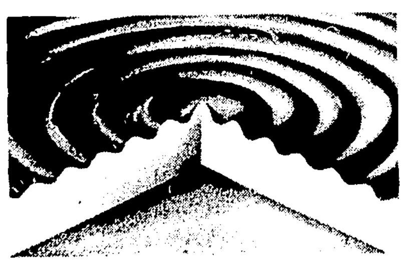
face at some angle. See Fig. 6.12. The downward component of the ball's speed is reversed by the reaction force exerted on the ball by the surface. Since no force acts parallel to the surface, this component of speed cannot change. Therefore, the angles must be equal.

How can refraction of light be explained by the particle theory of light? Newton advanced one such explanation after his discovery of the law of Universal Gravitation. His explanation was that the particles of matter in the denser transparent material exerted a stronger force or pull on the particles of light than did the particles in the less dense material. Hence, the particles of light would travel faster in the denser material and be bent or refracted from their original path.



Newtonian explanation of reflection and refraction.

The Newtonian theory of light was accepted by most scientists until about 1665 when an Italian scientist named Grimaldi made a discovery that cast serious doubt on the particle theory. What Grimaldi discovered was that light sometimes *does not* travel in straight lines! If we place a source of light so that its light falls upon a small circular hole in front of a screen, we find that the hole casts a circular bright spot on the screen. If we examine the region where the geometric shadow begins, we observe a sharp cutoff of the light. If, however, we replace the circular hole with successively smaller diameter holes, we soon discover that a certain point is reached where the bright spot *extends* beyond the geometric shadow, or into the dark region where no light could possibly reach without penetrating the material around the hole. Grimaldi called this phenomenon *diffraction*. Apparently, light is bent whenever it passes a very small aperture. The particle theory fails to adequately account for this phenomenon. Are there other phenomena in nature that behave in a similar manner to diffraction of light and could they be used to help explain this phenomenon? The answer is yes. Water *waves* are also diffracted by obstacles such as rows of pilings or breakwaters. If we pass a straight water wave through a small slit we discover that it emerges as a semicircular wave front, which extends into the geometric shadow of the slit. The Dutch scientist Christian Huygens, a contemporary of Newton, recognized similarities between the behavior of light and water waves, and developed a wave model theory of light, with which he challenged the Newtonian particle theory. Because of the stature of Newton in the scientific community, Huygens' theory was not well received until the beginning of the nineteenth century.



Cross-section of a water wave.

In 1801 Thomas Young discovered another phenomenon which supported the wave theory and cast further doubt on the particle theory. This phenomenon was *interference* of light. Young discovered that light passing through two closely spaced slits produced a pattern similar to the diffraction pattern produced by a single slit or hole. However, the pattern was not produced by the bending of light, but by interference between the two diffracted waves. See Fig. 6.13.

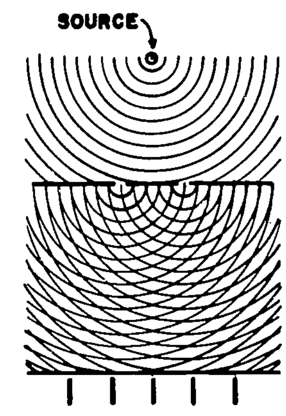


Fig. 6.13 Diagram of Young's diffraction experiment.

Consider two water waves produced by two stones dropped into a pond some distance apart. The resulting circular wave patterns spread out until they begin to overlap. Consider a cork floating on the surface of the pond some distance from the source of the two waves. As one wave crest reaches the cork, it will rise to a certain height as the crest passes and sink as the trough passes. As both waves produce crests at the cork simultaneously, the cork will rise to a height greater than that of a single wave crest, and will sink deeper than a single wave trough. If, however, a crest from one source and the trough from the other source arrive simultaneously at the cork, the cork will remain motionless, provided that the waves are the same size. This cancellation or reinforcement of waves is called interference. Two light waves behave in a similar manner. When two crests or troughs coincide, reinforcement occurs and bright spots or areas appear on a screen placed behind the two slits, and when a crest and trough coincide, cancellation occurs and dark areas appear on the screen. See Fig. 6.14.

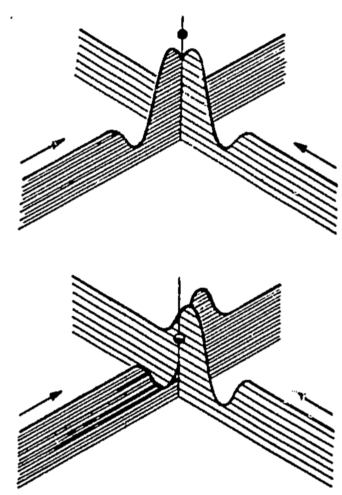
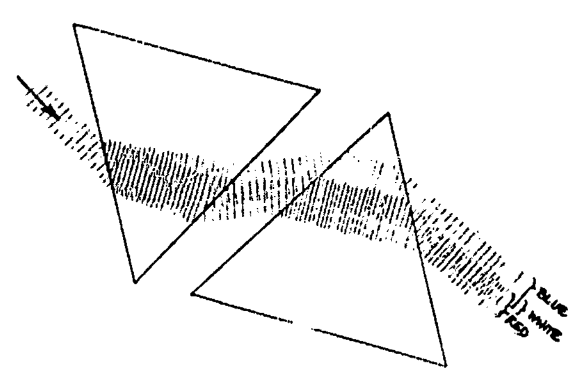


Fig. 6.14 Interference of two water waves. Cork floats higher when two crests overlap. It floats lower when crest and trough overlap.

Other examples of interference phenomena occur when light is reflected from the front and back surfaces of thin films such as soap bubbles or an oil film on a street after a light rain shower. Interference or waves reflected from the two surfaces produce colors. Where do these colors come from? Are they due to the reflecting materials, or are they present in light itself?

Newton was one of the first to attempt an answer to this question. Newton performed numerous experiments with triangular prisms to observe the spectrum of colors created when light passes through them. This phenomenon is called dispersion. Newton was concerned whether the colors were produced by the prism itself, or whether they were an inherent property of light which the prism revealed. Through several experiments, he was able to show that the colors separated by one prism could not be further separated by a second prism. Hence, he concluded that the light of the sun was a mixture of many colors. Newton reasoned that if a blue powder and a yellow powder mixed together appeared green to the eyes, but still could be distingushed as separate colors under a microscope, then,



The second prism cannot separate the colors further.

perhaps, white light could be composed of a mixture of the basic colors inherent in the spectrum produced by the prism. Newton thought that the prism separated the individual colored particles in white light by refraction, and that the same process accounted for the colors of the rainbow produced by rain droplets.

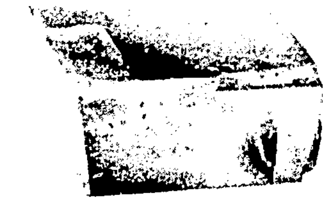
Newton was basically correct in assuming that white light was a mixture of colors; however, Young's experiment showed that only a wave theory could satisfactorily explain interference phenomena. Today we consider white light to be a mixture of different *wavelengths* of light instead of a mixture of different colored particles. Each wavelength travels with its own characteristic velocity through matter, and hence, transparent materials can separate each wavelength according to its velocity. A table of wavelengths for visible light appears in Table 6.1 below:

TABLE 6.1. SPECTRUM OF VISIBLE LIGHT

COLOR	Approximate Wavelength Range in Angstrom Units ($1\ \mathring{A} = 10^{-10}$ meter)
VIOLET	4000--4300
BLUE	4300--4900
GREEN	4900--5700
YELLOW	5700--5900
ORANGE	5900--6400
RED	6400--7000

Newton theorized that objects appear different colors due to the color of light that they reflect. Grass, for example, appears green because it absorbs all colors of light present in white light except the green color, which it reflected to the eye. Objects observed under illuminations other than white light appear different colors. If we shine red light upon a blue object, it would appear black because it absorbs the red light and reflects little.

POLARIZATION OF LIGHT Young's interference experiment established that light consists of waves; however, it did not establish the character of the waves--whether they are *longitudinal* like sound waves, or *transverse* like water waves or waves on a string. Another phenomenon, polarization, necessary to distingush between the two alternatives, was discovered by a French engineer named Malus in 1808. He accidently discovered this effect while looking through a crystal of Calcite, which possesses the property of double refraction (one ray of light entering the crystal is separated into two rays). He found that rotation of the crystal resulted in changes in the intensity of the two images of light source viewed through the crystal. As one image decreased in intensity, the other increased. This experiment indicated that light is transverse since rotation would not change the intensity of a longitudinal wave. Many other minerals also exihibit this phenomenon. Crystals of iodine compounds also act as polarizers if aligned in a preferred direction.



Iceland Spar Crystal
Double Refraction

Ordinary light is polarized at random, so that the mixture of waves are all vibrating in different planes. Passage of the unpolarized light through a polarizing filter removes all components except those vibrating at right angles to the absorbing crystals. The emerging light is *plane polarized* or vibrates in one plane only. See Fig. 6.15.

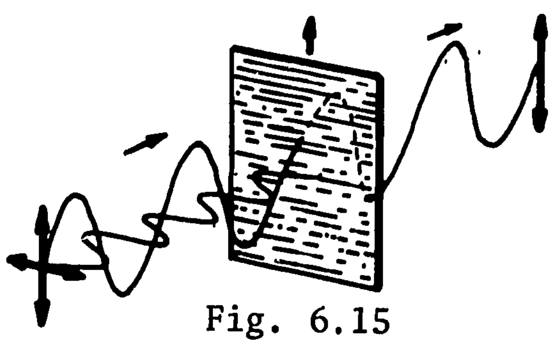


Fig. 6.15

Passage of this plane polarized light through a second polarizing filter results in maximum brightness if the two filters are in the same orientation, or if their planes of polarization are parallel. See Fig. 6.16 If the second polarizer is perpendicular to the first (see Fig. 6.17), it absorbs most of the plane polarized light coming through the first filter and very little light emerges from the second filter.

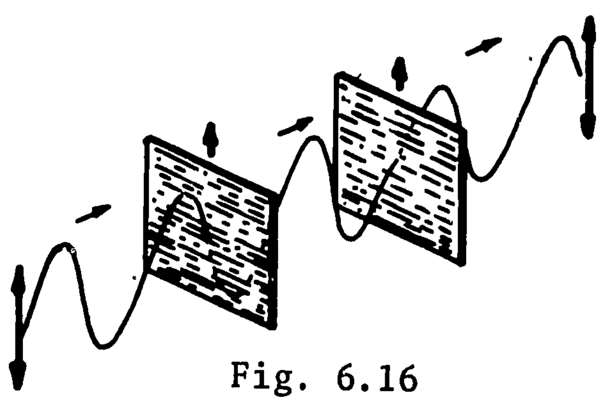


Fig. 6.16

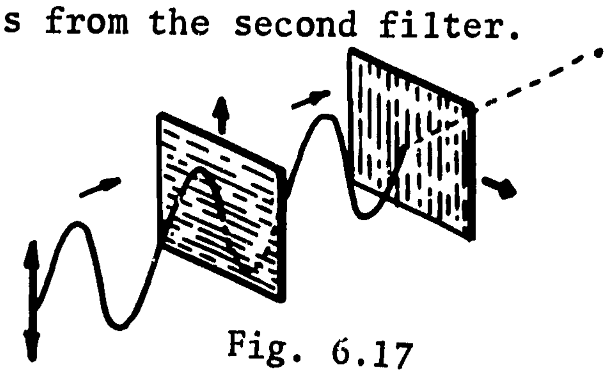


Fig. 6.17

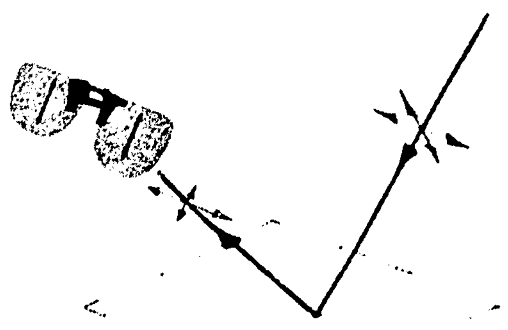


Fig. 6.18 Polariod sunglasses remove the component of light partially plane polarized by reflection.

In 1928, E.H. Land, developer of the Polaroid-Land camera, developed a polarizing material that he called "Polaroid" which is widely used today in the manufacture of sunglasses. The "Polaroid" sunglasses are superior to ordinary sunglasses that are only tinted to decrease the light intensity. Ordinary unpolarized light is partially plane polarized by reflection, and hence, "Polaroid" sunglasses reduce more of the glare. See Fig. 6.18.

Polarizers also have an important use in engineering and industry in the analysis of stress in structures. The branch of engineering which deals with the application of polarized light to these problems is called *photoelasticity*. The polarized light is used to show when defects may occur in transparent materials. An engineer may wish to find out how a bridge will support a given load of traffic and where reinforcement is needed in the structure. By building a scaled model of the bridge of transparent plastic, and shining polarized light through it, he can detect the stress area by analyzing the emerging light with another polarizer. See Fig. 6.19.

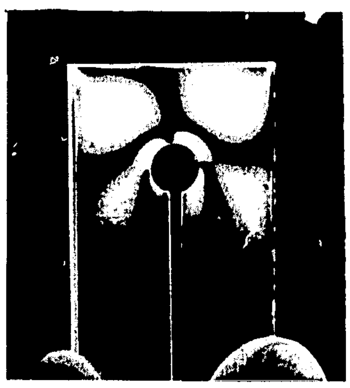


Fig. 6.19 Photograph of stress pattern in a plastic sample.

The areas of stress show up as light or dark areas in the plastic. This is due to the realignment of the molecules of the model when force is applied to these areas. In this manner, defects in the transparent materials can be analyzed prior to the construction of the full-scale object, and savings may occur as unnecessary material can be eliminated and other areas can be reinforced where needed. Using this technique, defects in glassware can easily be spotted and poor products eliminated before they reach the consumer.

THREE-DIMENSIONAL MOVIES As a child you may recall viewing a three-dimensional motion picture. This is another application of polarized light. Some companies still manufacture 3-D slide projectors for viewing stereoscopic slides on a screen. Normal stereoscopic vision is possible because each eye sees a slightly different view of the same object. The brain can fuse these two images together into a three-dimensional picture. In stereoscopic photography, one uses a camera with two lenses spaced about the same distance apart as the human eyes. Exposure of the film results in two pictures of the same object taken from slightly different angles. The developed pictures are mounted side by side for viewing with a stereoscopic viewer or for projection.

Using a special two-lens projector which projects the two pictures on the same screen, one gets two overlapping pictures which appear blurred to the normal eye. If two polarized filters are placed over the lenses so that each is at right angles to the other, one obtains pictures that are plane polarized at right angles to each other. If one views the pictures using a pair of glasses containing "Polaroid" lenses mounted at right angles to each other, each eye will see only one of the pictures and the brain will interpret this as a three-dimensional view.

LASERS To those familiar with Ian Flemmings' super spy, James Bond (agent 007), the word *laser* means a deadly ray of light capable of great harm, but to scientists, it represents an exciting, new form of light whose benefits are just beginning to be understood and developed.

First of all, the laser is not a death ray, but a unique type of light--one whose wavelengths are all *in phase* (or step). Such a light source is said to be *coherent*. In this light source all the atoms which emit light do so together. In fact, the word *laser* is a contraction of the term *L*ight *A*mplification by *S*timulated *E*mission of *R*adiation.

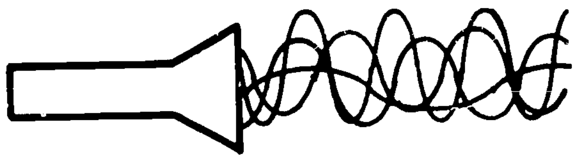


Fig. 6.20 Incoherent light source.

In an ordinary light source, atoms emit light radiation at random. Hence, all the wavelengths of light are all out of phase and of different wavelengths. See Fig. 6.20. In the laser, atoms are stimulated or forced to emit light together; hence, the light is of a single wavelength, and coherent. See Fig. 6.21. It is this unique property of the laser that makes it so useful.

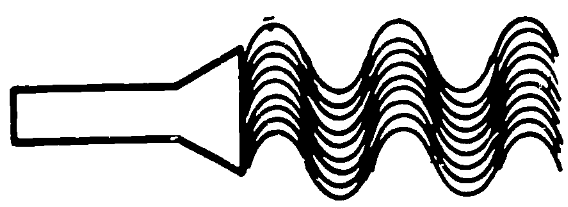


Fig. 6.21 Coherent light source.

Besides being a source of coherent light, which scientists can utilize in optical research, the laser is proving to be a useful tool in medicine. One use is in the reattachment of a detached retina of the eye. Before the laser, blindness usually resulted from this condition. The coherent light of the laser can be focused by the eye's lens and concentrated into a small, intense beam. Hence, it is possible to "weld" detached retinas back into place.

Future applications may be surgical knives of light that seal blood vessels as incisions are made, preventing loss of blood.

Applications in industry include micro-welding of micro-circuits in electronics, use as a surveying transit, use as a fine drill, and applications in the field of communications. One laser beam may someday carry thousands of messages in special pipes like telephone wires do today.

HOLOGRAPHY Another application of the laser is in the field of photography, or holography as laser photography is known. An object to be photographed is exposed to laser light and the scattered light is allowed to fall on a photographic plate. The plate is also illuminated by a direct laser beam. Since both beams of light are coherent, interference occurs between them in the film. Instead of getting a photograph of the object, an interference pattern is produced. The developed film appears much like a photograph of a lake surface which has been disturbed by several stones striking the surface together.

When the film is illuminated by laser light or a monochromatic (single wavelength) light source, a three-dimensional image of the original object is reconstructed.

Each part of the film is capable of reproducing the original scene in its entirety. By viewing the hologram at different angles, one can see the object in some other orientation or perspective. Someday you may be able to receive three-dimensional television programming if such a process using laser light proves possible and practical.

SUGGESTED OUTSIDE READINGS:

F. Hoyle, *Astronomy*, Doubleday, Garden City, New York, 1962.

B. Jaffer, *Michelson and the Speed of Light*, Doubleday, (Anchor), Garden City, New York, 1960.

W. E. Kock, *Sound Waves and Light Waves: The Fundamentals of Wave Motion*, Doubleday (Anchor), Garden City, New York, 1965.

R.T. Lagemann, *Physical Science: Origins and Principles*, Little, Brown and Company, Boston, 1963.

Project Physics, *An Introduction to Physics: Light and Electromagnetism*, Holt, Rhinehart and Winston, New York, 1968.

ARTICLES FROM *Scientific American:*

E.N. Leith and J. Upatnicks, "Photography by Laser" (June 1965).

S.E. Miller, "Communications by Laser" (January 1966).

R.S. Shankland, " The Michelson-Morley Experiment" (November 1964).

STUDY QUESTIONS:

1. What evidence exists that supports the idea that light travels in straight lines?

2. Give some common examples of refraction of light.

3. What is the difference between a Galilean telescope and an astronomical telescope?

4. How does a telescope produce an enlarged image?

5. What advantages does a reflecting telescope have over a refracting telescope?

6. Describe Newton's theory of light.

7. What evidence tends to support Newton's theory of light?

8. What evidence tends to refute Newton's theory of light?

9. What is meant by diffraction of light?

10. What is meant by interference of light?

11. Why do roses appear red in color?

12. What does polarization of light indicate about the nature of light?

13. List some applications for laser light.

CHAPTER 7 THE STORY OF ELECTRICITY AND MAGNETISM

PAGE

INTRODUCTION - FROM THE REALM OF THE GODS TO THE BENEFIT OF MAN The story of electricity and magnetism spans the entire scope of man's speculation about and investigation into the world in which he lives. Naturally occurring electrical phenomena were probably among the first effects of nature to intrude upon early man's efforts to find food and shelter for himself. For instance, his discovery of the effects of and his first source of fire very likely was due to the energy released by lightning striking a tree during a storm. These first acquaintances with electrical phenomena were often attributed to mystical manifestations of gods or demons and no further investigations were made into their nature and only hesitant uses were made of their application.

Systematic investigation and speculation into the nature of electricity and magnetism did not really begin until the work of William Gilbert (1540-1603) and has moved forward at an accelerating pace ever since. Major breakthroughs were made in the last half of the nineteenth century with the result that mankind has made greater strides in the past one hundred years in harnessing energy and developing time for other pursuits than in all of his previous time on the earth! Today, nearly every action or device that we use depends either directly or indirectly upon the harnessing and application of electrical energy in one form or another.

The story of electricity and magnetism is not merely a story to be used to catalog gadgets and how they work, however. It is also a story that can tell how scientific investigation developed, the nature and reasoning of the people who did the work, and some of the other influences that were at work in society at large when the developments were made. In some respect, this latter story is more important than the development of "gadgetry" in gaining a better insight into the field of science and the people who have developed it into the form that we have today.

We shall deal with both of these stories in our discussion here. Since the two related fields of electricity and magnetism were developed along similar, but separate lines until the nineteenth century, it will be easier to treat them as two separate stories until they were merged together in the work of Faraday, Oersted, Henry, Maxwell, and others. Occasional references will be made where possible from one story to the other to tie them together, but by separating them in this manner, it should be easier to ascertain what happened in each. Finally, we shall have a brief look at what causes some of the phenomena that we find in everyday life in our dealings with the manifestations of electricity and magnetism.

ANCIENT TABOOS. Philosophers, due to the nature of their efforts, are often given to speculating and reasoning over "unanswerable" questions. These have taken such forms as attempting to determine the sound of one hand clapping, the number of angels that can dance on the head of a pin, whether a tree falling in a forest makes a sound if there is no one there to hear it, and so on. Scientists base their theories on observation, logical reasoning, and experimentation. These theories should conform to all of the facts presently known and are examined to see if additional phenomena might be predicted or expected outside of what is presently known and, in this manner, find some answers that *might be correct* to some of the "unanswerable" questions of the philosophers.

Recent investigations of this sort have concerned the origin of the universe and the beginnings of life. All of these theories have been based upon the premise that the presently known physical laws, including those of electricity and magnetism, were in effect then in the same manner that they are today. These investigations have shown that - by combining water, chemicals, and radiation of the sort thought to be coming from the sun when life first began - proteins (the basic molecules of all life) can be produced by the catalytic action of an electrical discharge or bolt of lightning. Thus, we can confidently infer that natural electrical phenomena as we know them today, were present long before the coming of man.

What were some of the natural manifestations of electricity that prehistoric man found, and how did he probably react to them? The four most prominent effects that were apt to be encountered are lightning, static charges from rubbing one material against another, St. Elmo's Fire, and the electrical discharge from electric eels and other similarly endowed aquatic animals. Primitive people seem to have a universal predilection toward consigning any phenomena that are not readily explainable to the actions and providence of some gods or demons. Furthermore, the more vigorous and common the effect, the higher the ranking of the god in the local hierarchy that is supposed to be in charge of the phenomena. The result is that investigations into the origins of and reasons for the phenomena are blocked, as they were in the special providence of the gods and not something for mere man to know, understand, or even consider.

