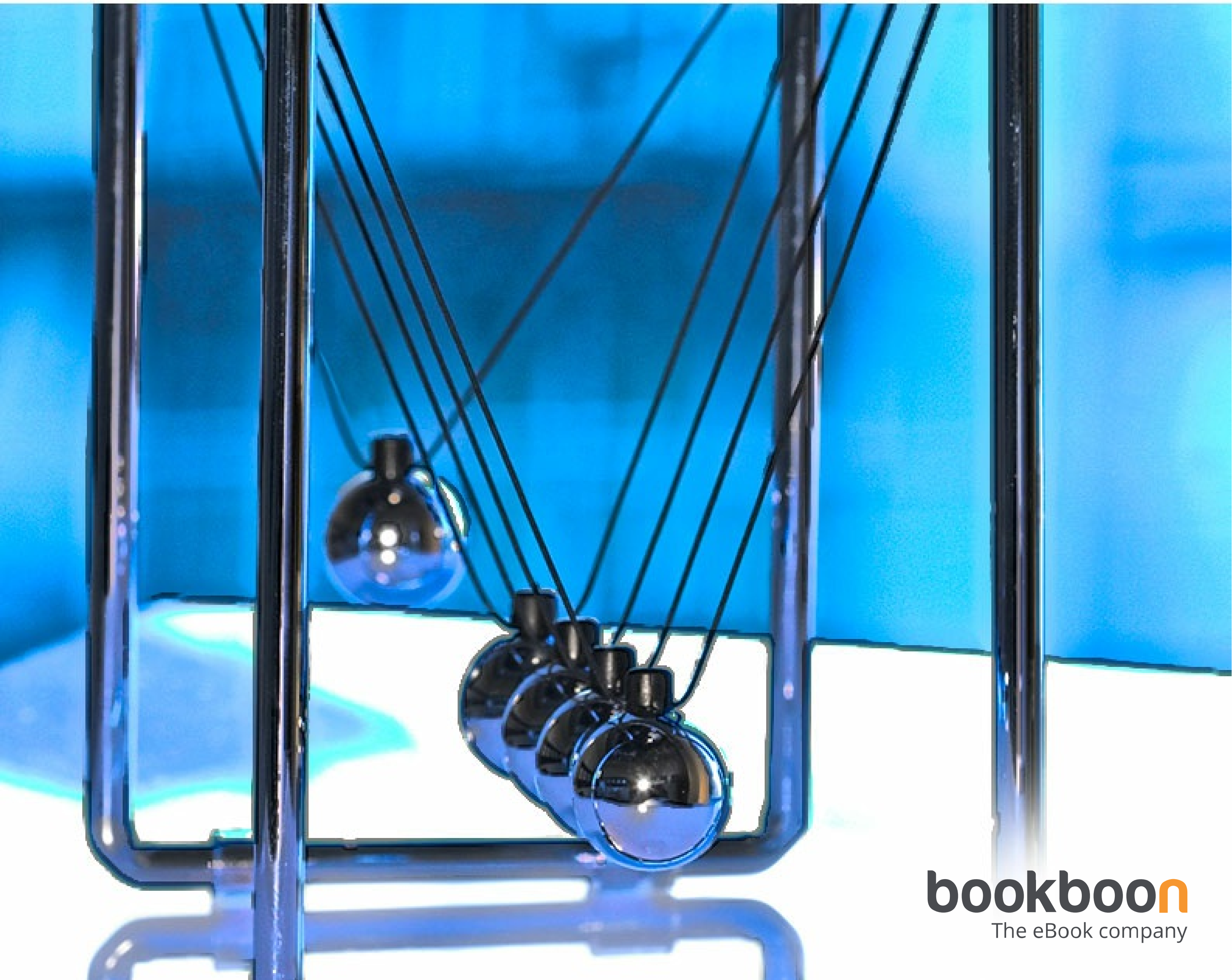
Mike Corwin

20th Century Physics



bookboon
The eBook company

MIKE CORWIN

20TH CENTURY PHYSICS

20th Century Physics
1st edition

ISBN 978-87-403-2878-3
Peer review by Dr Ulrich Zurcher, Professor of physics at Cleveland State University.