Thus, the ancient Greeks took the phenomena of lightning and put it under the special providence of Zeus, the head of all of their gods, and regarded it as a sign of his displeasure because of the destructive power resulting from the enormus amount of energy contained in a single bolt.

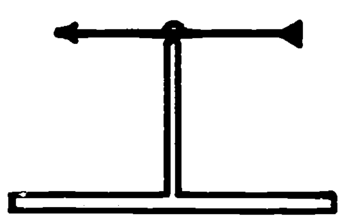
Around 70 A.D., the Roman author, Pliney, referred to two phenomena that are closely associated with large bodies of water and that were attributed to the sea gods. One of these concerned the torpedo, a form of electric eel. Biologists have not, as yet, determined the way in which this animal can produce large voltages between different parts of its body to defend itself and kill its food, but the effects have been known for at least 1900 years and probably much longer. Superstitious fishermen, encountering this animal and being shocked or even having one of their companions killed by it, would naturally fear it and believe that they had encountered some evil demon of the seas that should be avoided when possible.

The other phenomenon mentioned by Pliney and undoubtedly known much earlier - is St. Elmo's fire. It occurs during storms on or near the oceans and is a greenish-yellow glow that seems to flow from the top of high objects, such as the mast of a ship, along the object until it is all illuminated down to some flat surface such as the deck. Sailors have seemed to regard this phenomenon as a particularly bad omen for whatever venture they were engaged in down through the ages. Herman Melville in his novel, *Moby Dick*, had occasion tc use this phenomenon as a means of building suspense for the reader and disquiet among the crew of the "Pequod" before attacking the great white whale of the novel. His is an excellent description of the features of St. Elmo's fire and of the weather conditions in which it is found.

These and other phenomena related to other fields of science acted as a great deterrent to investigation and measurements of the natural manifestations surrounding early man. Instead, philosophical reasoning without experimentation dominated speculations into the nature or cause of these effects. This held even into those areas where the phenomena had not been attributed to a specific god or gods. One of these was the effect obtained by rubbing one material with another. The first of these materials mentioned in early writings was amber. After being rubbed it exhibited the ability to attract small bits of material - such as straw. Amber is a reddish-orange material with a waxy texture. Its Greek name, "elektron", gave us our present name for the smallest known unit of electrical charge, the electron, as well as a name that appears throughout the entire study of this phenomena.

The ancients' restriction to investigation of electrical phenomena did not, however, apply to discovering possible applications of the effects or - in the case of the generation of static electricity - to determining other materials that could be used to produce the same or similar effects. This resulted in wide-ranging efforts to use static electrical (as well as magnetic) effects in the fields of medicine, and alchemy to draw out such things as "bad humors" from people or gold from other materials. It is an easy matter to deride these efforts in the light of what is known today, but it must be remembered that when the experiments took place, there was no prior knowledge to guide tneir researches into the applications of the effects. Some of these efforts were indeed successful and these are the results that we tend to take for granted when or if we think of them today. This is especially true of some of the efforts in the field of magnetism which we shall discuss later.

MODERN BEGINNINGS William Gilbert (1540-1603), who was the court physican to Queen Elizabeth I of England, conducted research into magnetism and electricity over a period of about 18 years and published the first scientific book in England of any recognized importance, *De Magnete*, in 1600. Although this work was primarily a discussion of magnetic effects in materials and the materials that produced them, Gilbert included a section on electrostatics as he thought that the two phenomena were related since they both had the ability to attract other materials. His conclusions, which are still correct, were that *a magnet and a static electrical charge have no effect upon each other.* In addition, however, he developed a device which he called a "versorium" to detect the presence of an electrical charge. It consisted of a lightweight metal needle with a sharpened end that was balanced on a second needle so that it could rotate. When an electrically charged object was brought near the device, the needle would point towards the object. He also listed a number of materials other than amber that could develop an electrical charge when rubbed by another object of the group. These materials included diamond, glass, sulfur, sapphire, fur, flannel, silk, shellac, sealing wax, and hard rubber among others. While Gilbert discovered many of the facets of electricity and magnetism that we know today, he also overlooked a considerable number of things that we consider commonplace today. Included in these were the repulsive effects of electrical charges.



The story of the discovery of the repulsive effects of electricity and magnetism involves a short digression back to the efforts of certain religious groups' attempts to stifle inquiries that might result in changes in man's thinking about the universe around him and his place in it. At about the time that Gilbert published his discoveries, Galileo and Kepler were attempting to find evidence to support the theory of Copernicus that the earth was not the center of the universe. In order to do this, they needed a force that would occur naturally and that would cause the planets to move in their observed orbits. They thought that this might be accounted for by the magnetic forces such as those discussed by Gilbert in his hypothesis that the earth acted as a giant magnet. To refute this position and maintain the assertions of the Church that the earth was indeed the center of the universe, an Italian Jesuit priest, Nicolo Cabeo, discovered the repulsive effects of electricity and magnetism some thirty years after Gilbert's work. As well as providing an important discovery in the long road from man's early ignorance to today's supposed complete knowledge of the causes and effects of electrical and magnetic effects, this part of the history also illustrates that the motives for discovery in the field of science, as well as elsewhere - are sometimes less than altruistic. The early scientists, as well as those of today, were merely average people who were interested in a phenomenon, investigated it, and (at least those whom we remember today) had the intuition to recognize just what they had discovered and what the implications of the discovery were.

One final result of Gilbert's work that will affect the story of electrostatics and electricity in general was that he focussed attention on magnets and away from "electrics" as the effects were then called. In addition, it was much easier (as you will see in the laboratory) to work with magnets, and they were found to be useful to navigation. It was not until 1672 that Otto von Guericke reversed this trend with his discovery of a machine to produce "frictional" (static) electricity. His device consisted of a large ball of sulfur mounted on a rod with a handle so that it could be turned. A hand held against the ball produced the electrical charge by friction. Systematic investigation could then be followed as the charge was sufficient to attract small, light objects to the ball and then repel them after they had touched the surface of the ball.

The next important discovery that we shall examine took place in England some 50 years after von Guericke devised his generator. This was the discovery of electrical conduction by Stephen Gray (1696-1736). Gray noted that, when he rubbed a glass in a darkened room, he could observe sparks jumping from the tube to his fingers. From this discovery, he speculated whether "electrick vertue", as he called it, could be transmitted from one object to another. The result of these speculations was the discovery of electrical conduction. Eventually, Gray was able to transmit electrical charges nearly 300 feet along a piece of string. He was also able to distinguish between good and poor conductors of electricity and recognized that all materials could be classified as conductors or non-conductors of electricity. One of his experiments that caught the popular fancy of the day consisted of suspending a small boy in air by means of an insulator (non-conductor), placing an electrical charge on the boy, and then drawing sparks from him by bringing fingers near his body. This experiment became popular throughout Colonial America where it was often repeated with the result that development in the field of electrical research moved at a much faster pace than if it had been restricted to work in private laboratories.

Gray published his work in the *Philosophical Transactions* of the Royal Society where they were later read by Charles Dufay (1698-1736), a Frenchman. Dufay noted that electrically charged objects attracted each other under some conditions, yet repelled each other under other conditions. He suggested that there were two separate and distinct electrical phenomena or "fluids" in matter and that these fluids tended to attract each other when they were on different materials, but repel objects with a surplus of the same type of fluid. This was the birth of the *"two-fluid theory"* of electricity. The two-fluid theory became popular as a means of explaining the action of charged objects primarily in Europe by many experimenters.

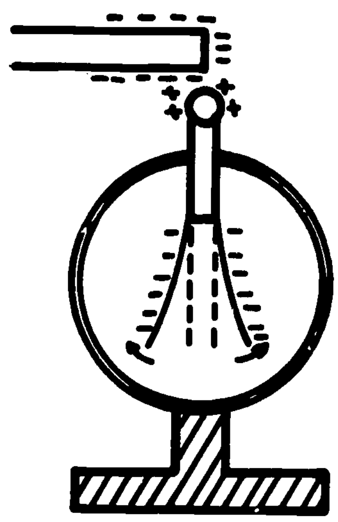
Meanwhile, in America, Benjamin Franklin (1706-1790) also performed various experiments in electricity and, as a result, proposed what has become known as the "one-fluid theory" of electricity. He proposed that every material in nature contained a single type of electrical material or fluid and that, when one object was rubbed by another, some of the material was removed from one of the objects and deposited as a surplus on the other. The two materials - each having a surplus or lack of the electrical fluid - would repel each other; whereas, if one had a surplus and the other a lack - they would be attracted. In order to distinguish between the two types of electrical charges, he called the charge that was deposited on fur after it had been rubbed with a piece of hard rubber *positive* to denote a surplus and the charge deposited on the rubber *negative* to denote a lack of electrical fluid. Franklin could equally as well have named the two charges in the reverse order. Many people, in fact, have expressed the wish that he had done so as we now believe that the material which actually is moved from one object to the other in this case is the electron. As a result of Franklin's choice, today, we say that the electron has a negative charge. The fur in the above example has the positive charge as a result of the removal of the negatively charged electron, leaving a lack of electrons which can move or a surplus of what are called protons which cannot move. Therefore, we say that the proton is *positively* charged.

Modern theories of electricity would have to say that both Dufay with his two-fluid theory and Franklin with his one-fluid theory were correct to some degree. There are two types of particles in all matter which are oppositely charged and attract each other which would support Dufay. However, only one of these is free to move from one place to another in most situations, which would support Franklin. Both of these theories, therefore, have validity in the present understanding of our universe and each should be given equal credit within the restrictions mentioned above.

It is perhaps unfortunate that this most important contribution of Franklin in the development of our understanding is practically unknown whereas one of his other contributions - which, while important, did not advance man's knowledge of the universe as much as the one-fluid theory - is something that might be classified as common knowledge. This, of course, is his celebrated experiment with the kite and the key. Franklin in this case was attempting to prove a theory that lightning was nothing more or less than a gigantic electrical spark. It was most fortunate, both for himself as well as for history, that he was not killed in the process of performing this extremely dangerous experiment. Most people in the vicinity of a lightning bolt are killed immediately, and Franklin - in the course of the experiment - was actually inviting a bolt of lightning to his kite and from the kite down to the ground where he stood. He was extremely lucky that none was forthcoming. An invention of Franklin's that resulted directly from this experiment also is not commonly realized. This was the lightning rod which has been responsible for saving an enormous amount of lives and property from this destructive manifestation of nature.

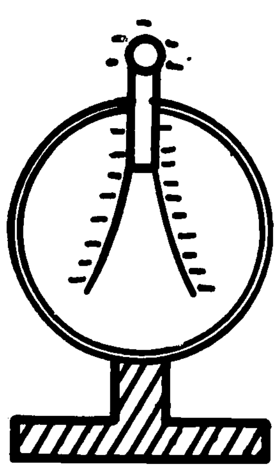
Before we leave the work of Franklin, it would be well to point out another feature of his work and life that was not very uncommon in his day, but that has almost completely disappeared from the life of the present century. Franklin was primarily a statesman, not a scientist. Yet, he made extremely important contributions to science. Until recently, people from all walks of life took an active interest in science in addition to their other occupations. The scientist, in other words, was not shut off from an uncomprehending society at large as he seems to be today. It is a most unhappy state of affairs today that there is such a large gap between the scientists and their works and the general public. This lack of interchange leads to a lack of public appreciation of the developments in science and, perhaps more importantly, to no exchange of ideas between the public and scientists that could lead to new applications and further developments for the benefit of society.

THE START OF MEASUREMENTS One further development took place in the 18th century that contributed significantly to the development of the understanding of electrical phenomena. This was the invention of a device to *quantitatively* measure the size of the electrical charge present in an experiment. Earlier work in the field was dependent upon the use of the experimenter himself as a measuring device to determine the size of the electrical charge present. This was unsatisfactory for two basic reasons. First, there could be a large variation between one experimenter's perceptions and another's, or even a variation from one day to the next. Second, when an experimenter never knew whether the next shock would merely knock him to the floor or kill him outright, there was a very real deterrent to further experimentation. The development of the leaf electroscope solved both of these problems and another as well.



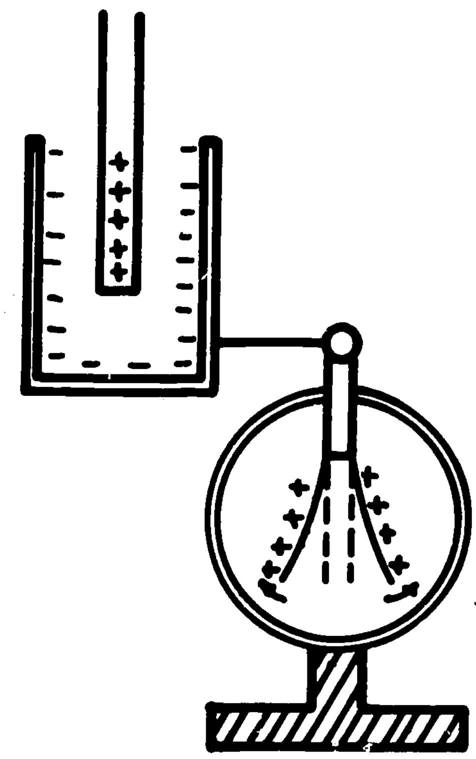
BEFORE TOUCHING

The electroscope consists of a metal rod with a knob that is mounted in an insulated holder and has two thin metal leaves attached to the bottom of the rod. The metal leaves are often gold since gold can be made very thin and flexible. Windows of glass or some other transparent material are often added to the electroscope to protect the leaves from air currents to give a more accurate reading. When a charged object is brought near the knob the electrons in the rod and the leaves will move due to the forces of attraction or repulsion between them and the charged object. If the object is negatively charged, the electrons will move toward the leaves; if it is positively charged, they will move towards the knob. In either case, the leaves of the electroscope will both have the same kind of electrical charge and they will repel each other. If the object is then touched to the knob of the electroscope, electrons will flow either onto or off of the knob to neutralize the charge on the knob. When the object is moved away from the knob, there will be either a surplus or lack of electrons on the electroscope so that the charge of the electroscope will correspond to the charge of the object that brought the charge to the electroscope. This device allowed the experimenter to determine whether or not a charge was present on an object without having to "draw a spark from it" with his body. Also, the amount of charge placed upon the electroscope could be determined roughly by the amount of deflection of the leaves.



AFTER TOUCHING

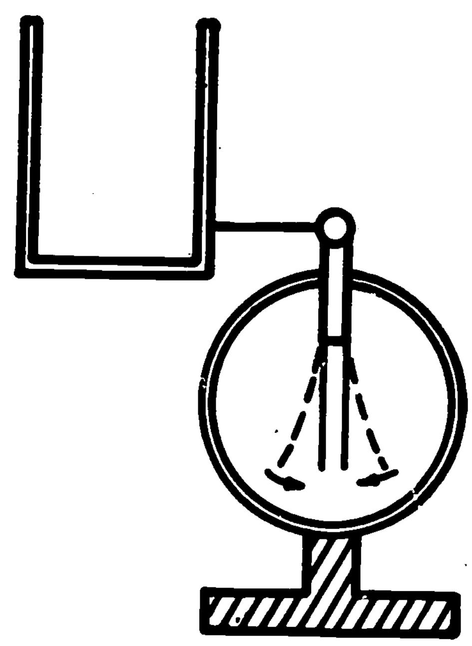
There was an additional benefit from the use of the electroscope. It enabled the experimenter to determine the *type* of charge present either on the electroscope or another object. It also allowed him to determine the *size* of the charge *without having to remove the charge from either the object or the electroscope.* By bringing the second object near the knob of the electroscope, the electrons would move as they did before and, if the two charges were the same kind, the leaves would move even further apart due to the increased force of repulsion; if the charges were opposites, the leaves would not be repelled as much and they would move closer together. The amount of deflection of the leaves would give a comparison of the size of the charges on the electroscope and the object.



START

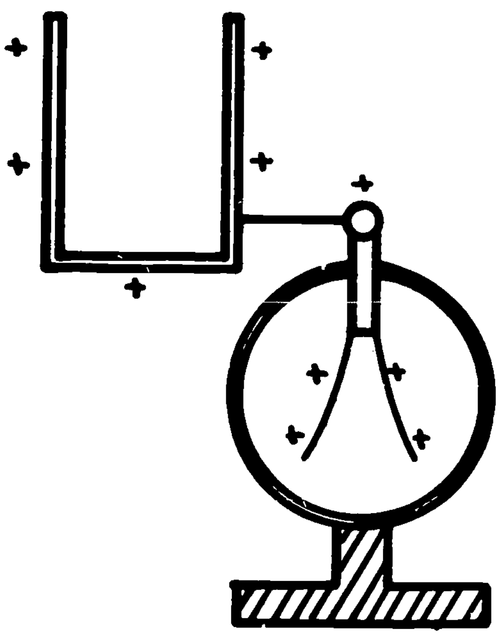
In 1843, Michael Faraday used the electroscope to make some important discoveries. He took a metal ice pail (actually, any metal can will do) and attached a wire to the outside of it and attached the other end of the wire to an electroscope. When he lowered a small, charged, insulated ball into the can carefully so that the ball did not touch the sides of the can, the leaves of the electroscope moved apart. If he removed the ball from the can still without touching the can with the ball, the leaves went back together. He then tried two other experiments.

In the first, he again lowered the ball into the can, but before he removed the ball he allowed it to touch the inside of the can. When he removed the ball from the can, the leaves of the electroscope still were spread apart indicating that the electroscope had become charged. This charging was by means of a process we now call *conduction*. The electrons were conducted directly between the ball and the can and thus to the electroscope. When he made tests to determine the charge on the electroscope, he found that it had the *same* charges as the ball. Therefore, he concluded *when one object is used to charge a second by means of conduction, the second object receives the same type of charge as the first.*



(I WITHOUT TOUCHING
(NO CHARGE)

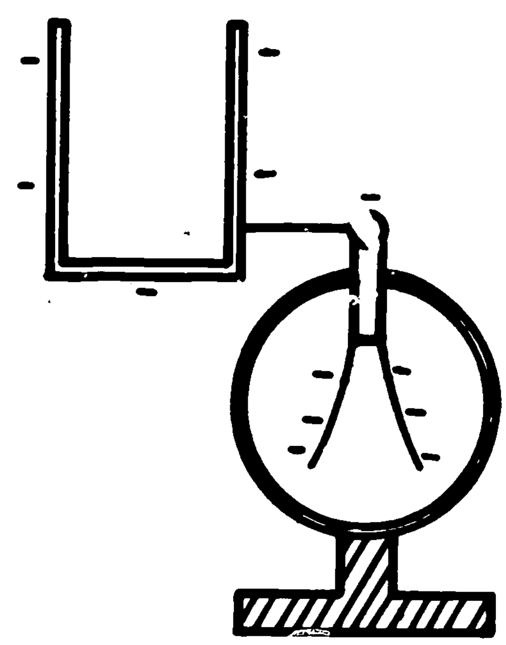
In the second, he lowered the ball into the can, but this time he touched the *outside* of the can with another conductor (in the laboratory with small charges, this could be done by touching the outside of the can with a bare hand). When the ball was removed, the leaves were again spread apart indicating that the electroscope was charged. However, when he made tests to determine the charge on the electroscope, he found that the charge was the *opposite* of that on the ball. The charge was placed on the electroscope by a process that we now call *induction*. In this case, the electrons had flowed along the conductor to neutralize the charge on the ball. When the conductor was removed from the outside of the can, an excess charge had been left on the can of opposite sign to the charge on the ball. This could have been either a surplus or lack of electrons depending upon the type of charge that was originally on the ball. *When one object is used to charge a second by means of induction, the second receives the opposite type of charge as the first charge.*



(2 CONDUCTION
(SAME CHARGE)

In addition to being able to use a charged object to give a second object *either the same or opposite types of charges*, Faraday was able to determine two other important results from this experiment. The first and most important discovery was that *electrical charge can be neither created nor destroyed*. This is known as the *Law of Conservation of Electrical Charge*. In each of the three cases discussed here, the electroscope did not receive a permanent charge when the ball was removed unless the charge had been allowed to flow onto the electroscope from some other object (from the ball in the case of conduction, or the other conductor touching the outside of the can in the case of induction). *There has never been a circumstance observed in which a net electrical charge has been created or destroyed.*

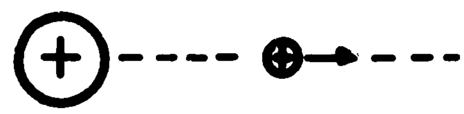
Secondly, and for nearly equal importance, *excess electrical charges are always found on the outside surface of a conductor*. At first glance, this result may seem to be a contradiction to what we have seen in Faraday's experiment, but let us look at the situation a little closer. Before the ball was placed in the can, there was no excess electrical charge present because the electroscope leaves were repelling each other. When the ball was placed in the can, there was an excess charge present due to the ball. Electrons then flowed in the can so that the charge on the inside of the can was equal to but of the opposite type of the charge on the ball. This meant that on the inside surface of the can (or anywhere else inside of the outer surface of the can - including the ball) there was no excess electrical charge. Instead of an excess electrical charge of the same type and size as that on the ball was on the outside of the can.

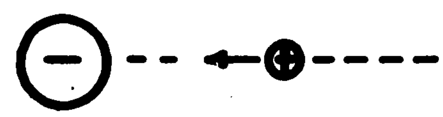


(3 INDUCTION
(OPPOSITE CHARGE)

Another way of looking at this rule to decide whether or not it seems to be reasonable is to take the case of a solid, metal sphere with a surplus of electrons placed on it. The electrons will be free to move anywhere on the surface or inside of the sphere, and they will also repel each other. The force of repulsion will drive the electrons as far away from each other as possible. Any electrons beneath the surface of the sphere will be repelled by each other but also by the rest of the electrons that are scattered *anywhere* on the surface. This force will drive the electrons to the surface where they cannot move any further away. Once on the surface, the electrons will still be repelling each other so that, in order to reduce the force to the greatest possible extent, they will distribute themselves evenly on the surface of the sphere. Similar arguments can be used to describe the process for a lack of electrons (positive charge) on a sphere and for objects of other shapes. (If the shape is *not* spherical, the distribution of the charge is not even, as will be discussed later).

A FUNDAMENTAL LAW While these investigations into the types of charges and methods of charging a conductor were being performed, other investigators were attempting to determine the nature of the forces of attraction or repulsion between electrical charges. It had been noticed that the force seemed to be much larger when the two charged objects were closer together than when they were further apart. This showed the investigators that the force was probably *not* inversely proportional to the distance alone, but to some other relationship between the force and the distance. The next possibility in terms of complication was that the force might be inversely proportional to the square of the distance between the two charged objects. The first mention of this possibility was probably made by Daniel Bernoulli around 1760. Joseph Priestly and Henry Cavendish then made independent experiments using methods similar to but earlier than those of Faraday, but in which they did not reach the same far-reaching conclusions. They did, however, realize that all of the electrical charge and the resulting electrical force on an object was found to be on the outside of a conductor. Using this information and the results that Newton had found for his Law of Universal Gravitation (in which the force was found to be inversely proportional to the square of the distance between the objects), they predicted that the form of the law of attraction or repulsion between two charged objects would likewise depend upon the reciprocal of the square of the distance. They were able to do this because of one of Newton's arguments and results. Using only geometrical considerations and not depending upon the cause of the force, they showed that *there would* be no force of attraction exerted upon a body inside of the attracting spherical body. (An excellent description of this reasoning is given on pages 550-552 of the book, *Physics for the Inquiring Mind* by Eric Rogers).





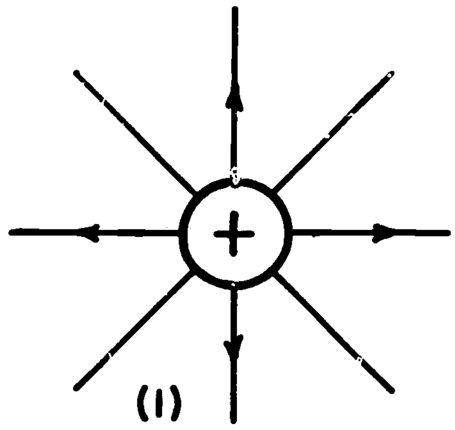
In 1785, Charles Coulomb (1736-1806) performed experiments in France to prove that the force of attraction or repulsion between the charges was indeed inversely proportional to the square of the distances between the objects. (A number of different laws of physics have been given the common name of *inverse square laws*). The final form of the law governing the forces between two charged objects have been named Coulomb's Law in honor of the work that he did and has the form;

$$F = C \frac{q_1 q_2}{r^2},$$

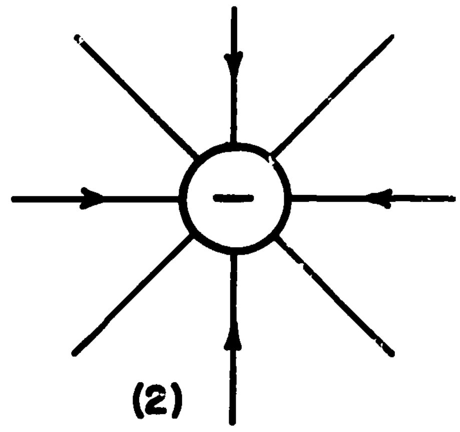
where q_1 and q_2 are the charges on the two bodies respectively, r^2 is the square of the distance between the two bodies, and C is the proportionality constant that converts the units of charge and distance into units of force. Over the years, this law has been constantly checked by experimenters with increasingly accurate equipment and has never been found to be anything other than what is stated above.

"INVISIBLE LINES" IN SPACE Another feature that is commonly used today in the field of electricity that was developed by Michael Faraday is the concept of the *electric field*. Faraday suffered from what we today would call a mental block. He found it extremely difficult to understand or work with the forces between two charged bodies without some medium connecting the two charges through which the forces could act. Many people have the the same problem when they first start to study electricity, and the concept that Faraday developed to solve his problem quite often helps them also. This concept is called the *electric field*, and it is used to illustrate the way that the forces due to a charged object would act on another charged body *if a second charged body were present*. Each separate electrical charge has an electric field associated with it, and the sum of these separate electric fields will describe what would happen to another charge if it were placed in the region of the previous charges. Since we are dealing with the forces that would be exerted on the "test" charge as this last, extra charge is called, the field must tell us the *direction* as well as the *size* of the force that would be exerted upon this charge. The test charge that is used is considered to be always *positive*.

For example, if we had a positive charge located by itself somewhere, a second positive charge would always be repelled by it. The second charge would be repelled whether it was "above", "below", to the "left", to the "right", or any where else in relation to the first charge. The electric field would, therefore, be pointing outward in all directions from the positive charge and, since the force is inversely proportional to the square of the distance from the charge, the field would get smaller as the distance from the charge increased. The electric field is often thought of as a series of lines showing where the force on the test charge would try to move it. Similarly, if we had a negative charge rather than a positive one, the electric field would again consist of a series of lines radiating from the charge in all directions, but this time the lines would be pointing inward as the test charge would be *attracted* to the charge. If we had two charges, one positive and the other negative, some of the lines would go directly from the positive charge to the negative charge as the test charge would be repelled by one and attracted to the other. These lines have some additional properties which may be roughly described as the following:



(1)

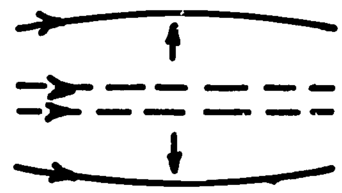


(2)

1. The lines *initially* move outward from a positive charge uniformly in all directions and perpendicular to the charged surface.

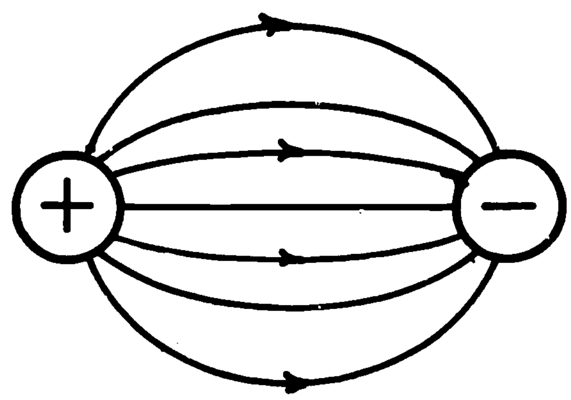
2. The lines *initially* move inward near a negative charge uniformly in all directions and perpendicular to the charged surface.

3. Two parallel lines *going in the same direction* will *repel each other.*



(3)

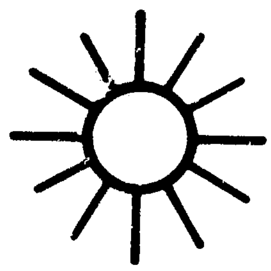
4. A line will go from a positive charge to a negative charge by the shortest possible route. (It may help to think of the lines as being stretched rubber bands going from the positive to the negative charge).

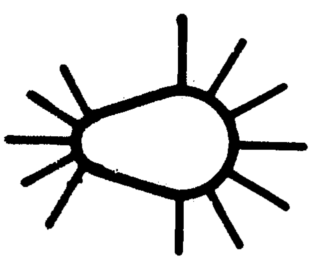


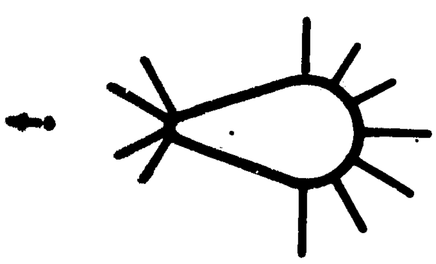
(4)

It must be remembered that these lines do not actually exist in the sense that they can be seen. They merely described the path that a small, positive electrical charge would take if it were placed on a line. It is possible to illustrate the location and shape of the lines, however, by using an insulating liquid (such as distilled water) with small particles of insulator (such as grass seed) floating on the surface and placing charged conductors on the surface of the liquid. The particles of the insulator will have their charges rearrange themselves so that one end will become slightly positive and the other slightly negative. The closer these particles are to the charged conductors, the greater the rearrangement. The particles will then form chains along the location of the electric field lines. (This is similar to the method of showing *magnetic* field lines by means of iron filings in the vicinity of a magnet).

Before leaving this discussion of the electric field, the definition of a field in science (this is true for any kind of a field; electric, magnetic, or gravitational) may be given as: *A field is a region in space where a force would act on a test object if the object were present.* A field by itself cannot be seen and its presence cannot be detected without the aid of some kind of test object. Its usefulness is restricted to *what would happen if the test object were present* and it is used in describing a given situation.







We may now apply the concept of the electric field to see how Franklin's invention of the lightning rod works. Consider a sphere with an electrical charge on it. The electric field inside of the sphere is zero as there is no force exerted upon an object inside a conductor. The field lines on the outside of the conductor must, therefore, bend outward from the surface of the conductor. These lines and those from the surrounding charges repel each other so that the charges will be evenly distributed all over the surface. Suppose, now, that one part of the sphere is drawn out from the surface of the rest of the sphere so that the conductor is now egg-shaped. At the small end of the object the curvature is greater so that more charges can be pushed in and still have the same amount of space between the electric field lines. If this small end is made sharper and sharper, eventually the charges at the top will have field lines that are almost the same shape that they would have if the rest of the object and other charges were not present. These charges still want to have their lines as far away from each other as possible, so they will be pushed off of the surface completely. The other charges will rearrange themselves after the first charges have left and the process will repeat itself until the number of charges remaining is such that the field lines do not have enough force between them to force the charges off of the surface. Thus, when there is a sharply pointed object, electrical charges will gather near the point and force each other off until the forces between the charges have been reduced to the point that they can no longer do so.

Let us now look at how this would apply to a lightning rod. When air passes over the ground or when two separate layers of air rub against each other, static electrical charges are produced. When these charges reach a certain density, there is a spark, or discharge, between the two layers that we call lightning. If this charge builds up between the air and ground, the lightning may strike the ground. Furthermore, the spark will take the shortest possible conducting path so that the lightning may strike a tree, house, or barn in preference to the flat ground. If we take a sharply pointed metal rod and mount it on the top of a house or some object that we do not want the lightning to strike and attach a metal wire from it to the ground, the excess charges on the ground will gather at the point of the rod and leak off into the air and, thereby, neutralize the charges in the atmosphere, therefore the electrical charge cannot build up to the point where a discharge can take place and the house or barn will not be struck by lightning.

There is one more important concept that we should examine before we move on from static electricity to magnetism. This is the concept of *electrical potential.* We mentioned earlier that the negatively-charged electron is, for the most part, the source of electrical current as it is the particle that is able to move in materials whereas the positively-charged proton is fixed in the atoms of the material. Suppose that we take an electron and move it along the direction of an electric field line. The further along the line we moved the electron, the greater the force of repulsion that we would have to overcome as the electric field lines always point away from positive charges and toward negative charges. We would, therefore, have to exert a *force* on the electron in order to make it move. This force is often referred to as the *electromotive force* or *emf.*

Whenever you apply a force to an object to move it through a distance, the object gains energy. If the energy gained is stored, it is called potential energy. An example of this would be the energy stored in a large rock that has been pushed up a hill. As long as the rock is sitting at the top of the hill, the energy is stored or has the potential for being released when the rock rolls back down the hill. In the electrical case, the principle is the same. When the electron is moved along the direction of the electric field line, it is gaining energy that is being stored and it has a higher potential energy. Potential energy is released by letting the object - whether it is a rock on a hill or an electron that has been moved along an electric field line - move from a position where it has a higher potential energy to a position where the potential energy is lower. It is often meaningless to refer to an *absolute* value of potential energy. The important value is the *difference* in potential energy the object would experience in moving from one point to another.

To return to our example with the rock, we could choose the place where the potential energy is zero to be whereever we want. For instance it could be a wall that is half way down the hill, a valley at the bottom of the hill, a valley at the bottom of the next hill that is lower, etc. Our choice of zero potential energy would make no difference in the amount of energy the rock would release in rolling down from the top of the hill to the wall half way down the hill. Similarly, we could choose zero potential energy for an electron to be anywhere along an electric field line from the positive charge to the negative charge. The important value would always be the difference in the potential energy from the original position of the electron to its final position. If the potential energy is higher at the end than the beginning, we have had to put energy into the electron (by means of the chemical energy in a battery or the mechanical energy in a generator; these sources will be discussed in detail later); if the potential is lower at the end than the beginning, we have gotten energy out of the electron (to turn a motor, operate a light, etc.). This change in the electrical energy is referred to as the potential difference. It is measured in units of the change in energy per electron in the circuit. The unit of electrical potential difference is the *volt*. In a flashlight battery, the amount of energy received by each electron is very small, 1.5 volts. In a wall outlet, the change in energy received by each electron from the generating station is much larger, 110 volts.

MAGNETISM The first discoverer of the effects of magnetism is unknown to history. As with the story of the antiquity of electrical effects, naturally occuring magnets have been found whose age greatly exceeds the length of time that man has been on earth. Magnetic effects have not been accorded as great a place in early history as have electric effects either. This is due to their not being as vigorous and awe-inspiring as, say, a bolt of lightning. Probably the first discovery of the phenomena of magnetism occured when an early man collected a special kind of rock which we now call magnetite for use back at his cave to fashion weapons and discovered that similar pieces of rock attracted or repelled each other.

As the effect was not very violent or spectacular, this led to these rocks being more of a curiosity than something to be feared. With the coming of the iron age to man, another feature of these strange rocks was discovered. They would attract iron implements to them.

As had happened to electrically-charged objects, these rocks were experimented with widely as a means of transforming or assisting in the attempted transformation of other metals into gold by the alchemists. Early doctors attempted to cure various illnesses by using these strange rocks to "draw out evil humors" from their patients. Since the attractive properties of the magnetic rocks did not have to be induced in them as was the case with electrically-charged objects, the magnetic rocks were preferred and had a higher value in their supposed properties in these attempts.

A GUIDE FOR SAND AND SEA Investigations into other possible properties and uses of magnets were carried on by many people throughout the years with the result that many misconceptions and fables sprang up. Examples of these were the following: Archimedes was supposed to have developed a device to attract the iron nails out of enemy ships to make them fall apart; rubbing an onion over the surface of a magnet was supposed to remove its attractive properties; and so on.

All of the early attempts to find an application for the effects of magnetism were not doomed to failure, however. Early man had long noticed that, north of the earth's equator, there was one star, called Polaris or the North Star, whose position in the sky did not vary during the night or from one season to the next as the other stars all did to some extent. Some ancient, unrecorded genius thought to make use of the magnetic rock as a means of determining the direction of north during the day when the North Star was not visible and, thereby, improve his navigation during the daylight hours. This was probably first done by someone who travelled with the caravans across the deserts. These first compasses as we call them today were probably very crude and consisted of the stone resting on a wooden raft in a container or bowl of water carried from place to place.

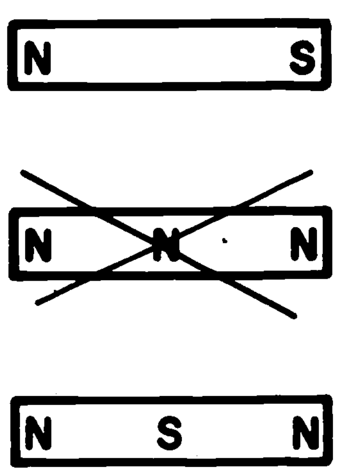
The next improvement in the use of these lodestones (or leading stones) as they were called came about when man was able to process mercury out of ores. Mercury is a metal that is a liquid at normal temperatures and has a very high density. As the density of mercury was greater than the density of the lodestones, the stones could float on the surface of the mercury and greatly simplify the construction of the early compasses. Furthermore, the mercury would not evaporate easily and thus provided an additional advantage over the use of water and wooden raft.

With this improvement, the use of lodestones as a navigational device spread rapidly among travellers both on the deserts and on the seas. Their use became known in Europe in the twelfth century, although no one has determined when they were first used or when. The first written record of the magnetic compass appears in a Chinese document in the eleventh century A.D. Later, it was discovered that, by rubbing a piece of hardened steel along the side of a lodestone, the steel would also become a magnet with the same properties as the lodestone. From this, we have obtained our modern magnetic compasses for use in navigation.

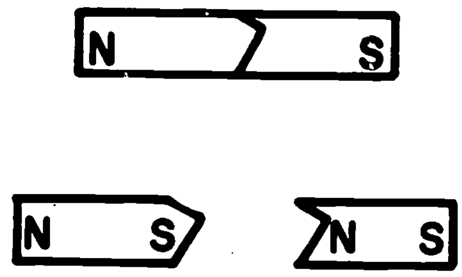
As was stated earlier, William Gilbert was the first person to systematically study the effects and properties of magnets and a great deal of his book on the subject was devoted to dispelling the misconceptions of his predecessors. One of the great advances he made on the subject was his discovery that the earth itself acts as a gigantic magnet, and that this is the reason for the compass to point toward the north.

Charles Coulomb, in addition to his discovery of the inverse square law of distance for electrical forces, made a similar discovery for magnetic forces. Thus, we find that all three of the known forces that act a distance (that is, the forces that can be transmitted from one body to another without actual physical contact between them) and which we can represent by means of fields, the forces diminish proportionally as the square of the distance increases.

There are other striking similarities between electrical and magnetic effects and some important differences. Therefore, rather than merely listing the magnetic effects and features separately, it might be best to compare them point by point with those of electricity.

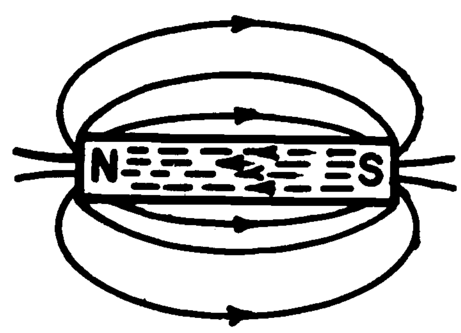


As with the electrical effect, there are two types of magnetic regions, that are known as *poles*. These have been defined as north-seeking (or north) and south-seeking (or south). Two similar poles repel each other,and and two opposite poles attract each other as with electrostatic charges. However, and most importantly, *a separate magnetic pole cannot exist in nature*. Whenever one magnetic pole is found in a material, the opposite pole is also found in the same material. Thus we can never have a single north pole, for instance, as we could have a positive charge on an object. We always have a north and a south pole in the same object. It is possible, however, to have an object whose two ends are both of the same type, but the center of the object will always be of the opposite type. Thus, we could have a bar of iron whose ends were north poles, but the center would always have to be a south pole, and vice versa. Furthermore, if we take a magnet with a north pole at one end and a south pole at the other and cut it in half, each of the two ends that were at the point where the cut was made would become opposite poles to the poles at the respective ends of the original magnet. (That is, the cut end attracted to the north pole would become a south pole and so on.)



Only certain materials are affected by magnets or attracted to them. Most of these materials, with a few exceptions, contain iron, This is a great difference when compared to the electrical effects that we looked at earlier in which most insulators, or non-conductors of electricity, could be acted upon by an electric charge and attracted or repelled. Furthermore, most of the materials that are attracted to magnets cannot become magnets themselves permanently, but exhibit properties of magnets only when in the vicinity of a magnet. Thus, a magnet will not affect bits of paper the way that they would be affected by an electrically charged object, but it will affect iron nails or paper clips or a hardened steel bar. When the magnet is near or actually holding the iron objects, the objects themselves become magnets with the end nearest the magnet becoming the opposite pole of the magnet that is nearest. When the magnet is removed from the objects, most of them will lose their magnetic properties except for a very few objects such as hardened steel. Thus a hardened steel bar, such as a screw driver or knife blade may become magnetized as a result of being in contact with a magnet, whereas a soft steel bar, such as a nail or a paper clip will not.

The explanation for this property of some materials and not others is that in iron and a few other materials, the atoms themselves behave like small magnets. These atoms form small regions in the material called *domains* which are normally oriented randomly in the material. When a magnet is brought near the material, the domains all shift around and align themselves with the magnetic field and the material becomes a magnet. In the soft iron-type materials, the domains shift back to their original random orientation when the magnetic field is removed, but the structure of the hardened iron is such that the domains are held in the new alignment. Hardened iron or steel rods are much more dufficult to magnetize, but they will retain their magnetic properties after the original magnet has been removed. Even they can lose their magnetic properties if they are subjected to a shock when there is no magnetic field present. If a "permanent" bar magnet is rapped against a table top when there is no other magnet near it, it will lose some or all of its magnetic properties.



In order to explain and predict the actions of magnets in attracting and repelling each other, we can define magnetic fields in a similar manner to electric fields. These magnetic fields are defined by both magnitude and direction just as they were for electric fields. In this case, the magnetic field is defined as the direction that a small "test" magnet points, and the force that would tend to align the "test" magnet represents the size of the magnetic field at that point. It is much easier to look at magnetic fields than it is to look at electric fields. First of all, if we have a "permanent" magnet and do not disturb it, we do not have to worry about the magnetic effect deteriorating in time as we do with electrical effects. Also, we can use small pieces of iron, called iron filings, to show the shape of the field near the magnet. By placing a piece of paper over the magnet and sprinkling the filings in the paper, we can see where the magnetic field lines are by tapping the paper so the fillings, which become small temporary magnets themselves, can realign themselves with the magnetic field of the larger magnet. Once again, it should be emphasized that there is actually nothing present in a magnetic field that can be seen. *The only way that we can determine if a field is present or not is by placing a test object in the region and seeing if it is affected by the field.* The use of the field concept is, however, very useful in explaining some of the effects of both electricity and magnetism, as we shall see later.

THE WORLD'S LARGEST MAGNET There is one important magnetic field that we should look at more closely. It is the earth's magnetic field. William Gilbert was the first person to realize that the reason that magnets were affected by their orientation on the earth was that the earth itself was acting as a gigantic magnet. Measurements made since his day have come up with the following;

1. The earth behaves as an extremely large, stubby magnet mounted near its center and oriented in a nearly north-south direction.

2. The cause of the earth's magnetic field is not known exactly, but it is thought to be due to an iron nickel core located at the center of the earth.

3. The poles of the earth's field are not located exactly along the north and south poles and they have apparently shifted over the years since rocks first hardened on the surface of the earth.

4. *The pole closest to the geographical north pole is a south-seeking pole* and vice versa. (This is true because it attracts a north-seeking pole and only opposite poles will attract).

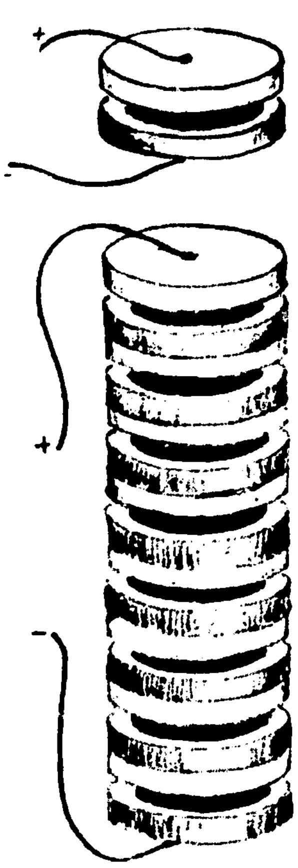
5. The magnetic north pole (to which north-seeking compasses are attracted) is located presently about 1300 miles or so from the geographical north pole (about which the earth rotates) in northern Canada.