CONTENTS

	Preface	**7**
1	**Measurements and Theories**	**8**
1.1	What is Physics?	9
1.2	Which Came First, the Data or the Theory?	13
1.3	Can't You Speak Plain English?	16
1.4	Fundamental Physical Quantities	17
2	**Conservation Laws and Energy**	**24**
2.1	Conservation of Charge	25
2.2	Conservation of Momentum	26
2.3	Energy	31
2.4	Conservation of Energy	37
3	**Modern Physics: A Brief History**	**46**
3.1	Observations That Appeared To Defy Analysis Using Classical Physics	46
3.2	Twentieth-Century Physics	55
3.3	The Correspondence Principle	62

4	**Physical Laws in Different Frames of Reference**	**66**
4.1	Frames of Reference	66
4.2	Inertial Frames of Reference	72
4.3	The Principle of Relativity	72
5	**The Special Theory of Relativity: Time Dilation**	**78**
5.1	The Special Theory of Relativity	78
5.2	Time Is a Relative Physical Quantity	80
6	**The Special Theory of Relativity: Length Contraction and Velocity Addition**	**93**
6.1	Length Contraction	93
6.2	Space-Time	99
6.3	Some Relativistic Examples	101
6.4	The Twin Paradox	105
6.5	Velocity Addition	107
7	**The General Theory of Relativity**	**116**
7.1	The Principle of Equivalence	118
7.2	The General Theory of Relativity	121
7.3	Evidence for the General Theory of Relativity	125
7.4	The General Theory of Relativity Today	127
8	**Electromagnetic Radiation: The Classical Theory**	**135**
8.1	The Nature of Light	137
8.2	Electromagnetism	140
8.3	Properties of Electromagnetic Waves	142
8.4	The Electromagnetic Spectrum	146
9	**The Photoelectric Effect**	**152**
9.1	The Photoelectric Experiment	153
9.2	Einstein's Quantum Hypothesis	156
9.3	Particles or Waves?	163
10	**More Evidence of Photons**	**170**
10.1	The Compton Effect	170
10.2	Antimatter	173
10.3	Pair Production	174
10.4	Pair Annihilation	178
10.5	Wave-Particle Duality	180

11	**Classical Models of the Atom**	**184**
11.1	Dalton's Indivisible Atoms	185
11.2	The Nuclear Model of the Atom	186
11.3	The Classical Planetary Model of the Atom	193
12	**The Bohr Model of the Atom**	**201**
12.1	Atomic Spectra	201
12.2	Quantizing the Planetary Model of the Atom	206
12.3	Difficulties with the Bohr Model	217
13	**Quantum mechanics**	**224**
13.1	De Broglie Waves	225
13.2	Schrodinger's Wave Equation	232
13.3	Heisenberg's Theory of Matrix Mechanics	235
13.4	What is a Matter Wave?	239
13.5	The Probability Interpretation of the Wave Function	241
13.6	The Copenhagen Interpretation	250
14	**The Nucleus**	**259**
14.1	The Composition of the Nucleus	259
14.2	Quarks	260
14.3	Observing Subatomic Particles	262
14.4	Nuclear Notation	263
14.5	The Force Between Nucleons	266
14.6	The Deuteron	267
14.7	Mass Defect and Binding Energy	269
15	**Induced Nuclear Reactions**	**274**
15.1	Induced Nuclear Reactions	274
15.2	Fission	279
15.3	Fusion	284
16	**Radioactivity**	**288**
16.1	The Radioactive Decay Law	290
16.2	Age-Dating with Radioactive Isotopes	295
16.3	Gamma Decay	298
16.4	Alpha Decay	302
16.5	Beta Decay	303
	Appendix A: Scientific Notation	**311**

PREFACE

I believe that the events that marked the transition from classical to modern physics are without historical parallel as illustrations of the power of rational and imaginative thought and thus play an important role in the education of both scientists and nonscientists alike.