6. In the United States, this means that compasses in the eastern states will point to the west of the true north and in the western states they will point to the east of true north. Since the magnetic pole is still moving slowly, this error also will change slowly from year to year. The greatest errors are about 20° each in the states of Maine and Washington. The error is, of course, much greater in Alaska.

7. Since there must be field lines from the south magnetic pole going to the north magnetic pole, both of which are located deep in the earth. There must be a vertical component of the earth's field as well as a horizontal one. In the northern hemisphere it will point upward.

PRACTICAL APPLICATIONS Electrons are free to move in conductors (which are primarily the metals) and go from a point of higher potential to one of lower potential. If we can change the potential energy of electrons by some means, we can form what is known as an electrical circuit and use electrons to transfer energy from one place to another where the energy can be transformed into some other type and be of benefit to us. The first step in this chain is, however, finding some way and means to change the potential energy of the electrons in some region. The first clue to this came in a very unexpected way.

An Italian doctor, Luigi Galvani (1737-1798) was studying the nervous system of a frog in his laboratory one day when he happened to touch the scapel he was holding to the exposed nerve of the frog's leg. A violent twitching occurred and an assistant in the laboratory reported that he thought that there had been a spark from a frictional electrical device nearby. Galvani thought there might be a direct relationship between the frog's leg and the device. Actually, there was no relationship, but further investigation revealed that the muscle of a frog would convulse whenever two different metals were attracted to the ends of the muscle and connected together. Galvani believed that this was due to some sort of frictional electricity in the frog.



A contemporary of Galvani's, Count Alessandro Volta (1745-1827), who had been doing his own research into the nature of electricity that had resulted in the invention of the electrophorous, a device for producing electrical charge, investigated Galvani's theory of "animal electricity" and found it to be false. Volta discovered that whenever two dissimilar metals were placed in contact with the muscle of any one of a variety of other moistened materials, an electric current would flow due to chemical reactions taking place between the metals and the other material. He published these results in his book, *On the Electricity Excited by the Mere Contact of Conducting Substances of Different Kinds* in 1800. The first batteries as we now know them were called "Voltaic Piles" and consisted of discs of copper and zinc spaced alternately with pasteboard discs soaked in water to which a few drops of acid had been added. From this start, all of the modern batteries have been developed. In honor of this accomplishment, the unit of electrical potential difference has been named the *volt*.

Due to the reactions taking place in a battery, there always are two separate regions in the battery that are of different electrical potential. Electrons receive chemical energy from the reactions and move from the point of lowest potential to the point of the highest potential. If a conductor of electricity is connected between these two points, the electrons will flow from their place of highest potential energy to the place where the energy is lowest. In the process of flowing, they give up their energy in the wire.

The rate at which the electrons flow past a given area in a conductor is defined as the *current* of electricity. It is measured in terms of units called *amperes*. Georg Ohm discovered in 1826 that there is a direct proportionality constant which depends upon the type of conductor and the length of the conductor. This direct proportionality holds for a wide variety of conductors and has been given the name *electrical resistance* as it may be thought of as resisting the flow of electrons from the higher to the lower potential energy. This may be stated as the ratio

$$R = \frac{V \text{ (potential difference measured in volts)}}{I \text{ (current measured in amperes)}}$$

The unit of electrical resistance is called the *ohm* in honor of Ohm's work. A resistance of one ohm will result in a current of one ampere when the potential difference is one volt.

Another useful quantity is the total energy being given up by the circuit or being transformed from electrical energy to some other form or forms of energy in the circuit in a period of time. Energy per unit time is called power and the unit of power that had already been developed for use in heat and mechanical energy cases was adopted as the unit in electricity. This unit is called the *watt*. Since a volt is the measure of the amount of energy contained in each electron in passing from a point of one potential energy to another potential energy and the ampere is the measure of the number of electrons flowing through a given area per second, the power released in a circuit is merely the product of the two, or

$$P = IV .$$

Therefore, one watt equals one volt times one ampere.

We may take the electrical energy that can be obtained from chemical energy in a battery or from mechanical energy in a generator (to be discussed later), transport it great distances, and then transform it into other forms of energy at a desired location. The transformation may be into light, heat, mechanical, or other forms of energy. There are certain common characteristics that may be found in any transportation of energy by this method.

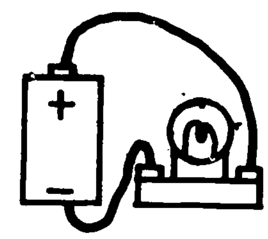
THE PATH OF A STREAM. The method of transportation is referred to as an *electrical circuit*. In some ways, the flow of electrons through a conductor may be considered to be the same as the flow of water through a pipe. (Many people like to draw this analogy when discussing simple electrical circuits with others, but as with most analogies a point is quickly reached where the comparison is no longer valid. Therefore, we shall probably err on the safe side to keep the problems to a minimum.) A conductor could be compared to a pipe that is already filled with water. The source, let us consider only batteries for the moment, would then be similar to a pump to supply pressure (energy) to the water. Finally, we could use a tap or valve to act as a switch to start or stop the flow of water. Let us assume that the pump is always running so that the water drops (electrons) on one side of the pump (battery) have a higher pressure (amount of energy) than those on the other side. When the valve is opened, the water will start to flow from the side of the pump with the higher pressure to the side with the lower pressure if the pipe containing the valve is connected directly between the two sides of the pump.

Unlike water, electrons cannot merely flow out of the end of the conductor, so that if the conductor for the pipe is not connected in a closed loop in this manner, it would be the same as if the valve had never been opened. That is, there would be no flow, but the difference in pressure would still be present. Also, the flow would start all around the circuit much faster than the time required for a single drop of water or a single electron to move from any one part of the circuit all around the loop due to the particles with a slightly higher potential energy pushing those with a slightly lower potential around the circuit. If the valve is closed for the water or there is a break in the chain of conductors around the loop for the electrons, all of the motion stops for either case. The current will always be flowing in the same direction, so we would say that we are dealing with a *D. C. or Direct Current system*.

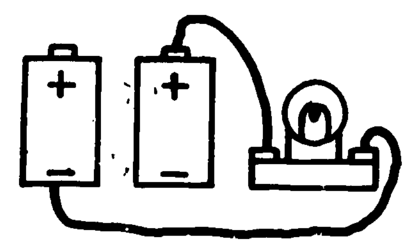
A LIGHT GOES ON In order to investigate the properties of an electrical circuit, let us choose the simplest possible circuit components that will still allow us to find out all of the peculiar features that might be present. For our power source, we can use batteries; for the object to use the energy present, we shall use light bulbs as they will indicate whether or not the current is flowing by being either on or off. Furthermore, the brightness of the bulb will give an indication of the amount of energy that is being transferred from each electron as it passes through the bulb (that is, the voltage of the system). some wires and switches will complete the equipment to guide and control the flow of the current to where we want it.

First, let us consider the wire. Most of the wires that are used in electrical circuits consist of a metallic conductor with a outer covering of an insulating material so that the flow of electrons cannot be diverted accidentally by having two conductors touch each other where they are not supposed to. Next, we have the switch wiich is merely a device that allows us to conveniently make a break in the path of the electrons around the circuit that can be easily reconnected. We do not actually have to use a switch in the circuit if we want to disconnect one of the wires instead - we merely use it for convenience and safety.

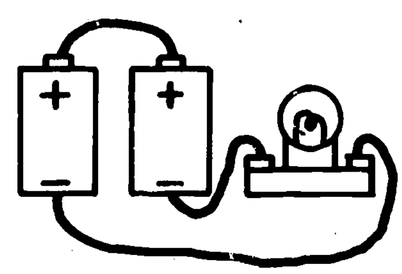
The simplest possible circuit we could make from our equipment would consist of connecting a wire from one end of a battery to one end of a bulb and then connecting a second wire from the other end of the battery to the other end of the bulb. When the last connection is made, electrons will flow from the negative end of the battery through the bulb and back to the positive end of the battery. When the electrons pass through the bulb, they lose their energy in heating up the small wire in the inside of the bulb (through which they must flow), called the filament of the bulb. This filament becomes white hot and produces the light that we can see. If there is a break in any part of the path, the electrons will not be able to flow and the bulb will not light even though there are still electrons in the filament of the bulb. It requires a *flow* of electrons from the battery to the bulb and back again for there to be energy transferred to make the bulb filament light up.



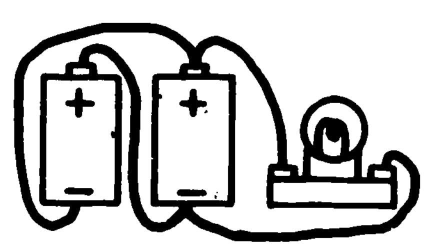
THE LIGHT GOES OFF Instead of just one battery, let us see what we can accomplish by adding a second battery to our circuit in various ways. First, let us disconnect the wire that we had connected to the positive end of the first battery and re-connect it to the positive end of the second battery. Notice that there is now a potential difference between the negative end of the first battery and the positive end of the second. But that *the bulb does not light*. Look again at this circuit. There is *no closed path* or loop that an electron may follow from the source to the object and back to the original source again. *Each electron must be able to return to its original starting point in order for it to flow at all.*



Remove the wire from the positive end of the second battery and attach it to the *negative* end of the second battery. Connect another wire between the *positive* ends of the two batteries. The bulb still does not light. There is a closed path for the electrons to flow around in this case, but the potential at each side of the bulb is the same. *In order for energy to be transferred by electrons from a source to an object, there must be a difference in electrical potential from one side of the object to the other* as well as a closed path for the electrons to flow around.



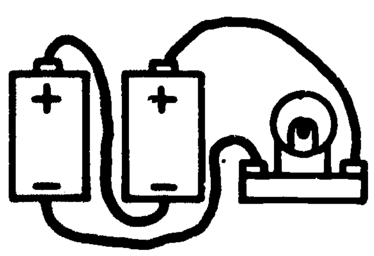
Let us now remove the second battery and the extra wire and return to the original circuit that we had using one battery and one bulb where the bulb did light up. Now let us connect a wire from the positive end of the second battery to the negative end of the first battery and another wire from the negative end of the second battery to the positive end of the first battery. The bulb goes out again! Look carefully at this circuit. You should be able to see that either of the batteries by itself is connected in such a way that there is a complete



loop for the electrons to flow along and that there is a difference in potential between one side of the bulb and the other due to *either* battery.

But look again at how the batteries are connected. An electron leaving the negative end of either battery can move immediately to a point of lower potential at the positive end of the other battery, gain energy in passing through it, and return to its original battery by losing potential energy in passing from the negative end of the second battery to the positive end of the battery it originally started from. By following this path, the electron not only is able to gain and lose more energy by passing through *both* batteries, but also is able to follow a path of *lower resistance* than it would have encountered by passing through the bulb. This brings up a third point: *an electron will always tend to follow the path of least resistance*. We have now seen three possible ways in which to make up a circuit involving two batteries and one bulb in which *the bulb will not light*. There are two other possible ways to hook up two batteries and one bulb in which the bulb will light. These involve the important concepts of *series* and *parallel* circuits.

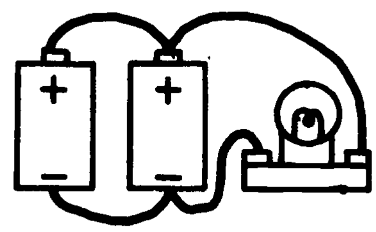
IN SINGLE FILE A *series electrical circuit* can be easily remembered because it is just what the name implies the elements are arranged in a series, one *after* the other. A series circuit usually implies two or more of the same kind of circuit elements connected so that the electrons have to flow through each of the elements in turn. They do not have a choice of flowing through one or another of the elements. The elements involved could consist of two batteries, two bulbs, or two or more of any other type of circuit element. As an example, let us connect two batteries in a series with one bulb. Start with the original circuit of one battery and one bulb in which the bulb was lighted. Disconnect the wire from the negative end of the first battery and connect it to the negative end of the second battery. Connect another wire from the positive end of the second battery and connect it to the negative end of the first battery. The bulb will now light and be about twice as bright as it was before.



Look at the circuit. An electron flowing from the bulb passes through one battery and gains energy from it. It must then pass through the *other* battery where it gains still more energy. When it emerges from the two batteries to flow towards the bulb again, it has had to pass through *both* of the bateries so that there is *twice* the potential difference between one side of the bulb and the other as there was when there was only one battery in the circuit.

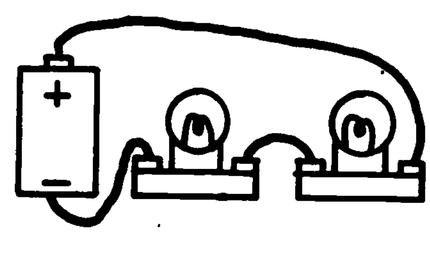
Any number of batteries may be connected in series in this manner. As long as the negative end of one battery is connected to the positive end of the next in the chain of series-connected batteries, the resulting potential difference available will be equal to the *sum* of each of the individual potential differences in the batteries. In an automobile battery of 12 volts, there are six 2 volt batteries connected together in series to raise the voltage available for starting the car, operating the lights, etc. If you examine one of these batteries carefully you may be able to see where the connections are made between one battery and the next.

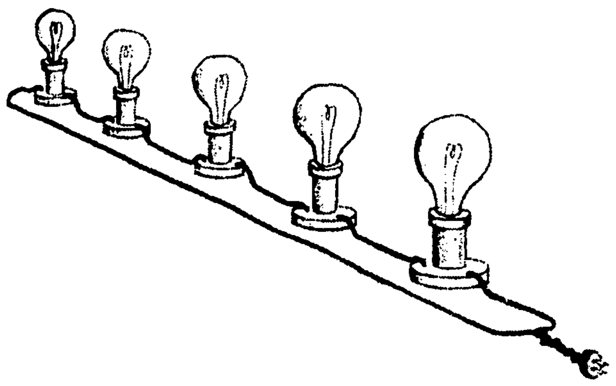
OR SIDE BY SIDE In a like manner, a *parallel electrical circuit* consists of any number of similar elements connected together one *beside* the other. An electron must, therefore, flow through one *or* another of the elements. Once again, the elements could be batteries, bulbs, or something else. As an example of a parallel circuit, let us now connect the two batteries in parallel with one bulb. Start once more with the circuit of one battery and one bulb in which the bulb was lighted. Connect a wire from the negative end of the second battery to the negative end of the first and another wire from the positive end of the second battery to the positive end of the first. The bulb is now as bright as it was in the original case in which there was only one battery connected to the circuit. If you examine the circuit, you will see that any particular electron can go through only one battery *or* the other to gain energy and that the potential difference across the bulb as a result is the same as it was with only one battery. Parallel circuits for batteries are used to prolong the *life* of a battery. Only a fixed amount of chemical energy can be stored in a battery and when it is transformed the battery is "dead". By connecting batteries in parallel they last longer.



So far, we have seen how we can either increase the potential difference or increase the life of batteries by connecting them into electrical circuits in series or parallel respectively. There is a corresponding difference if we connect two or more bulbs in series or parallel with single source as some of these differences are important, we shall look at them next.

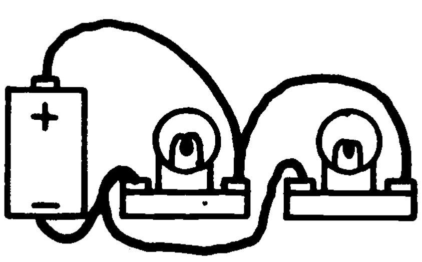
Let us start with a battery connected with two bulbs that are in series. This means that electrons leaving the negative end of the battery pass first through one of the bulbs and then the other. If both of the bulbs are are of the same size as the bulb that we used earlier, we would notice that the two bulbs are of equal brightness and that each is about one-half as bright as when we were using only one of them. If we unscrew one of the



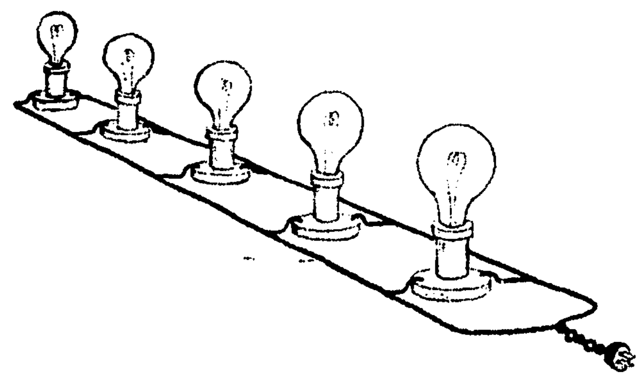


bulbs or it burns out, the other bulb must go out also as there would no longer be a closed path for the electrons to follow from the battery to the remaining bulb and back to the battery. Suppose now that we unscrew one of the bulbs and replace it with another that is larger (that is, one that has a higher power rating - such as exchanging a 60 watt bulb for a 100 watt bulb). The brightness of the remaining bulb that had not been changed would be slightly brighter than before while the larger bulb would be even dimmer than the bulb that was present originally. This is explained by noting that the current through the circuit depends upon the sum of the resistances in the circuit and that the total potential difference across both of the bulbs remains constant. When we put in the larger bulb (with a smaller resistance) the total current was increased so that the total power was increased. This power was divided proportionally to each of the bulbs according to their resistance and the smaller bulb with the higher resistance received the larger amount of power. The difference of potential across each of the bulbs was also directly proportional to the resistance of the bulbs so that the larger resistance had the greater potential difference across it and the sum of the potential differences across the battery.

This feature of series circuits is used in homes in the system of fuses that keep circuits from releasing more power than the wires and insulation can withstand. By placing a small piece of metal with a low resistance that melts easily in series with the rest of the circuit, practically none of the power is released in the fuse and is left for the rest of the lights, motor, and other appliances. However, when the current through the fuse is large enough to be dangerous for the rest of the wiring, the fuse receives enough power to melt the metal and break the connection. Since fuses are carefully designed to protect the electrical circuits they are included in, it is never a good idea to replace them with a larger size fuse designed for greater currents or with a penny as the protection would no longer be present and there would be a possiblitiy of a fire if anything went wrong with the circuit.

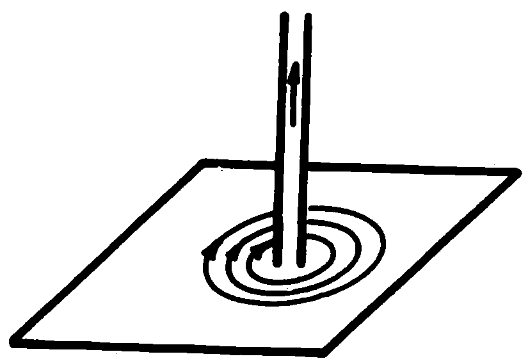


Now, let us look at two bulbs in parallel connected to a battery. In this case, each of the bulbs has a separate path that does not involve the other bulb through which it can receive electrons from the battery. Thus, if one of the bulbs is unscrewed or burned out, the other is not affected. Also, the potential difference across each of the bulbs is the same as that of the battery itself, so the bulbs burn just as brightly as if the other were not present. The current that must be supplied by the battery is larger, however. In fact, it is equal to the sum of the currents that would be supplied to each of the bulbs separately. Since each of the bulbs is completely independent of the other, this is the best type of wiring circuit to use in a house where you want to be able to control a number of bulbs with the same circuit. There is another advantage to the parallel circuit. If one of the bulbs in a ceiling light fixture were to burn out, the rest would still be working;

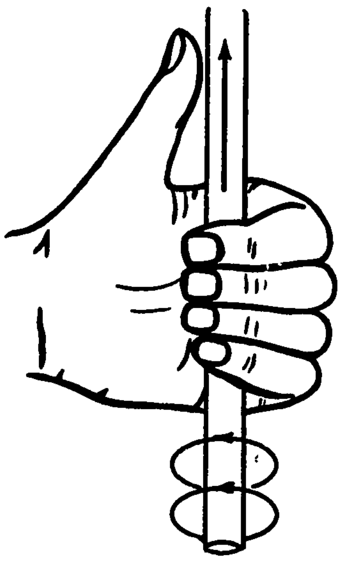


There are other, more complicated, circuit elements that we can include in our discussion of electrical circuits, but they would be connected either in series or parallel with the existing elements and we have to understand the relationship between electricity and magnetism before we can see how they would work. Therefore, we will discover how electricity and magnetism are related first and then we will return to look at these electrical circuits afterwards.

ELECTROMAGNETIC INDUCTION In 1819, the Danish physicist, Hans Christian Oersted (1777-1851) was explaining the flow of electrical current in a wire to a group of students. In this particular case, he was using a large battery so that there was a large flow of current through the wire. A compass happened to be lying on the table near the wire and one of the students noticed that the needle of the compass was not pointing towards magnetic north as usual, but was being affected by the current flowing through the wire. Further investigations showed Oersted that the *current flowing through* the wire set up a *magnetic field around* the wire, that the strength of the magnetic field was related to the size of the current, and that the direction of the magnetic field depended upon the direction of the current.

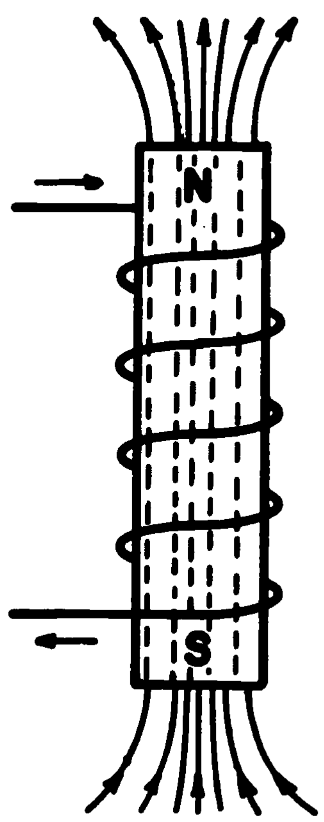


After Oersted published his findings, they were quickly translated into many languages. One of the physicists who read of his discovery was Andre Marie Ampere (1775-1836) in France. Ampere investigated Oersted's results and carried them further. He distinguished between static electrical phenomena and dynamic (or moving) phenomena. Furthermore, he was the first to distinguish between an electrical current and the cause of the current. He also found that one of the results of the magnetic field around a current-carrying wire was that there would be a force of attraction or repulsion between two current-carrying wires that were parallel depending upon whether the currents were parallel or anti-parallel respectively.



One of the further results that came out of all of the these investigations was that the direction of the magnetic field can be found easily once the direction of the electron flow is known in the wire. *Taking your left hand, if you place your thumb in the direction of the flow of the electrons, your other fingers will wrap around the wire in the direction of the magnetic field lines.* This result can be used further. If you have a loop of wire in which the current is flowing in a clockwise direction, while lying on a table, the magnetic field will be coming out of the table in the center of the loop and going into the table outside of the loop. By adding additional loops in which the current is going in the same direction, the strength of the magentic filed will increase. If we then add a bar of soft iron to the center of the coil formed by our loops, we will magnetize the iron and have an electromagnet in which the strength of the magnetic field will be much stronger than that caused by either the iron or the coil by themselves.

Prior to Oersted's discovery, it had been thought that no relationship existed between electricity and magnetism. After his discovery, a number of investigators returned to the attempt to produce either an electrical charge or an electrical current by means of a magnetic field. Most of the attempts were doomed to failure, because the investigators were overlooking one basic factor. They considered that the magnetic field was being induced by a *constant* current flowing through the wire - which it was - without considering what was implied by having a constant current. Thus, when they used stronger and stronger magnets in the vicinity of wires, they were unable to induce any current in the wires.



Almost simultaneously (and without each other's knowledge) two men, Joseph Henry (1797-1878) in the United States and Michael Faraday in England, discovered the method by which a magnet can induce a current to flow in a wire. Henry did his work in 1830, but he did not publish his findings until later. Faraday worked a year later and managed to publish his results first. For this reason, Faraday has been given almost exclusive credit for the discovery of the process of *electromagnetic induction* as it is now known. Basically, the process realized that a constant current implied a flow of electrons so that the electric field due to any particular electron in the region of the wire was changing. Thus, it could be said that *the magnetic field produced was caused by a changing electric field.* Using this information, it might be expected that the converse might be true. Both Henry and Faraday worked on the problem from this angle and they found that *an electric current can be induced in a wire if there is a changing magnetic field in the region of the wire.*

BRINGING EVERYTHING TOGETHER The discovery of these two relationships between electricity and magnetism have formed the basis of almost all of our present electrical devices. After these discoveries were made, a large number of physicists throughout the world began to study these effects and implications that they had on man's knowledge. We shall pass over most of these efforts and mention the work of only one more man. James Clerk Maxwell (1831-1879) has emerged as the most celebrated scientist in the field of electricity and magnetism and their relationship to each other. Almost every other accomplishment in science is attributed to different people in different countries of the world. That is, a particular physical law is known as A's Law in one or several countries, but known as B's Law somewhere else. Maxwell has the enviable position of being known as the discoverer of his equations throughout the entire world. Such was the originality and value of his work in the field of electricity and magnetism.

Maxwell's accomplishment seems almost too simple and obvious after he had made it. He managed to take all of man's knowledge of electricity and magnetism and their relationship to each other and express it in terms of four single, simple equations! That he was able to do this in a manner that could at once contain everything and in such a simple form is a true measure of his genius. The equations are given to illustrate their simplicity, the mathematical operations involved, however, are more complicated than we have time to explain:

$$(1)\ div\ \vec{E} = 0$$

$$(2)\ div\ \vec{H} = 0$$

$$(3)\ curl\ \vec{E} = -\frac{1}{c}\frac{\partial \vec{H}}{\partial t}$$

$$(4)\ curl\ \vec{H} = \frac{1}{c}\frac{\partial \vec{E}}{\partial t}$$

Equation 1 states that electric field lines can be neither created nor destroyed in the absence of electrical charges. It also can be changed into another form to give Ohm's Law. Equation 2 does the same for magnetic field lines, but goes further in that it also states that there can be no separate magnetic poles (that is, "magnetic charges" cannot exist). Equation 3 says that a changing magnetic field will produce an electric field (induce an electric current) in the proper direction.

Equation *4* says that a changing electric field (an electric current) will produce a magnetic field in the proper direction. It has been said many times before and worth repeating here for emphasis of the value of Maxwell's work that *if you understand Maxwell's Equations and how they apply to a situation, you will be able to derive almost everything there is to know about electricity and magnetism.*

There is one final feature that was predicted from Maxwell's equations. In order to make the equations balance so that the size of the effects produced correspond properly in his equations, Maxwell had to introduce some constants that referred to the particular material in which the changing electric or magnetic fields were located. These constants appear in one form in Equations *3* and *4* as the letter c. These constants have units of velocity and tell how rapidly a change in one area of the material will become apparent in another area. *The velocity found is the velocity of light in the material.*

Notice that, if there is a change in the electric field in an area, a magnetic field will be produced near the area. But, a changing a magnetic field being produced would create an electric field in turn. This process is self-perpetuating and results in a wave that moves outward from the source at a speed corresponding to the speed of light. Electromagnetic waves that are produced in this manner have made possible radio and television signals. Furthermore, as a direct result of this process, we have been able to discover that light itself is nothing more than an electromagnetic wave of a particular frequency range.

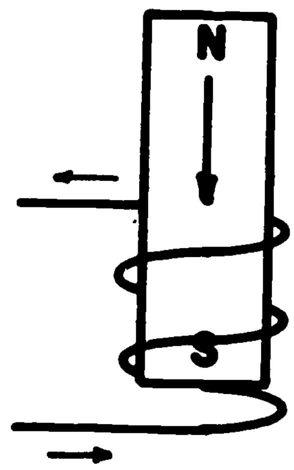
ELECTRICAL CIRCUITS In this section, we shall examine some devices that make use of the properties of changing electric and/or magnetic fields and see how they operate. To do this, we will make use of the three following rules: the Left-Hand Rule for the direction of the magnetic field produced by a current; the rule that electric or magnetic field lines tend to repel each other when they are parallel; and a variation of the Law of Conservation of Energy. It might be well at this point to summarize these briefly before we go further -

1. The magnetic field produced by a current flowing through a wire is formed around the wire in such a way that if the left thumb is pointed in the direction of the current, the left finger will wrap around the wire in the direction of the magnetic field.

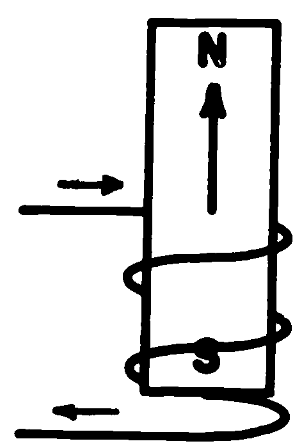
2. If two field lines of the same type (either or electric or magnetic) are parallel, they will repel each other if they are in the same direction and attract each other if they are in the opposite direction.

3. Whenever a change in one type of field produces the other type of field, the field that is produced will change in such a way as to *oppose* the change in the field that is producing it.

The third rule may seem a bit complicated at first, so an example may help. Let us go back and look at the magnetic field that we produce by running a current through a coil of wire. We saw that a current flowing through the wire produces a magnetic field in a particular direction inside of the coil which depends upon the direction of the current and the direction in which the coil was wound. Before the current started to flow, there was no magnetic field in the wire. After it was flowing, there was a magnetic field. Therefore, when the current first started to flow, the magnetic field had to be changing from zero to some other value. Our third rule states that the way that the magnetic field changed had to be in the direction that would try to *keep the electric current form flowing*. Likewise, when the current was turned off, the magnetic field was changing in the opposite direction (from being present to not being present), so the change in the field then would have to be in the direction that would try to keep the current flowing.



We can test this rule experimentally by passing a current through a coil of wire and seeing the direction of both the current and the magnetic field when the current is flowing. Then, by disconnecting the source of the current and replacing it with a meter that will show currents flowing in either direction, we can change the magnetic field inside the coil by pushing a permanent magnet into the coil whose field is in the same direction as the original field when the current was flowing. When the magnet is pushed in, the current in the coil will be induced to flow in the opposite direction as the original current. When the magnet is removed the current will be induced in the same direction as the original current. (These induced currents will also be setting up their own magnetic fields to oppose the motion of the magnet.) If the opposite were true, the magnetic field produced by the current would have to induce a current in the same direction as the original current, which would make the magnetic field even stronger, which would



make the current even larger, with the result that we would be creating energy. This is impossible according to the Law of Conservation of Energy.

The use of coils with their property of producing a variety of effects due to changing electric and magnetic fields make them an important circuit element in electrical and electronic circuits. For this reason, we will look at some of the things that we can use coils for, and where we can expect to find them in our every day life. First though, we should realize the following properties of coils. If one loop of wire wound around a piece of soft iron will produce a given magnetic field, then two loops will produce twice the effect, three loops will produce three times the original effect, and so on. *The number of turns in a coil will be proportional to the size of the magnetic field in the coil, all other things being equal.* Similarly, if a changing magnetic field in the vicinity of a single loop of wire will produce an emf or potential difference to cause a current to flow through the loop, then two loops in the same region will produce twice the potential difference. *The number of turns in a coil will be proportional to the size of the potential difference between the ends of the coil, all other things being equal.* Let us now see how these two properties can help us in the transfer of energy from one place to another.

Suppose we have only a single coil that we can connect to a circuit. If we start a current flowing through the coil, by our third rule the magnetic field produced will oppose the flow of the current so that the coil of wire will appear to be a resistance when the electric current through the wire is changing. When the current has reached a constant value, the magnetic field will be a constant and the resistance will become zero due to this effect. When the current stops, the resistance due to the magnetic field will try to keep the current flowing through the coil. So far, this is merely a restatement of our example to see how the third rule might be applied and to explain what was meant by it. However, we can use this result to come to some conclusions about any coil that was first discovered by Joseph Henry. For this reason the size of the resistance found in these situations is measured in units called *henrys*. The conclusions are that:

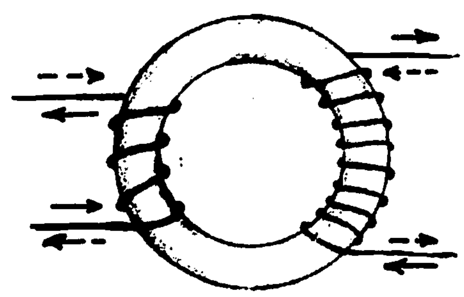
1. Any coil will develop a resistance to a change in current flowing through it, regardless of what the previous current was or how it is changing. That is, we can use a coil to try to maintain a constant current no matter what current we have flowing through the coil.

2. The larger the coil the larger the resistance.

3. The faster the changes in current, the larger the resistance.

Let us consider the third conclusion. If the current is changing rapidly back and forth (larger and smaller and back again), the coil could be used to "smooth" out the current and keep it flowing at a constant rate. It may thus be used in an electrical circuit to keep a changing current on one side of the coil in the circuit while allowing a constant current to flow from the part of the circuit on one side of the coil to the part on the other side without interference. This effect on currents that coils have is called *self-inductance or merely inductance*. It is measured in units of henrys, as we mentioned, that describe the amount of resistance present due to changing currents.

THE VOLTAGE TRANSFORMER If we have two coils placed near each other, the change in the magnetic field of one will produce a current in the other that will correspond to this changing magnetic field and oppose it. This effect that one coil has upon another is called *mutual inductance* for similar reasons. The mutual inductance between two coils will change as the coils are moved closer together, further away, or if a loop of soft iron passes through each of the coils. Recall that the magnetic effect of a current passing through a coil of wire could be made much larger by having a piece of soft iron in the center of the coil. By having the soft iron pass through the centers of both of the coils and be in the form of a closed loop, the mutual inductance can be even stronger than if it were not present or if it were in only one of the coils. Coils can be used for a wide variety of applications when they are arranged in this manner.



If we have two coils placed near each other and pass a current through one, the changing magnetic field will produce a current in the second and, according to our third rule, the current in the second will be in such a direction as to produce a magnetic field that will oppose the magnetic field due to the first. This rule, therefore, will allow us to determine the direction of the current on the second coil that is being produced by the current in the first. The size of the current and/or the potential difference across the ends of the second coil will be determined by the mutual-inductance of the combination of the two coils and the ratio of the number of turns in the first and second coils. Let us assume for the rest of the discussion that the mutual inductance between the two coils is perfect. That is, that the change in the magnetic field of one of the coils will result in a change in the magnetic field in the other coil that is exactly the same size. (This will, of course, actually never happen, but it makes the discussion easier to follow). Then, the effect that will be seen at the ends of the coils where we can make our connections will depend only upon the number of turns in each of the coils. After taking into consideration the number of turns in each of the coils, we will find that the ratio of the potential difference causing the current to start flowing and of the potential difference that we measure across the ends of the other coil is directly proportional to the ratio of the number of turns in the corresponding coils or

$$\frac{N_1}{N_2} = \frac{V_1}{V_2}$$

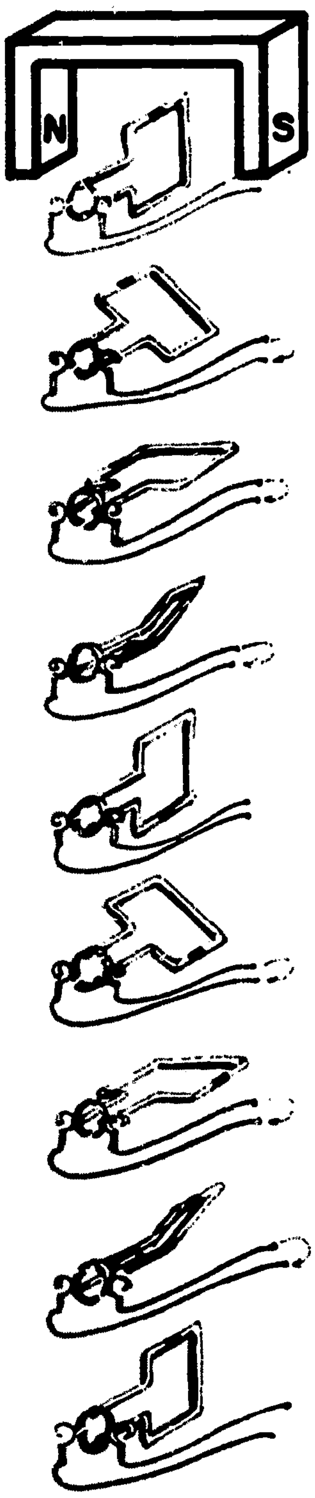
where N is the number of turns and V is the potential difference or voltage. A system of two coils that is used in this fashion is called a *transformer* because it can be used to transform voltages from one value to another.

A COMMON USE An everyday example of this effect is found in automobiles. Most cars today have electrical systems that are operated by 12 volt batteries, but a potential difference of about 10,000 volts is needed to make the spark at the spark plugs in order to ignite the mixture of gasoline and air in the cylinder to make the engine run. This spark is caused by two devices. First, there is the ignition coil, which is merely a form of transformer that has a ratio of turns of about 1,000 to 1 to raise the potential difference from the 12 volts of the battery to about 12,000 volts when the current is interrupted. Second, there is the distributor which successively makes electrical connections to each of the cylinders. When the distributor makes contact to one of the sparkplugs, there is a change in the current in the coil and the change produces a potential difference of 12,000 volts or so that causes the spark in the cylinder.

Thus far, we have dealt only with currents that are moving in a single direction, or D. C. currents. However, as we have seen, the use of coils in the form of transformers generates an entirely different type of current that moves first in one direction and then in the other. This occurs because the current produced in the second coil (or *secondary* of the transformer) will flow first in one direction when the current starts to flow in the first coil (or *primary* of the transformer) and then flows in the opposite direction in the secondary when the current stops flowing in the primary. For this reason, and because it avoids the need of having a device to start and stop the flow of the current to the primary coil, we make use of the type of current that we call *alternating current* or *A. C.* These currents may be obtained by having someone operate a reversing switch to make the current flow first in one direction and then in the other and back again, but it is much easier to produce currents that behave in this manner by means of an electrical generator. In the United States, currents that are produced commercially in this manner vary at the rate of 120 changes per second or 60 complete cycles, so that they are referred to as 60 cycle currents. In Europe, the change is often 100 times per second or 50 complete cycles. The more rapidly the current changes direction, the more efficient a particular transformer will be, so aircraft often have 400 cycle per second electrical systems in order to reduee the weight of the transformer thay they use.

One of the biggest uses of transformers that immediately affects our daily lives is in the area of power transmission. There is often great distances between the point where the commercial power is generated and the point where it is used. If the power were transmitted at the same voltage from the generator to the consumer, there would be a great deal of it "lost" in heating up the transmission wires and there would be very little electrical energy to reach the consumer. However, if the voltage is raised from 110 volts up to 110,000 volts, most of the energy will be transmitted to the consumer because the increase in voltage lowers the current in the wires which causes the heating effect. This is done by means of transformers. At the generating stations, there are large transformers to raise the voltage; outside of cities, there are other transformers to lower the voltage to values that are safer for populated areas (down to 2200 volts) and then on the poles along the streets, there are futher transformers to lower the voltage still futher for even greater safety (to 220 volts for stoves, washing machines, etc. and 110 volts for other household uses). When Edison first introduced commercial electrical power to New York City, he used direct current, later he had to change to alternating current to be able to raise the voltage for transmitting, as the loss of electrical energy in the wires transmitting the energy was too great.

Let us turn now to the operation of motors and generators. Each of these devices consisted of a magnet that is permanently fixed and a rotating coil of wire called an *armature*. The armature is connected to the rest of the circuit by means of a commutator which is a sliding contact that allows current to flow from or to the armature to the circuit. The commutator consists most simply of a piece of metal that rubs against the metal parts of the armature. If the armature is moved by mechanical energy so that energy is transformed into electrical energy to be used elsewhere, the device is called a generator; if the energy is supplied by an electrical current that is used to turn the armature and transform the energy into mechanical energy, the device is called a motor. That is the only basic difference between the two devices. Often, there is a slight difference in the design of a generator as opposed to that for a motor. and vice versa, in order to make it function more efficiently at the task that is set for it, but there is practically no difference in their actual operation. To reduce costs, the motors used on electric busses, trolly cars, and subways are often designed to act as generators when the brakes are applied both to help stop the vehicle and to return some of the electrical energy that had been supplied to it back to the circuit from which it was obtained.



Let us examine motors more closely and then, by means of the third rule that we have established, we can see how a generator would also be explained by means of the same device. The simplest motor would be a D.C. motor that consisted of a permanent magnet and a rotating coil between the poles of the magnet. If the coil is horizontal when the switch is closed to allow the current to flow, one of the ends of the coil would become a north pole and the other a south pole. The north pole would be attracted to the south of the magnet and repelled by the north pole. The south end of the coil would react in the opposite manner. This would cause the coil to rotate to a vertical position. Just as the coil reached the vertical position, the commutator would switch the current in the coil and reverse the magnetic field of the coil. The coil itself is moving so that it would go a little bit past vertical due to its own motion. Now, each end of the coil would be near a pole of the magnet that repelled it, so it would be forced to move in the same direction that it had originally until the coil was again vertical but at the opposite ends whereupon, the current would again be switched by the commutator. Instead of a permanent magnet, we could have used a second coil of wire that always had a current flowing in the same direction.

Suppose now, that instead of a motor, we wanted to make a generator out of this device. We would need to have the permanent magnet in this case to supply the magnetic field that was needed. As the ends of the doil moved from one pole to the other, the magnetic field would be changing in the vicinity of the coil. This would mean that a current was being generated in the coil. The direction of the current would be just the opposite of the current that was required to make the device operate as a motor.

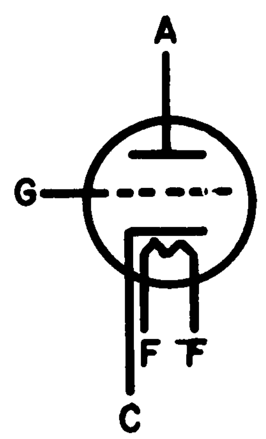
Most A.C. motors have to operate at a fixed speed (as opposed to the D.C.motors whose speed depends upon the voltage at which the motor is being operated). This is because the rotating portion of the motor, the armature, is always connected to the same wire leading to it and the change of direction of the current in the wires is used to change the direction of the magnetic field in the armature to make the armature rotate in the fixed field of the magnet. Thus, most of the A.C. motors that you will encounter have a speed rating of 3600 rpm which is the same as 60 revolutions or cycles per second. (This is why an electric razor will not work as well in Europe as it will in the United States due to the difference in the frequency of the commercial power between the two areas. The razor or other motor will operate at 1/6th slower speed there than it does here). A.C. generators are designed in the same manner as A.C. motors and their operation is similar to that of D. C. generators with the difference being in the type of electrical current that is being produced by them.

There is one further device that should be discussed. It is called a capacitor. Recall the manner in which a coil of wire will try to maintain a *constant current* through it and can be used to "smooth" out rapid changes in current in an electrical circuit. A capacitor operates in a similar manner, but it trys to maintain a *constant voltage* itself and to filter out any rapid changes in voltage. It consists of two metal plates that are placed parallel to each other. (Often these plates are in the form of metal foil with an insulator between them and the whole device is wrapped into a roll to save space). If there is an electrical charge on one of the plates, there will be an equal and opposite charge on the other plate as both of the materials are conductors and a current will be free to flow onto or off the plates. Once there is some kind of a charge on the plates, the force of attraction between the two plates will attempt to keep that charge on them due to the resulting field that is built up. The charge on either plate is most easily measured as being proportional to the potential difference between the two plates. Any attempt to either increase or decrease this potential difference would have to result in a change in the electric field between the two plates that would, in turn, result in a magnetic field that would be produced in such a way as to oppose the change in the original electric field. Rapidly changing electric currents would not be effected, however, as long as the potential difference across the capacitor did not change.

In summary:

> An inductor will maintain a constant current through it by allowing a changing voltage across it, while a capacitor will maintain a constant voltage across it by allowing a changing current through it.

Thus far, the devices that we have investigated have only reacted directly on the current or voltage that was applied to them and have done nothing to influence the current or voltage in any other way. For this reason, they are referred to as *passive circuit elements*. There are other circuit elements that act in an entirely different manner on the current passing through or the voltage applied across them. These other elements are called *active circuit elements*. This is because they behave in such a way that they can use one current or voltage to control another rather than merely reacting to the qualities of a single current or voltage. Active circuit devices are used extensively in the apparatuses that have been developed for communication in all of its various forms.



A - ANODE (POSITIVE)
C - CATHODE (NEGATIVE)
F - FILAMENT (HEATER)
G - GRID (SIGNAL)

The device that was developed first around the start of the twentieth century is the vacuum tube. Although vacuum tubes are becoming less and less common today through the development of other devices to replace them, it might be best to start with them, as they are easier to understand. Basically, a vacuum tube consists of three elements: a piece of metal that has a negative charge on it called the cathode; another with a positive charge called the anode or plate; and a wire mesh that is placed between them. All of these parts are enclosed in a container, usually glass, from which the air has been removed. Electrons flow off of the cathode (often due to heating the cathode to "boil off" the electrons) and flow toward the anode. In the process, they flow through the wire mesh which is called the "grid".