Twentieth Century Physics: A Liberal Arts Approach is the story of a scientific revolution and those uniquely creative individuals who led it. Because this new physics demands radical, and at first incomprehensible, changes in our worldview, it is also the story of a philosophical revolution. Very little of our classical commonsense approach to the physical universe has survived. The clockwork universe consisting of tiny mechanical parts marching inexorably forward under the control of rigid laws is no longer a part of scientific thought. Physics has progressed beyond the final answers of the nineteenth century to a belief that the universe will never surrender her full beauty and depth to our relentless probing. Progress in science is now seen as successively more accurate and comprehensive approximations to the true nature of reality. Each generation of seekers will find new and enticing secrets to pry from nature's infinite store of wonders.

This book is organized into four parts: Introduction (Chapters 1 through 4), relativity (chapters 5 through 7); quantum physics (chapters 8 through 13); and nuclear physics (chapters 14 through 16). Classical physics is introduced only where it is needed to understand certain concepts in twentieth-century physics.

At the end of each chapter there is a chapter summary and some questions related to the chapter. You may find it helpful to go over the summary before (as well as after) you read the chapter. This will give you, in a few brief paragraphs, a sense of the most important ideas to be developed in the chapter

This book does not attempt to give a complete picture of modern physics as a body of knowledge. Instead, I have selected material primarily on the basis of what I found exciting and thought provoking as a student. I have stressed the historical and philosophical aspects of modern physics throughout, not only because they are interesting, but because they contribute to the understanding of the physics.

1 MEASUREMENTS AND THEORIES

At the beginning of the twentieth century, physicists looked back with pride on more than two centuries of remarkable progress. In 1687, Isaac Newton (1642 – 1726) published *Principia*. This monumental achievement included both his theory of mechanics, the nature of matter and motion, and also a quantitative expression for gravity.

Together, they showed that the fall of an apple from a tree was caused by exactly the same gravitational force responsible for the orbit of the moon around the earth and the orbits of the planets around the sun. Earlier, Johannes Kepler had, by trial and error, determined mathematical formulas for the planetary orbits. However, he had no explanation for why they moved in these particular orbits. Gravity and Newton's laws of motion explained the motion and allowed the orbits to be determined directly by theoretical calculation.

By 1900, Newton's mechanics (gravity and his laws of motion) had been supplemented by the theories of thermodynamics and electromagnetism. These provided coherent explanations of such phenomena as heat, light, electricity, and magnetism. The resulting body of knowledge is now referred to as classical physics. There were still a few puzzling facts that stubbornly resisted explanation, but physicists were confident that these puzzles would soon be solved using the theories of classical physics. All that seemed to remain for twentieth-century physicists was the routine task of filling in the details of this grand picture of the universe

Only a decade later, the situation had changed completely. The new theories of relativity and quantum mechanics were being developed. Not long after that, atoms and atomic nuclei began to yield their secrets. These new theories provided answers to those few puzzles that had resisted earlier efforts. However, they also forced a rethinking of the most basic concepts of reality, the concepts which were the foundations of all our earlier understanding of the world. The commonsense, mechanistic-deterministic model of the universe, the very foundation upon which classical physics had rested, has been completely swept away.

The focus of this book is the exciting adventure in human thought that is represented by the transition from classical physics to our current view of the universe. However, the information gathered by earlier physicists and many of the concepts they created play an important role in our story. To set the stage before plunging into the story of modern physics, we begin with a survey of the structure and the basic tools of physics.

1.1 WHAT IS PHYSICS?

Physics is the science that attempts to discover order within the universe. It deals only with the portion of human experience that lends itself to quantitative measurement. Its task is the coherent description of that experience. Physics assumes the possibility of knowledge, the genuineness and universality (at least in principle) of sense perceptions, and a systematic and orderly nature for the events occurring in the universe. In the words of Albert Einstein, "God is subtle but not sneaky!" A regular order does exist among our experiences and our task is to describe it.