As long as the grid is kept neutrally-charged (no excess positive or negative charge), it will have no effect upon the flow of the electrons. The size of the current will then be determined only by the potential difference between the cathode and the plate. If, however, we place a charge on the grid, the flow of electrons will either increase or decrease according to whether the charge is positive or negative respectively. If the charge is positive, the electrons from the cathode will be accelerated towards the plate before they pass through the grid. If the charge is negative, they will slow down or be decelerated before they pass through the grid. Due to the design of the device, a very small charge on the grid can make a large change in the flow of the current through the tube. This ratio of the size of the controlling current going to the grid to the size of the current passing through the tube from the cathode to the anode or plate is referred to as the *amplification factor* of the tube or the circuit in which it is located. The amplification factor may be as high as several thousand, meaning that a signal of a given size placed on the grid will result in a change in the size of the signal passing through the device that is several thousand times as large. Most of the time, however, the size of the amplification factor is only two or three. This is because the larger the amplification factor, the greater the error in the signal that comes out, or the greater the distortion. By properly designing the device and the rest of the circuit in which it is located, it is possible to have a very large amplification of the signal with a very small amount of distortion. This is why a small, cheap radio does not sound nearly as good as a large hi-fi set. The amplification is made through a much larger number of stages than the small radio so that the distortion will be kept to a minimum. Amplifying devices of this type are used in all phases of communication.

The way that the grid of a vacuum tube acts to change the size of the current flowing through the tube can be thought of as being similar to a change of the resistance of the tube. In the late 1940's, scientists at the Bell Laboratories discovered a new type of device that they called the *transistor* that operates on a similar principle to that of the vacuum tube, but is able to do the work much more satisfactorily. In the vacuum tube, the elements can wear out so that distortion of the signal increases or the entire tube itself can simply stop working. By using the transistor, the change of the characteristics of the device through age are reduced and there is nothing to "burn out" as there is in a vacuum tube, so that the device is able to work for much longer periods more effectively. Basically, a transistor is a special type of crystal that has three parts of the simple vacuum tube. A large current passes between two of the connections that is controlled by the signal placed upon the third connection. Once again, the mechanism operates by having the third connection charge the resistance of the entire device as far as the larger current path is concerned. A further benefit of the transistor over the vacuum tube is that all of the parts are contained in a single solid crystal rather than being separate metal parts mounted in an enclosed evacuated tube of glass or other material. This is why transistors and other similar devices are often referred to as "solid-state" devices. Because of the solid, crystalline nature of the transistor, it is better able to withstand shocks than the standard vacuum tube.

The development of the active circuit elements has made possible all of the modern communication devices that we use so freely today, just as the development of the passive circuit devices has made possible the generation, transmission, and utilization of electrical energy from one place to another for heating, lighting, etc. The result of these developments in electricity and magnetism has been that the average person's life has changed and improved more in the past one hundred years than in all of man's history up to that time. Future developments in the use of electricity and magnetism may be expected to change our lives to at least as great a degree in the next one hundred years, as the end of these developments is not yet in sight.

SUGGESTED OUTSIDE READINGS:

F. Bitter, *Magnets: The Education of a Physicist*, Doubleday (Anchor), Garden City, New York, 1959.

D. G. Fink and D. M. Lutyens, *The Physics of Television*, Doubleday (Anchor), Garden City, New York, 1960.

R. Galambos, *Nerves and Muscles*, Doubleday (Anchor), Garden City, New York, 1962.

R. T. Lagemann, *Physical Science: Origins and Principles*, Little, Brown and Company, Boston, 1963.

D. K. C. MacDonald, *Faraday, Maxwell. and Kelvin*, Doubleday (Anchor), Garden City, New York, 1964.

J. R. Pierce, *Electrons and Waves: An Introduction to the Science of Electronics and Communications*, Doubleday (Anchor), Garden City, New York, 1964.

Project Physics, *An Introduction to Physics: Light and Electromagnetism*, Holt, Rhinehart and Winston, New York, 1968.

ARTICLES FROM *Scientific American:*

L. O. Barthold and H. G. Pfeiffer, "High Voltage Transmission" (May 1964).

K. W. Ford, "Magnetic Monopoles" (December 1963).

H. W. Lewis, "Ball Lightning" (March 1963).

G. de Santillana, "Alessandro Volta" (January 1965).

STUDY QUESTIONS:

1. Why were electrical phenomena of more interest to the ancients than magnetic phenomena?

2. Give some common everyday examples of static electrical phenomena.

3. What is meant by a "field" in science? Why is it a useful concept?

4. Contrast the "one fluid" and "two fluid" theories of electricity.

5. Give modern views on the production of static electricity.

6. Discuss the similarities and differences between Newton's law of gravitation and Coulomb's law of electrostatic force.

7. If the charge on two spheres is doubled, by what factor does the force between them change?

8. If the distance between two charged spheres is doubled, by what factor does the force between them change?

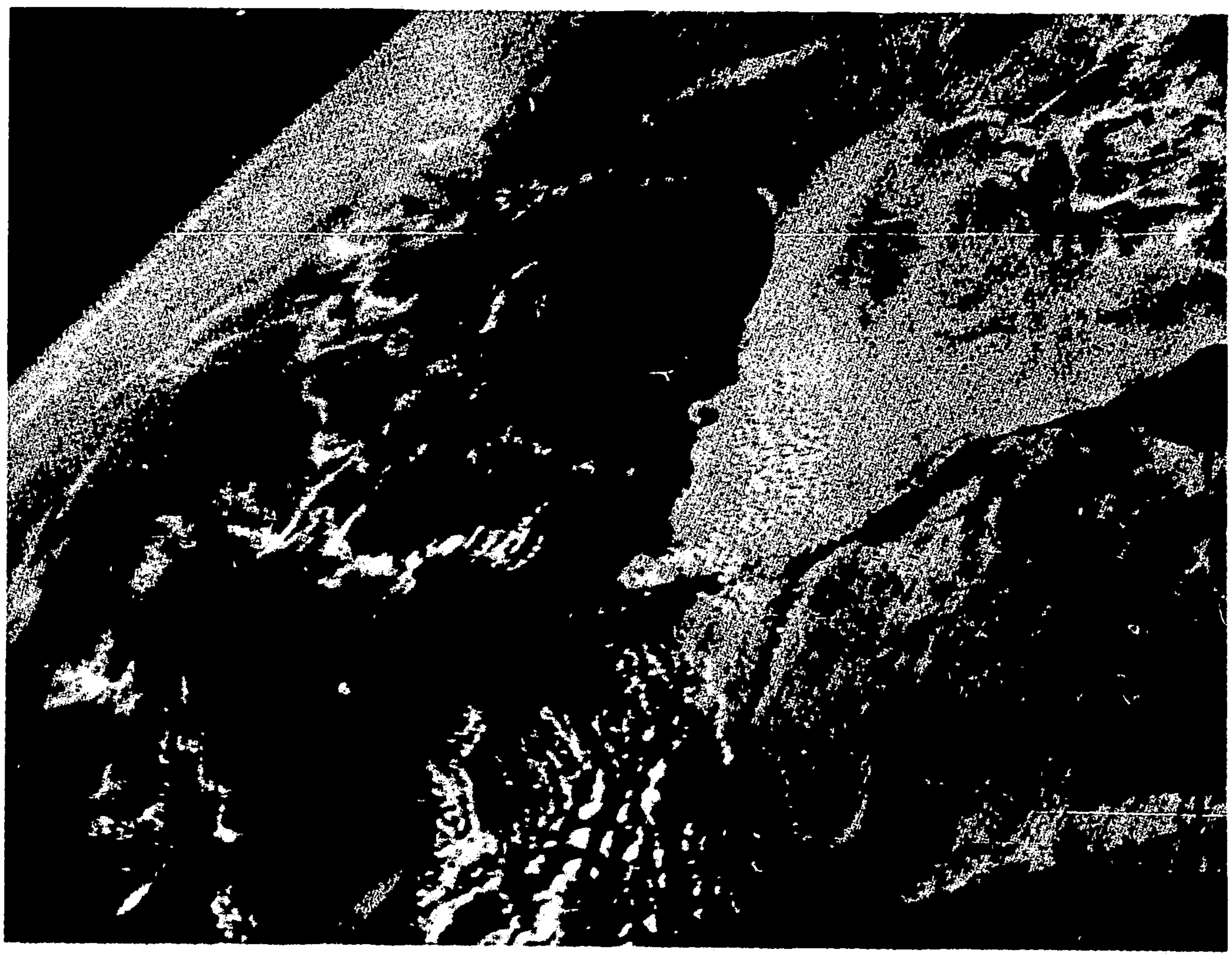
9. State Faraday's law of electromagnetic induction.

10. In what respects are electrical motors and generators similar?

11. For a typical day in your life, list in order from morning to night all of your activities that in any way involve the use of electricity

CHAPTER 8 THE EARTH IN TIME AND SPACE

PAGE



"IN THE BEGINNING..." Thus begins the Biblical account of the story of creation in the first chapter of *Genesis*. This account of the creation of the earth was accepted as literally true by most people until the early 1800's when the science of geology, or the study of the earth, came into being. Prior to this, a few bold thinkers such as Herodotus, the famous Greek historian of the fifth century B.C., Aristotle, and Leonardo da Vinci attempted to explain the appearance of fossil deposits and certain geological structures, but they were ahead of their time. Man has always been more intrigued by the moon and stars than by the earth itself. Perhaps, this was because the heavenly bodies were remote and more of a mystery than the earth upon which he dwelt. Later, the power and authority of the Church branded as heresy any attempt to explain the origin of the earth in any way that conflicted with the Holy Scriptures. According to Bishop Usher, who did a study of Biblical chronology in 1650-54, the creation of the earth took place on a Tuesday morning, 4004 B.C. Many believed that the many fossils that were periodically unearthed by archeologists were remains of plant and animal life that perished in the cataclysmic Noachian deluge. Still, others believed that these were remnants of an earlier age before man existed on the earth.

Today, a great many Biblical scholars do not insist upon a literal interpretation of the story of creation as related in *Genesis*, but they believe that the writer of Genesis was attempting to show in his own way how God is working through history to reveal Himself to man. As Galileo remarked, "The Holy Spirit intended to teach us how to go to heaven, not how the heavens go." What then, can geology teach us about the origin of the earth and its place in the scheme of things in the universe? The study of the earth's structure, its chemical composition, and changes that have occurred in its form in the past can all teach us something about the way things were before man existed on the earth if we take the time to investigate them. Unfortunately, the record written in the earth is not as clear as one would like. Over the centuries and millenia, forces of nature have been hard at work in an attempt to erase the record of the earth's birth through the erosive agents of rain, wind, and water. Changes in the earth's crust, or outer skin, have uplifted mountains from shallow seas, worn them down, and uplifted them time and time again. Thus, it is no easy task to decipher the events following the earth's birth. This is one reason why scientists look forward to the return of the Apollo Mission and its sample of lunar rocks. They hope to learn something about the early history of the Solar System through the study of their composition. Unlike the earth, the moon has not had the weathering agents present on its surface to erase the birth record. This is why some of the future astronauts who explore the lunar surface will be trained geologists. But even a comprehensive study of the moon will not tell us all we would like to know about the earth. We must study the earth for that. First, let us look at some of the contributions of scientists in the past as they sought to uncover the secrets buried in the earth's crust.

Among the ancient Greeks, Herodotus, Eratosthenes, and Aristotle recognized that deposits of sea shells and salt in various locations on the earth indicated that these areas once were inundated by the sea. The Roman scholar Strabo (54 B.C.-25 A.D.) wrote a comprehensive work on earthquakes, volcanism, and fossils, in which he acknowledged a similar view. These authors were not taken seriously because of the Judeo-Christian traditional account. Little progress was made in the science of geology until the Renaissance when the scientific method was applied to the origin of fossils. Descartes believed that layers of rock that appeared at various places on the earth's surface were indicative of dislocations of the crust and postulated that the earth once was in a molten state. Perhaps the most outstanding geological scholar of the seventeenth century was Nicolaus Steno (1638-86) who advanced three laws explaining geological stratigraphy. These were the *Law of Superposition* which states: *all strata are in age order with the oldest order on the bottom*; the *Law of Original Continuity*: *all strata were originally continuous in nature and any breaks were due to displacement or erosion*; and the *Law of Original Horizontality: all strata were originally horizontal and any change is due to a tilt.*

Moving on to the eighteenth century, three scientists that contributed significantly to the development of geology were Buffon, Werner, and Hutton. The latter two men were founders of two competing schools of geological thought.

Buffon published two volumes on the history of the earth entitled *Natural History* (1749) and *Epoques de la Nature* (1778), in which he attempted to trace the history of the earth from the time of creation to his time. In these volumes he compared the six days of creation as related in *Genesis* to six epochs of indeterminate duration. He advanced the theory that the earth was formed from a hot mass that detached from the sun and gradually cooled during the six epochs. In the First Epoch, the molten earth cooled and a crust slowly formed over the liquid center. In the Second Epoch, surface features such as mountains formed, while water existed only in a vaporous state. During the Third Epoch water began to condense into seas.

The Fourth Epoch saw the emergence of land while the waters sank into hot cavities to form volcanos. During the Fifth Epoch animal life came into being. In the Sixth Epoch the continents separated and some land areas submerged.

During most of the eighteenth century geologists were divided into two competing schools of thought concerning the origin of rocks. One school, known as the Neptunists, after Neptune, god of the sea, was founded by A. G. Werner (1750-1817) who believed that all rocks were formed by precipitation from the primeval ocean that surrounded the earth. Werner arranged the order of precipitation to account for the deposit of the oldest rocks first, such as granites, followed by slates, basalt, limestones, and sandstones.

The opposing school was known as the Vulcanists, or Plutonists, after the gods of fire and of the underworld, because they believed that most rocks, such as granites and basalts, were igneous or formed by interior fires within the earth. These schools probably had an influence on the French writer Jules Verne, who wrote the story *Voyage to the Center of the Earth* in 1864. The most well known member of the Vulcanist school was a Scotsman named James Hutton (1726-97), who advanced the Principle of Uniformity, which states that natural forces have been operating to change the surface of the earth with uniformity in the past and continue to do so at present. This concept contradicted the Calaclysmic Theory which supported the Biblical tradition.

The study of geology also attracted the great German poet Goethe to its ranks from time to time. He founded and named the branch of geology that deals with the systematic study of rock formation and transformation - morphology.

In the nineteenth century, geology came into its own as a science. The contributers to this were many, but one, Charles Lyell, is of particular interest because he developed the systematic nomenclature and geological definitions utilized in geology today, besides contributing greatly to our wealth of knowledge. Lyell stands out as the first true geologist. The following definitions and terminology are listed for your edification:

Stratigraphy - the study of rock layers and by the deductive reasoning therefrom obtaining the geological history of a region.

Fossils - An impression of any form of animal or plant life, such as a footprint or complete body form preserved in rock. Fossils are characteristic of the strata in which they are found, and therefore one of the major tools used by the stratigrapher in the study of the earth's past.

Sedimentary rocks - rocks formed in layers at the bottom of bodies of water, e.g., limestone, sandstone, quartz, mica, and silica.

Igneous rocks - solidified molten interior material extruded to fill earthquake cracks or blown up to the surface by volcanic action, e.g., basalt, granite, and lava.

Metamorphic rocks - sedimentary rocks recrystallized without fusion, e.g., slate, schist, and gneiss.

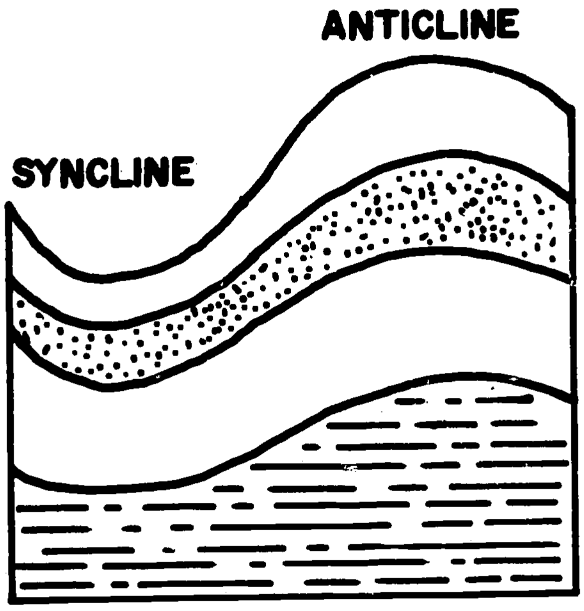
Anticlines - large convex folds in the crust due to diastrophic activity accompanying the rise of liquid rock magma deforming the sedimentary rock layers above it.

Synclines - large concave folds in the crust formed at the same time and under the same conditions as anticlines.

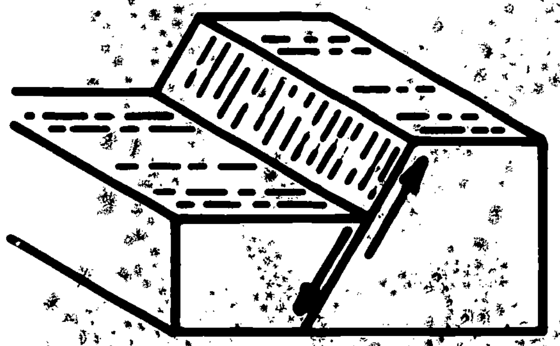
Batholith - a large mass of igneous rock intruded into the surface layers, sometimes being extruded into a crack or fissure to form a dike.

Fault - a large fracture in the rock strata as in the case of a normal fault (up-down slippage), thrust fault (one layer pushed over another) and slip-strike fault (horizontal slippage). Faultage is slow and discontinuous, occurring over thousands of years. One of the most famous is the San Andreas fault crossing through California. Earthquakes usually occur along faults.

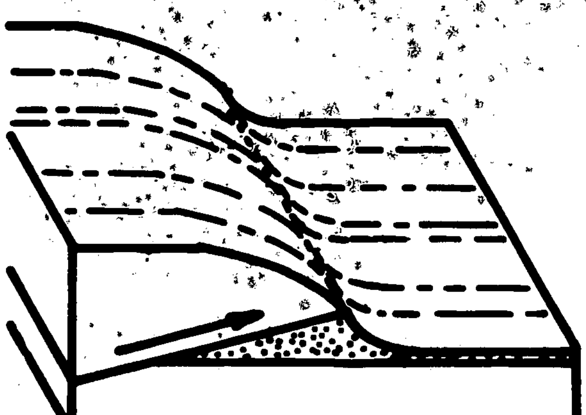
In the twentieth century, the major contributions to geology have been of a technological nature: improvements in dating of fossils by analysis of Carbon-14 half-life data, better analysis of specimens by x-ray diffraction techniques, and location of mineral deposits by radiation detection and airborne gravimetric measurements. A practical application of the space age is the use of earth satellites to photograph and map the earth for future mineral exploration. Geologists can locate deposits by comparing satellite photographs of certain areas with photographs of areas that have known deposits.



STRIKE-SLIP FAULT



NORMAL FAULT



THRUST FAULT

THE ORIGIN OF THE EARTH What have we learned from the science of geology concerning the origin of the earth? Today most scientists believe that the earth came into existence approximately 4.5 billion years ago. This estimate is arrived at through a study of the relative abundance of certain radioactive elements present in rocks. The formation of the earth begins with the formation of the Solar System itself probably about six billion years ago. A cloud of hydrogen gas and dust began to contract under the action of gravity into a *protosun* of roughly spherical shape. Perhaps due to rapid rotation, some of the gas spun off to form the *protoplanets*. As the protoplanet that became the earth contracted under the pull of gravity, the heavy elements accumulated at its center and its interior temperature increased.

The protosun continued to contract until its temperature and density was high enough to initiate a thermonuclear reaction, converting the hydrogen into helium and radiating away the energy into space. Meanwhile, the planets continued to contract and increase their interior temperatures, partly due to gravitational contraction and partly due to radioactivity. As the earth cooled by radiating away energy, its crust began to form, and gradually hydrogen and oxygen in its atmosphere combined to form water which fell to earth as rain, to be converted back to steam by the hot crust. This process probably continued for millions of years as the earth gradually cooled. Eventually the earth cooled sufficiently to allow the rain to remain and the first primeval oceans were formed. The rains turned into rivers and through the eons washed away the land and minerals into the steaming primeval seas. Under the constant bombardment of radiation from the sun and stars, and lightning, these substances in the seas were converted into more complex molecules and chemicals, and finally into simple organic molecules. As the eons passed these molecules combined into more complex structures and eventually formed the first living things. As these creatures died they fell to the ocean floor where they were eventually covered by sand and debris. As life continued to evolve, hard shell animals appeared in the sea, and as they died they drifted to the ocean floor to be covered up by deposits. These are the remnants of fossils that we have uncovered today. In Figure 8.1, is a diagram indicating the various geological ages of the earth, which shows how far geologists have been able to extrapolate into the past and reconstruct the primeval history of the earth.

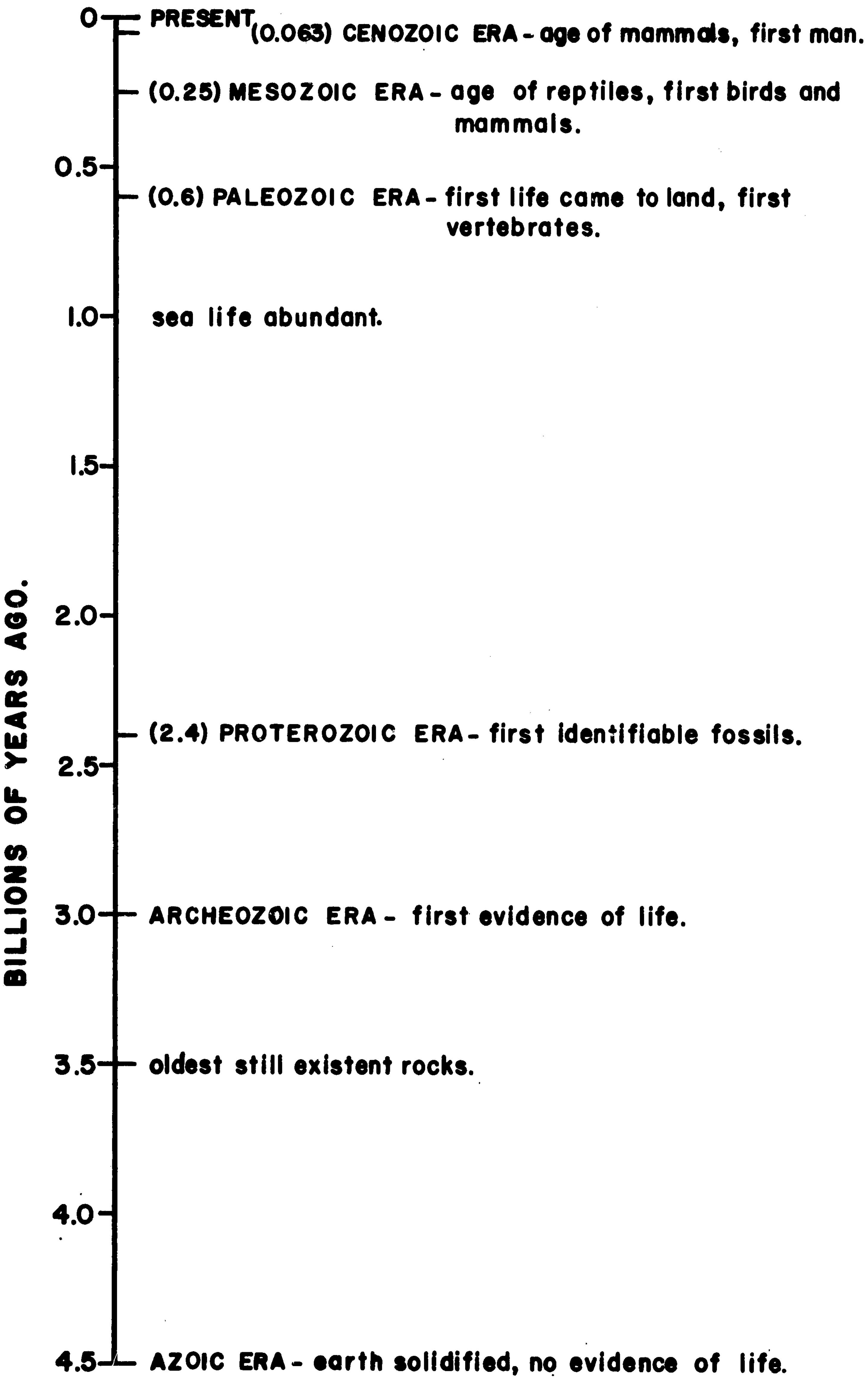

BILLIONS OF YEARS AGO.
0 — PRESENT (0.063) CENOZOIC ERA - age of mammals, first man.
(0.25) MESOZOIC ERA - age of reptiles, first birds and mammals.
0.5
(0.6) PALEOZOIC ERA - first life came to land, first vertebrates.
1.0 — sea life abundant.
1.5
2.0
(2.4) PROTEROZOIC ERA - first identifiable fossils.
2.5
3.0 — ARCHEOZOIC ERA - first evidence of life.
3.5 — oldest still existent rocks.
4.0
4.5 — AZOIC ERA - earth solidified, no evidence of life.

Fig. 8.1

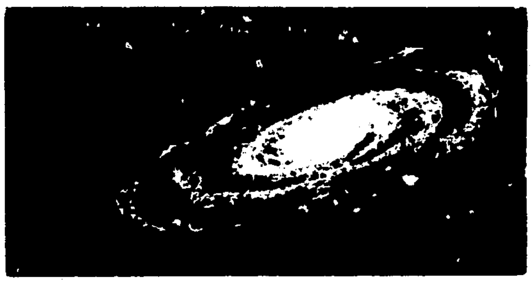
THE ORIGIN OF THE UNIVERSE Man has always been curious about the origin of the universe, and from time to time various peoples have developed cosmological theories to explain it. The ancient peoples of Mesopotamia, notably the Babylonians and Semitic tribes, developed a cosmology which pictured the world as a disk-shaped earth completely surrounded by a moat of sea beyond which an inverted bowl formed the vault of heaven. The vault of heaven, or the firmament as it is referred to in the *Book of Genesis*, divided the waters of the heavens (rain) from the waters below. The celestial bodies moved across the firmament daily from the east and set in the west. Beneath the earth existed the realm of Hades, or the abode of the dead.

The ancient Babylonians believed that the sun, moon, planets, and stars were deities and it was the task of the priest-astrologers to observe their movements to ascertain their intent toward men. These astrologers began naming certain constellations after gods and creatures that they believed inhabited the heavens as early as 3000 B.C. Later, this practice was transferred to the Greeks who renamed some of these after characters from their own mythology and most of these have come down to us today in such names as Gemini the Twins, Taurus the Bull, and Leo the Lion.

The ancient Egyptians pictured the universe populated with gods and goddesses. To them the sky represented a celestial Nile, navigated daily by the sun-god, Ra, who at night rested in Hades. They explained solar eclipses as Ra being attacked by a great serpent. Similar explanations are found in other civilizations, such as the Chinese.

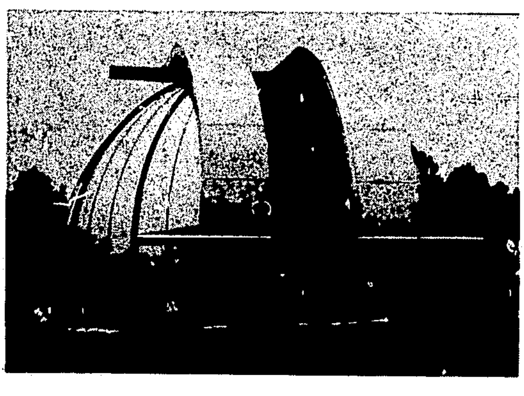
The universe of the Hindu religion was one of vast size and duration. It was pictured as passing through continuous cycles of development and decline. The cycle was known as *kalpa*, or "day of Brahma", and lasts 4.32 billion years. The cycle begins with the god Vishnu asleep upon a cobra, Shesha, who symbolizes endless time and who rests on the cosmic sea. A lotus blossom springs from Vishnu's navel and from this is born the god Brahma who created the universe for Vishnu to control during the remainder of the kalpa. At the end of this period, the universe is destroyed by the god Shiva, and is assimilated into Vishnu's body. Vishnu sleeps for a kalpa, and after this the cycle is repeated. After 100 kalpas, Vishnu's personality is absorbed into the one World Spirit of the Brahman. After an indeterminate period of time, a new Vishnu reappears and the cycle is repeated ad infinitum.

We have already in the first section of the text discussed the Greek and medieval cosmologies, so we will move on to the present to review the modern theories of cosmology.



Today, most scientists believe that the earth is a typical planet of a typical star belonging to a typical spiral galaxy which is only one of billions of such galaxies in the universe. Our own star, the sun, belongs to the Milky Way Galaxy, which is a system of approximately 100 billion stars, comprising a spiral galaxy approximately 100,000 light-years across and about 10,000 light-years thick. Our sun occupies a position about 30,000 light-years from the galactic center, or nucleus. The nearest star to our sun is approximately 4 light-years away.

Today's modern optical and radio telescopes have been able to "see" galaxies at a distance of approximately 10 billion light-years, which is believed to be the radius of the observable universe. A discovery of modern astronomy which allows one to make this estimate was the *red shift* observed in spectral lines of distant galaxies. Most astronomers believe this to be due to galaxies receding away from us at tremendous speeds. One obtains an analogous shift in sound frequency whenever a fire siren or train whistle passes by an observer. This red shift is interpreted to indicate an expanding universe. Everywhere one looks into space galaxies appear to be moving away from us with speed increasing with distance. This fact has led some scientists to develop what is generally referred to as the "Big Bang" cosmological model of the universe. According to this model, based on the observed rate of expansion, the universe must have originally occupied a small volume in space, sometimes referred to as the *primeval atom*, consisting of closely-packed neutrons. For some unknown reason this atom exploded and the neutrons decayed into protons and electrons. The expansion that we observe today is the result of this original explosion. From the protons and electrons, atoms of hydrogen formed and these later condensed into galactic clouds under the attraction of gravity. Further contraction led to the birth of stars within these clouds. Today, some astronomers believe they are observing such processes going on in certain dust clouds.



Scientists cannot at present decide between a "one shot" "Big Bang" model or a cyclic "Big Bang" model, due to inconclusive data. While some maintain that the universe began with one explosion, others think that this may have occurred over and over. If the explosion that caused the expansion was not violent enough, then the expansion might eventually slow down and the universe may begin to contract under gravitational influences until it is compressed into the primordial atom again where it may again explode and expand. This cycle may be repeated indefinitely. Does this sound familar? Scientists have calculated on present knowledge that such a cycle would require approximately 30 billion years to complete, so cheer up, we have 20 billion more to go!

Still, other scientists, notably Hoyle, Gold, and Bondi, have proposed a "steady state" model. This model asserts that the universe was always the way it appears now, and that there is no beginning and no end. This theory is based on the uniformity of the universe, or the fact that it appears the same in all directions. To account for the observed expansion, it was proposed that as galaxies move away from each other new galaxies are created in space. The energy of expansion is believed to be converted into matter at the rate of approximately one hydrogen atom per century to account for the appearance of new matter in the universe. Before we can decide between these alternative models or some new model, we have to make better observations and experiments.

QUASARS, PULSARS, BLACK HOLES, AND ALL THAT In recent years astronomers have discovered star-like objects in space that emit enormous amounts of energy in the radio-frequency region of the electromagnetic spectrum. These are called *quasi-stellar* radio sources or *quasars* for short. These are puzzling objects because they seem to be enormous distances away, some exhibiting red shifts comparable to speeds of approximately 80% the speed of light, and emitting energy comparable to the conversion of one billion solar masses into energy. Astrophysicists are yet unable to propose a mechanism that would account for the generation of such tremendous quantities of energy. Some have proposed that these are galactic explosions occurring between galaxies of ordinary matter and galaxies composed of antimatter. Others think these may be remnants of the primeval atom. Some think they may be a new class of star known as a *neutron* star which consists entirely of neutrons.

More recently, another class of strange objects has been observed by radio astronomers. These objects emit radio waves in extremely precise pulses. These have been given the name *pulsars*. Their nature is still unknown but some scientists think they may be neutron stars rotating at high speeds and emitting a radio pulse characteristic of their rotation, similar to a flashing light on a police car or emergency vehicle. So far, these strange objects, appearing at such enormous distances away, seem to support the "Big Bang" theory rather than the steady state theory. Only time will give us the answer.

Lately, someone has suggested that perhaps not all of the stars in the universe are visible and this may have some effect on the outcome of the two models. It is proposed that some large neutron stars might undergo gravitational collapse, or contract under the influence of their great mass into a singularity point, or "black hole" in space-time. The star becomes so dense as it contracts that it curves space-time around it so that no light can escape because, once emitted by the star, it follows a curved path and cannot leave the vicinity of the star. Such a star would disappear from sight and leave no evidence to the outside universe.

Even if we were able to decide between these cosmological models or must develop new ones, the question will still remain, where did the matter come from in the first place? Perhaps, we can find no better answer than "In the beginning God...".

SUGGESTED OUTSIDE READINGS:

M.W. Avenden, *Life in the Universe: A Scientific Discussion,* Doubleday (Anchor), Garden City, New York, 1962.

L. Barnett, *The Universe and Dr. Einstein,* The New American Library, New York, 1952.

H. Bondi, *The Universe at Large,* Doubleday (Anchor), Garden City, New York, 1960.

G. Gamow, *The Creation of the Universe,* The New American Library, New York, 1957.

H. Grayson - Smith, *The Changing Concepts of Science,* Prentice-Hall, Englewood Cliffs, New Jersey, 1967.

F. Hoyle, *Astronomy,* Doubleday, Garden City, New York, 1962.

P.M. Hurley, *How Old is the Earth?,* Doubleday (Anchor), Garden City, New York, 1959.

D.W. Sciama, *The Unity of the Universe,* Doubleday (Anchor), Garden City, New York, 1961.

ARTICLES FROM *Scientific American:*

G. Burbridge and F. Hoyle, "The Problem of Quasi-Stellar Objects" (December 1966).

J.L. Greenstein, "Quasi-Stellar Radio Sources" (December 1963).

M. Miller and W. Miller, "Planetary Nebulae" (April 1963).

P. J. E. Peebles and D. T. Wilkenson, "The Primeval Fireball" (June 1967).

M. S. Robert, "Hydrogen in Galaxies" (June 1963).

A. R. Sandage, "Exploding Galaxies" (November 1964).

T. Takahashi,, "The Composition of the Earth's Interior" (June 1965).

STUDY QUESTIONS:

1. What is the Uniformatarianism Principle in geology?

2. Describe the role of erosion in the formation of surface features of the earth. Name three erosive agents of nature and describe their relative effectiveness.

3. How could you account for surface features such as the Grand Canyon, Stone Mountain, and Bryce Canyon?

4. What evidence is there to support the theory that large areas of the earth were once covered by shallow seas?

5. What evidence exists that mountains were once uplifted?

6. What topological features suggest that glaciers once covered parts of the United States?

7. What evidence is there to suggest that the moon never had much of an atmosphere or much water on its surface?

8. How does the moon influence geological features of the earth?

9. From outside readings, describe the nebular theory of the origin of the solar system of Kant and Laplace. What were some of the weaknesses of this theory?

10. How do scientists determine the distances of stars?

11. What observations support the hypothesis of an expanding universe?

12. What are neutron stars and why are they of interest to the astronomer?

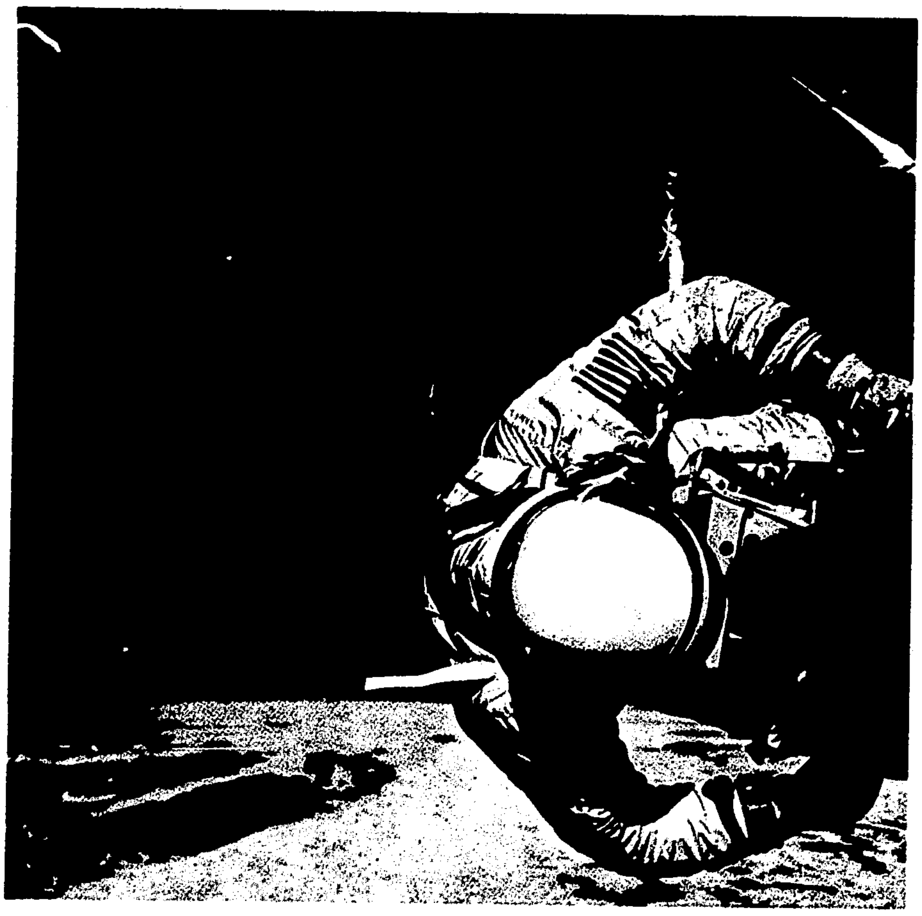
13. How many kilometers are there in a light-year?

14. What is the number of hydrogen atoms per cubic light-year of space if there is an average density of one hydrogen atom per cubic centimeter?

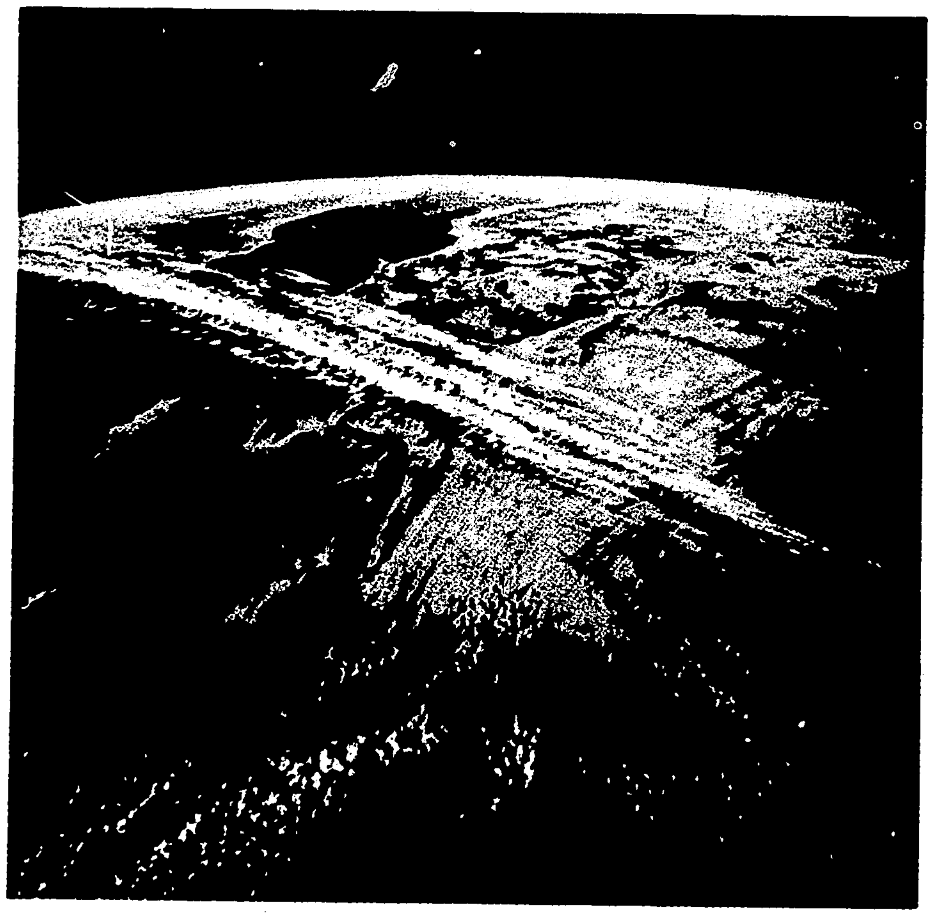
15. Describe the "big-bang" theory of the origin of the universe. What are some of its weaknesses?

16. Describe the "steady-state" theory of the origin of the universe. What are some of its weaknesses?

GEMINI XII EARTH-SKY VIEWS

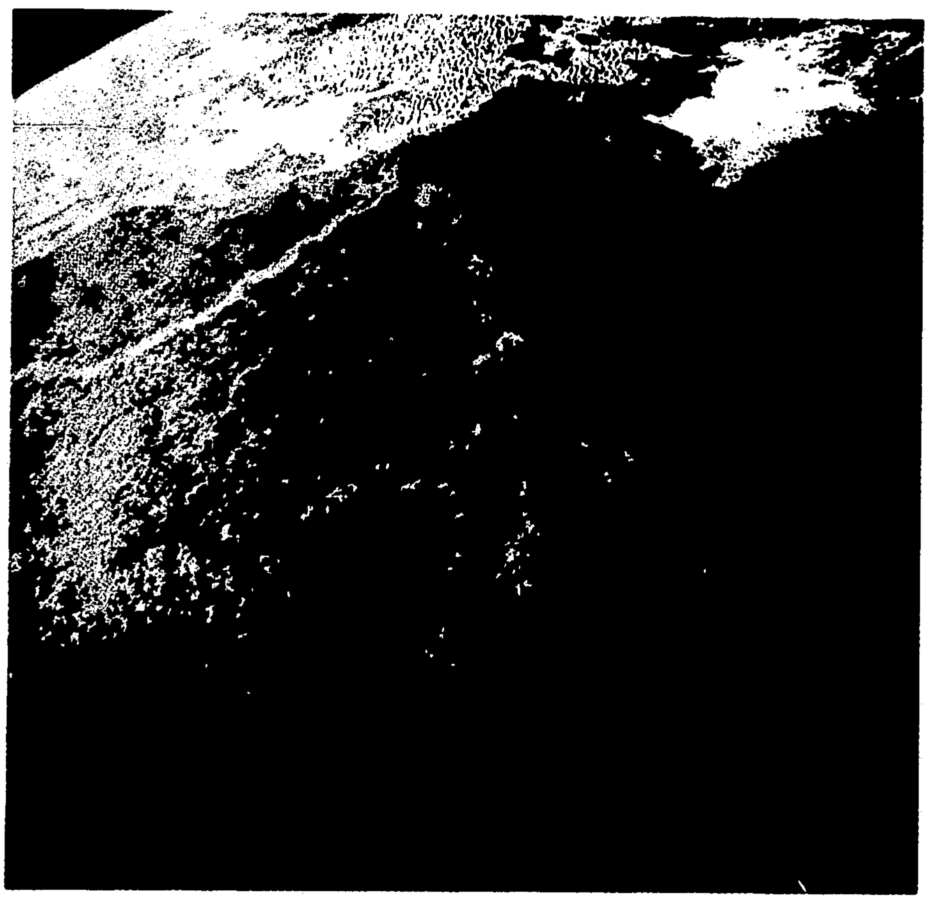


Astronaut Edwin E. Aldrin Jr. during his Gemini XII Extravehicular Activity.



United Arab Republic (Egypt), Nile Valley, Red Sea, and Arabian Peninsula. View is looking southeast.

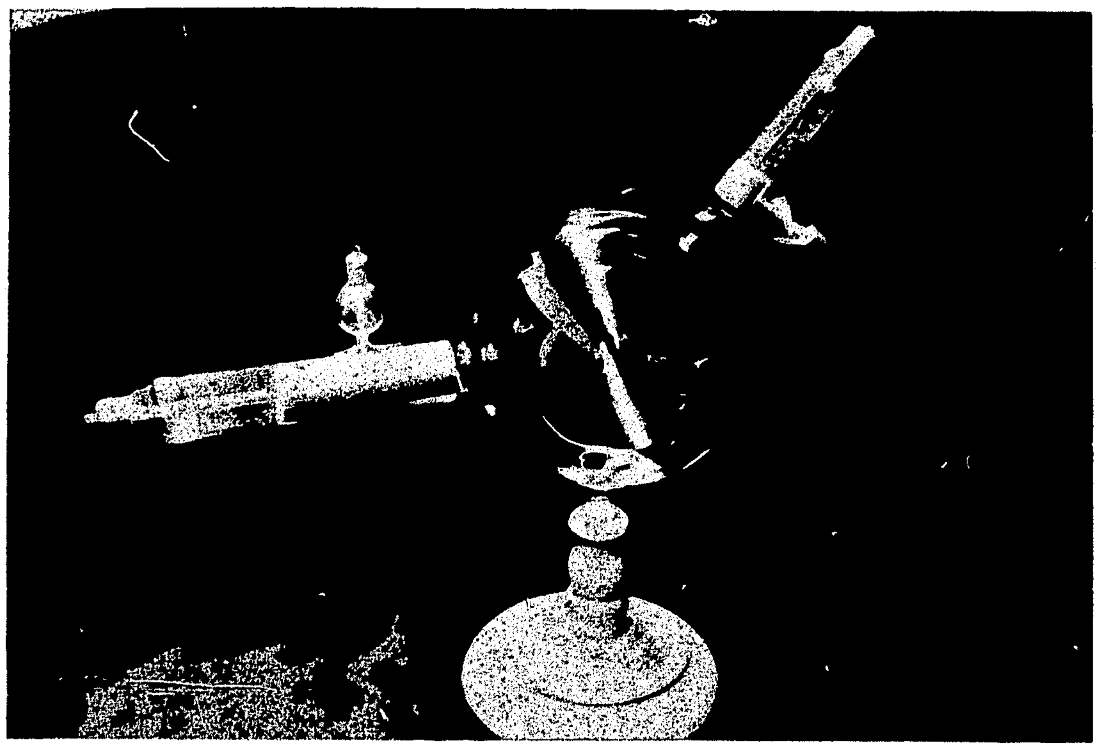
Agena Target Docking Vehicle tethered to the Gemini XII spacecraft over Baja California and Gulf of California. View is looking south.



Texas-Louisiana Gulf Coast looking east with view of Houston, Galveston and Manned Spacecraft Center area. Coast line is seen from Matagorda Bay to Vermillion Bay.

CHAPTER 9 THE NEW PHYSICS

PAGE



INTRODUCTION The turn of the nineteenth century saw the beginnings of a revolution in physical concepts that greatly affected the physical sciences in many areas. Many of the cherished concepts of Newtonian mechanics were shown to be inadequate in their explanation of newly discovered phenomena, and new concepts were developed. In the waning years of the nineteenth century, physicists and chemists turned their attention primarily to unlocking the secrets of the atom through the study of atomic spectra, or the radiation emitted by atoms. Research into this area brought up many new questions concerning the structure of matter and its relationship to energy. We have already briefly discussed the historical development of these concepts and ideas and how they reached fruition in the work of Roentgen, Rutherford, Planck, Einstein, and others. In this section we will concentrate our attention upon the physical phenonema themselves and see how they have led to many of the modern technological achievements of today.

CATHODE RAYS, ELECTRONS, AND TELEVISION We previously discussed Sir J. J. Thomson's experiments into the nature of cathode rays and their subsequent identification as electrons. Let us see how these investigations led to the invention of one of the most common means of communication that we have today - the television.

You may recall that cathode rays were streams of particles that were emitted from the cathode of a gaseous discharge tube and traveled in straight lines toward the anode. These cathode rays possessed some very interesting properties that, once understood, opened up new and better means of communication for all peoples of the world. Due to the fact that cathode rays were electrically charged, they were capable of being deflected from their paths by electrical and magnetic fields. Sir J. J. Thomson utilized a combination of electrical and magnetic fields to deflect these particles and thereby obtained a method for determining one of their physical properties, the ratio of their charge to their mass - called their e/m ratio. This he found to be a constant value for all gases used in gaseous discharge tubes. Later, other scientists showed that this ratio was only a constant for a certain energy and decreased as the energy of the electrons increased. We will see later that Einstein explained this peculiarity by assuming that energy could be converted into mass and vice-versa.

In 1909 R. A. Millikan determined experimentally that the charge on the electron existed only in discrete units, another quantum effect. He determined this by balancing the weight of a charged oil droplet between charged parallel plates against the electrical forces existing between the plates. By measuring the forces and weights for many oil droplets, he showed that charges on the droplets existed in discrete units. The charge on each electron he determined to be 1.6×10^{-19} coulombs, a unit of charge named after Charles Coulomb who verified the inverse square law of force for electrical charges. It has been determined that approximately 6.2×10^{18} electrons make up one coulomb of charge, which makes the charge on each electron 1.6×10^{-19} coulombs. Electrons possess the smallest unit of electrical charge found in nature although some theoretical physicists think that charges of 1/3 of an electron's charge may exist in nature which they call *quarks*.

Some other properties of the electron are:

1. Electrons have a mass approximately 1/1836 that of the proton. One electronic mass is approximately 9.11×10^{-31} grams.

2. Since they are electrically charged, electrons can be deflected by electrical and magnetic fields.

3. Electrons produce ionization and heat in molecules with which they collide.

4. Whenever electrons are accelerated, they give off electromagnetic radiation. For large accelerations, the radiation may be in the form of x-rays. Recently, some television receivers were ordered off the market because they were found to be emitting low energy x-rays, caused by the inadequate shielding of components which accelerated electrons through high voltages.

A variation of the gaseous discharge tube became known as the *cathode ray tube* and this became the forerunner of the modern television tube. In order to measure their e/m ratio, Thomson deflected the cathode rays from a straight path by an electric field by placing two charged parallel plates inside the tube. The electrical force between the plates deflected the electron beam downward when the bottom plate was positively charged and the top plate was negatively charged. If two magnetic coils were placed at right angles to the parallel electric plates and an electric current applied to them, a magnetic field could be produced within the cathode ray tube. The beam of electrons in the tube constitutes an electrical current and hence it will be affected by the magnetic field coils and be deflected from its straight line path. By choosing the proper direction of current through the coils, the magnetic field can deflect the electron beam in a direction opposite to that of the electric field. The magnetic force exerted on the electrons depends upon the charge and velocity of the electrons and the strength of the magnetic field. By applying the electric and magnetic fields together and adjusting their values, Thomson was able to obtain a value of e/m.

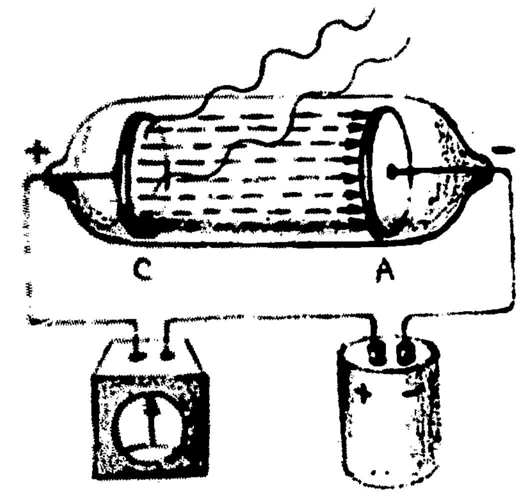
The modern cathode ray tube used in television reception is closely related to the tube that Thomson used to make his measurements. A cathode emits an electron beam which is accelerated toward an anode whose potential is sometimes as high as 25,000 volts positive. The electrons pass through the anode which is often a cylindrical electrode and travel in a straight line until they enter the region of the deflecting plates. Here the electric plates may defect the beam in a vertical plane while the magnetic coil may deflect it in a horizontal plane. By rapidly varying the current in the coils and the electric charge on the plates, the beam can be made to scan up and down and back and forth many times per second. If the inside of the front surface of the cathode ray tube is coated with a fluorescent material, the energetic electrons will excite its atoms and cause them to emit light. By controlling the number of electrons being emitted by the cathode, one can vary the light intensity and produce a picture on the screen.

THE PHOTOELECTRIC EFFECT In 1887, Heinrich Hertz, who discovered Maxwell's electromagnetic waves, discovered another effect which was to have far-reaching effects on the concepts of physics - the photoelectric effect. Hertz observed that light incident on certain metals could produce an electrical current if the light consisted of waves of short wavelength such as ultraviolet light. Maxwell's theory of electromagnetic waves predicted that only the strength or intensity of the light would have an effect on the emission of electrons from these metal surfaces. Further investigations, however, showed that such was not the case. The number of photoelectrons did depend upon the light intensity but only if the light was of the correct wavelength or frequency. If the wavelength was very large, no photoelectrons were emitted regardless of the amount of light incident upon the metal.

Scientists were perplexed by this failure of Maxwell's theory to adequately explain this effect when it apparently correctly explained all other phenomena. In 1900, Max Planck proposed a new concept to explain the spectra obtained from hot bodies, or so-called black body radiation. He proposed that atoms could only emit energy in discrete units or packets called *quanta*. The energy emitted could only occur in integral multiples of the frequency of vibration of the atom, or

$$E = nhf, \qquad (9.1)$$

where n is an integer, h is a constant called Planck's Constant, and f is the frequency of the vibration of the atom.

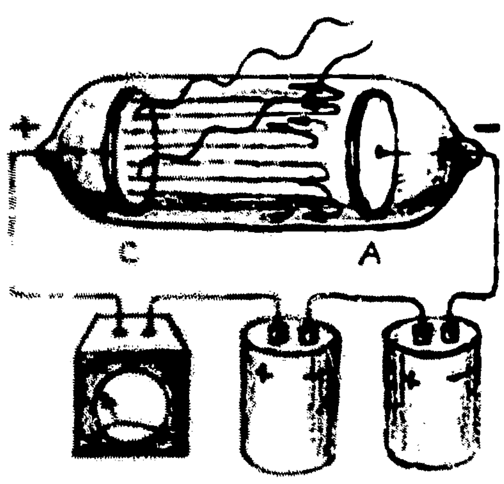


When the anode, A, is negative with respect to the cathode, C, some photoelectrons are still energetic enough to reach the anode.

This concept of quanta of energy was applied in 1905 by Einstein to the photoelectric effect with success. Einstein proposed that light itself consisted of discrete units or quanta, which later were called *photons*. In the absorption and emission of light, atoms only accepted or emitted photons of a particular frequency which were integral multiples of Planck's Constant, h, or of energy, $E = nhf$.

Experiments showed that the number of photoelectrons emitted depended upon the number of photons of a particular frequency. If the frquency of the light decreased, a point was reached where no more photoelectrons were emitted. Einstein theorized that atoms could only absorb light energy in units of hf, and below a certain frequency, called the *threshold frequency*, f_o, no emission occurred.

One could also perform experiments to show that the kinetic energy of the emitted photoelectrons depended upon the energy of the photons and the type of material the photons were incident upon. If a photocathode (an electrode capable of emitting photoelectrons) is sealed in a glass tube with an anode, one can apply a voltage across the tube and collect the photoelectrons emitted when photons of the proper frequency strike the photocathode. Such a device is called a photoelectric cell or photocell. If an ammeter is connected in series with the photocell and voltage source, one can measure the photoelectric current produced. Arranging the circuit so that the anode is positive, one collects all of the photocurrent. If the anode is made slightly negative, one finds the photocurrent decreases but does not go to zero as one might expect. The reason for this is that some of the photoelectrons possess enough kinetic energy to overcome the negative repulsive force of the anode. Increasing the negative voltage of the anode will finally stop all photoelectrons from reaching the anode. The value of the voltage for this to occur is called the *stopping potential*, V_s.



Increasing the negative potential of the anode to $-V_s$, causes the photoelectrons to be repelled.

Einstein found that the maximum kinetic energy of the photoelectrons, the energy of the photons, and the stopping potential were related by the following relationship:

$$(K.E.)_{max} = hf - W = V_s\ (1.6 \times 10^{-19}\ coul.), \qquad (9.2)$$

where W is a constant for each metal, called its work function, or electronic binding energy. W represents the energy or work done in removing the electron from the atom to which it belongs in the metal. Some electrons within the surface of the metal lose some kinetic energy in reaching the surface, while those emitted at the surface have maximum kinetic energy. This relationship explains why only light of a particular frequency will result in photoelectron emission. If $hf = W$, the kinetic energy of the photoelectrons is zero and no emission occurs. Only for hf greater than W will photoelectrons be emitted. $W = hf$ is the threshold energy when f is the threshold frequency.

Photocells today are used extensively to control circuits and operate various devices. A common use is the control of street lighting by photocell circuits. By employing a photocell in an off-on circuit, one can utilize the sunlight to turn the street lights on at dusk and off at dawn. The sunlight creates a photocurrent which is used to prevent the lights from operating as long as it is produced. On cloudy days or nights, the photocurrent stops and the circuit turns on the street lights.

Another useful application in industry is the control of products on automatic conveyor belts and the counting of items as they flow past the photocell. Each time the light beam is interrupted, a counter registers another item. Some automobiles also utilize photocells to automatically dim high-beam headlights whenever light from oncoming cars is bright enough to activate its circuit.

A very common use is found in photography. Here, photocells are used to measure the amount of light intensity present so that the photographer does not need to guess what the proper exposure setting for his camera should be. The photocurrent is used to operate a meter which is calibrated against a standard light source and correlated with camera exposure settings.

For his explanation of the photoelectric effect, Einstein was awarded the Nobel Prize in Physics in 1921, but it was for his Special Theory of Relativity, which he published the same year as the photoelectric effect theory, that he achieved his greatest fame. His theory of relativity, more than any other single work, stands as one of the greatest achievements of the human intellect.

THE SPECIAL THEORY OF RELATIVITY In 1905, Einstein published his Special Theory of Relativity which greatly altered man's concept of matter and energy and his view of the physical universe. His theory was formulated to explain certain discrepancies occuring in electromagnetic theory concerning the propagation of light. We have already discussed briefly the Michelson-Morley experiment and its failure to detect the speed of the earth relative to the ether, the medium through which light was believed to propagate in the nineteenth century. To explain the null results of this experiment and other data, Einstein made two bold assumptions in his Special Theory that called for radically new concepts concerning space and time, and energy and matter.

The first assumption was that *the speed of light in a vacuum is a constant value for all observers regardless of their motion;* the second stated that *the laws of physics acquire the same formulation in all systems moving with uniform speed relative to each other.* Without going into the mathematical details, we will present several important consequences of these two assumptions.

The first assumption requires that all observers obtain the same value for the speed of light in a vacuum, regardless of their speed relative to the light source itself. Since speed is defined in terms of distance and time, this means that distance and time are no longer absolute quantities, but have different values for different observers. This result leads to the concept of a unification of the Newtonian concepts of distance (space) and time into a new concept of a space-time continuum where space and time are interrelated.

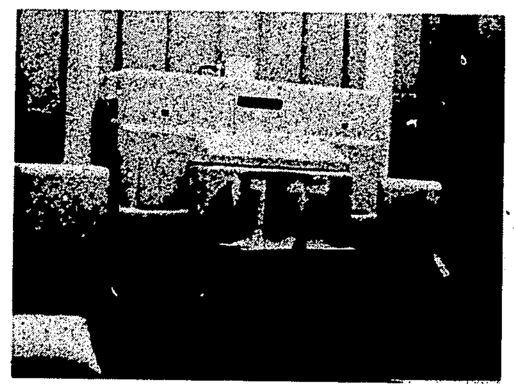
If a stationary observer measures the speed of light in a vacuum by measuring the distance d it travels in time t, he obtains the value $v = d/t = c$. If a moving observer measures the speed over a distance d' and a time t', he also obtains the value $v' = d'/t' = c$. In order for both of these observers to obtain the same value, c, their distance and time scales must be altered. This leads to the concept of length contraction and time dilation. A stationary observer would view lengths in a moving frame of reference as shorter than when the frame is at rest. Clocks in the moving frame would appear to run slower than in the stationary frame. Likewise, the moving observer sees clocks running slower in the stationary frame and lengths in that frame contracted. The reason for these effects is that light has a finite, constant speed for all observers, so that they must reinterpret their ordinary concepts of length and time to account for this fact. Such ideas may appear nonsensical at first, but such effects only become noticable at very large speeds approaching that of light.

Another effect predicted by special relativity theory is that the mass of an object is dependent upon the speed of the object. An object at rest is said to possess a *rest mass*, m_o. If it is moving with a speed v, its mass is no longer m_o, but it is a relativistic mass, m. Einstein showed that these masses are related by the equation,

$$m = \frac{m_o}{\sqrt{1 - (v/c)^2}} . \qquad (9.3)$$

At speeds when v is very small compared to the speed of light c, the relativistic mass is equal to m_o, because the expression under the square root sign becomes unity. As the speed of the object approaches the speed of light $(v \rightarrow c)$, the expression under the square root sign approaches zero, and the relativistic mass m approaches a very large value or tends toward infinity as $v = c$. To accelerate an object with finite mass to the speed of light, therefore, would require an infinite force since its mass would become infinitely large as it approached the speed of light. Hence, only particles of small mass, such as electrons and protons can be accelerated to speeds approaching that of light.

Today, we can accelerate objects such as the Apollo spacecraft to speeds of about 7 miles per second, which is a very small fraction of the speed of light (c = 186,000 miles per sec). In the laboratory, however, physicists have accelerated electrons to speeds of about 0.99999 times the speed of light in devices known as β- particle accelerators, or betatrons. As more energy is added to the electron, its speed increases by only an infinitesimal amount but its mass increases rapidly. At speeds of 0.99999 c, an electron would have a relativistic mass about 220 times its rest mass. Where is this increase in mass coming from? Einstein's theory predicted that mass and energy are related by his famous equation,



A betatron used in medical research.

$$\Delta E = \Delta m \; c^2,$$

where ΔE is the change in energy, Δm is the change in mass, and c^2 is the speed of light squared. The increase in mass appears because the energy added to the electron is converted into mass and is not used to accelerate the electron to higher speeds.

Due to this increase of mass with speed, the expression given earlier for the kinetic energy of an object, $KE = \frac{1}{2} mv^2$, is not valid at high speeds. Instead, the kinetic energy is

$$KE = (m - m_o) \; c^2,$$

or

$$KE = mc^2 - m_oc^2. \qquad (9.5)$$

The term mc^2 is called the total energy of the object, E, or

$$E = KE + m_oc^2. \qquad (9.6)$$

When the object is at rest, its speed and kinetic energy are zero, but it still has an amount of energy $E = m_oc^2$, called its rest-mass energy. It is this energy that is released in a nuclear reaction in a nuclear weapon or a nuclear reactor when mass is converted to energy. This equivalence of energy and mass also explains the observation that the e/m ratio of electrons decreases with increasing energy. As the electrons gain energy, some of the energy is converted into mass, causing a decrease in the e/m ratio.

THE NEW MECHANICS Einstein's success in explaining the photoelectric effect by assuming that light possessed both wave and particle properties and his relationship between matter and energy prompted other scientists to question the simple mechanistic models of matter of nineteenth century physics. One of the first to do so was Louis de Broglie, who in 1924 proposed that matter may possess both wave and particle properties. Other scientists began to look for these properties and soon verified de Broglie's hypothesis. In 1926, Davisson and Germer succeeded in diffracting electron waves by crystals in a manner similar to x-ray diffraction.

About this time, theoretical physicists - notably Heisenberg, Schrödinger, Born, and Dirac, were developing a new formulation of mechanics to explain these phenomena, *quantum wave mechanics*. This theory was a mathematical model instead of a physical model and called for a reinterpretation of the classical concepts of matter.

Quantum wave mechanics replaced the mechanistic concepts of describing the position and speed of particles uniquely by a *probabilistic* description. Heisenberg introduced his Uncertainty Principle - the fact that one can not measure both the position and speed of a particle to any desired degree of accuracy - which placed a limit to the accuracy of physical measurements. This principle is not important for large objects such as a baseball or planet, because their wavelength is so small compared to the measuring apparatus, but for subatomic particles, such as electrons, the uncertainty in the simultaneous measurement of position and speed becomes a major problem. This uncertainty comes about due to the interaction of the measuring apparatus with the object whose properties it is attempting to measure.

Max Born interpreted the inability to measure all properties of matter by either particle or wave properties as evidence that matter was probabilistic in nature and sometimes behaved like a wave and at other times like a particle. He regarded these two ways of viewing matter as complementary ways of viewing a more fundamental process. Today, most scientists accept this wave-particle duality of matter as a fundamental property of all matter and energy. Other scientists, notably Einstein, refused to accept the probabilistic interpretation of physical processes, stating that he did not believe that God threw dice in determining the behavior of the universe. Today, physicists and chemists utilize quantum mechanics in determining the structure and behavior of matter with great success. The theory adequately explains all physical phenomena from the motion of the planets to the motion of electrons and gives as close agreement with experiment as one chooses. Whether quantum mechanics will have the final word in describing the physical universe or whether it will be replaced by a more comprehensive theory which contains it as a special case, as it does Newtonian mechanics, is still a question for future scientists to answer.

SUGGESTED OUTSIDE READINGS

L. Barnett, *The Universe and Dr. Einstein*, The New American Library, New York, 1948.

L. de Broglie, *The Revolution in Physics*, Noonday Press, New York, 1953.

G. Gamow, *Mr. Tomkins Explores the Atom*, Cambridge University Press, New York, 1945.

G. Gamow, *Mr. Tomkins in Wonderland*, Macmillan, New York, 1947.

G. Gamow, *Thirty Years That Shook Physics: The Story of Quantum Theory*, Doubleday (Anchor), Garden City, New York, 1966.

ARTICLES FROM *Scientific American:*

H. Alfvén, "Antimatter and Cosmology", (April 1967).

P. A. M. Dirac, "The Evolution of the Physicist's Picture of Nature", (May 1963).

G. Gamow, "The Principle of Uncertainty", (January 1958).

E. Schroedinger, "What is Matter?", (September 1953).

STUDY QUESTIONS

1. Compare the revolution in society produced by the discovery of the electron with that produced by the invention of the steam engine.

2. Why did the value of e/m decrease as the energy of the electrons increased?

3. It is sometimes said that scientific theories and discoveries often have an impact upon social and political philosophies. Compare the social and political philosophies of the eighteenth and nineteenth centuries with the prevailing deterministic and mechanistic scientific outlook. What effect, if any, has the relativistic and non-deterministic scientific outlook of the twentieth century had on social and political philosophies?

4. Why don't we see interference effects for large objects such as baseballs?

PICTURE CREDITS

Cover photograph, NASA (Designed by Hans Bhalla)

P. 3	NASA
P. 4	NASA
P. 5	NASA
P. 7	Harvard Project Physics
P. 8	Harvard Project Physics
P. 9	Harvard Project Physics
P. 10	NASA
P. 11	Harvard Project Physics
P. 12	Harvard Project Physics
P. 14	Harvard Project Physics
P. 15	Harvard Project Physics
P. 17	Harvard Project Physics
P. 18	Harvard Project Physics
P. 19	NASA
P. 22	Harvard Project Physics
P. 23	Harvard Project Physics
P. 24	(top) Harvard Project Physics; (center, bottom) NASA
P. 26	NASA
P. 39	Harvard Project Physics
P. 43	Harvard Project Physics
P. 45	Harvard Project Physics
P. 58	Emory University Clinic
P. 61	Neely Nuclear Research Center, Georgia Institute of Technology
P. 79	Harvard Project Physics

P. 80	(top) Harvard Project Physics
P. 94	Harvard Project Physics
P. 104	(center, bottom) Harvard Project Physics
P. 105	(center, bottom) Harvard Project Physics
P. 107	(top) Harvard Project Physics
P. 108	(top) Harvard Project Physics
P. 132	Harvard Project Physics
P. 134	Harvard Project Physics
P. 140	(top) Harvard Project Physics
P. 141	(top) Harvard Project Physics
P. 147	Harvard Project Physics
P. 150	Harvard Project Physics
P. 156	NASA
P. 164	(top) NASA: (center, bottom) Fernbank Science Center
P. 174	Harvard Project Physics
P. 177	Emory University Clinic

All photographs and diagrams not credited above were made by the staff of the Cooperative General Science Project.