Consider the following observation: "If an object is released from rest near the surface of the earth, it will fall toward the earth." This is a qualitative rather than a quantitative statement, and as such is of limited use in physics. It does serve, however, to suggest certain physical quantities that can be carefully measured to provide information about the process of free fall. Each measurement (or related set of measurements) provides a physical datum (plural data). The basic elements of physics are physical quantities, where a physical quantity is anything that can be measured.

The above example, suggests the appropriate physical quantities to measure in order to understand free fall are time and distance of fall. With appropriate measuring equipment, we might obtain the following physical datum: after four seconds of free fall, the falling object has traveled a distance of 85 meters. This datum represents simultaneous measurements time and free fall distance. If we gather data for various times of free fall, we can hope to find some regularity or pattern that will allow us build a coherent description of an object in free fall. We might organize the data into a table such as that shown in Table 1.1 or we might plot them in a graph such as that in Figure 1.1. It is clear from the graph that a pattern does exist.

Time (s)	1	2	3	4	5	6	7	8	9	10
Distance (m)	5.1	21	47	85	110	169	237	332	393	499

Table 1.1

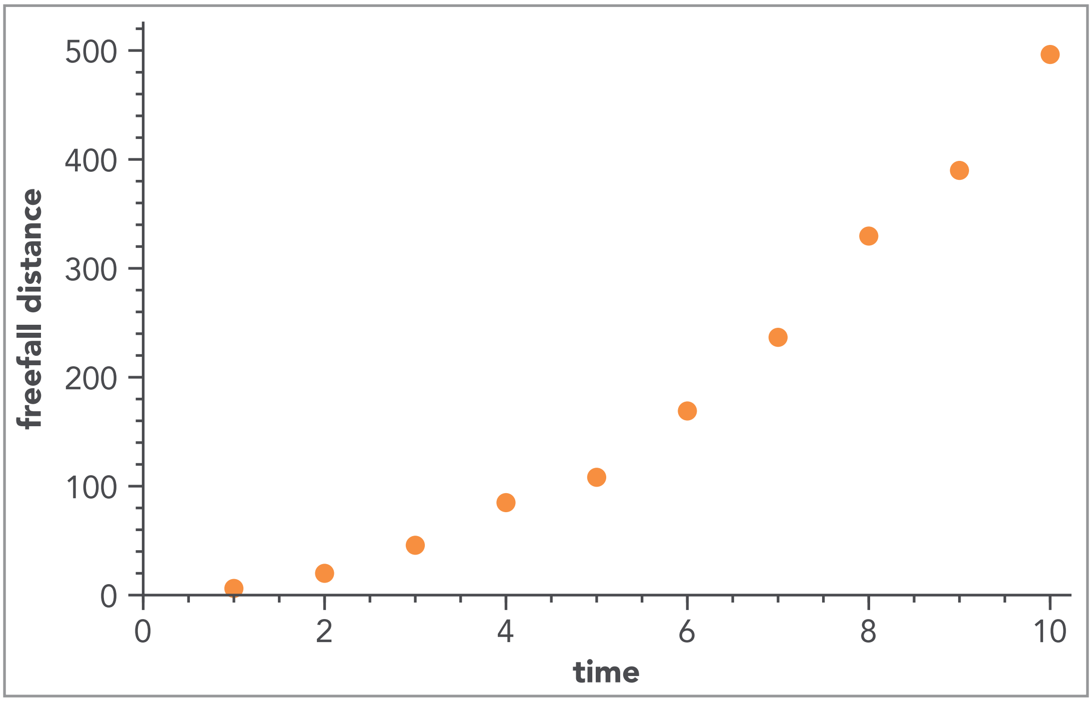


Figure 1.1

From similar data, Galileo Galilei (1564 – 1642) found the following relationship:

> **The distances traveled by a body falling from rest are to each other as the square of the free fall time.**

The language of algebra is very useful for describing such mathematical patterns. Using the symbol “d” to represent a measurement of distance of fall measured in meters and the symbol “t” to represent the corresponding time of fall measured in seconds, we can express the pattern in the data of Table 1.1 with a simple algebraic equation:

$$d = 4.9\, t^2$$

A plot of this equation matches the free fall data. The equation really says the same thing as Galileo’s statement, but it says it much more compactly.

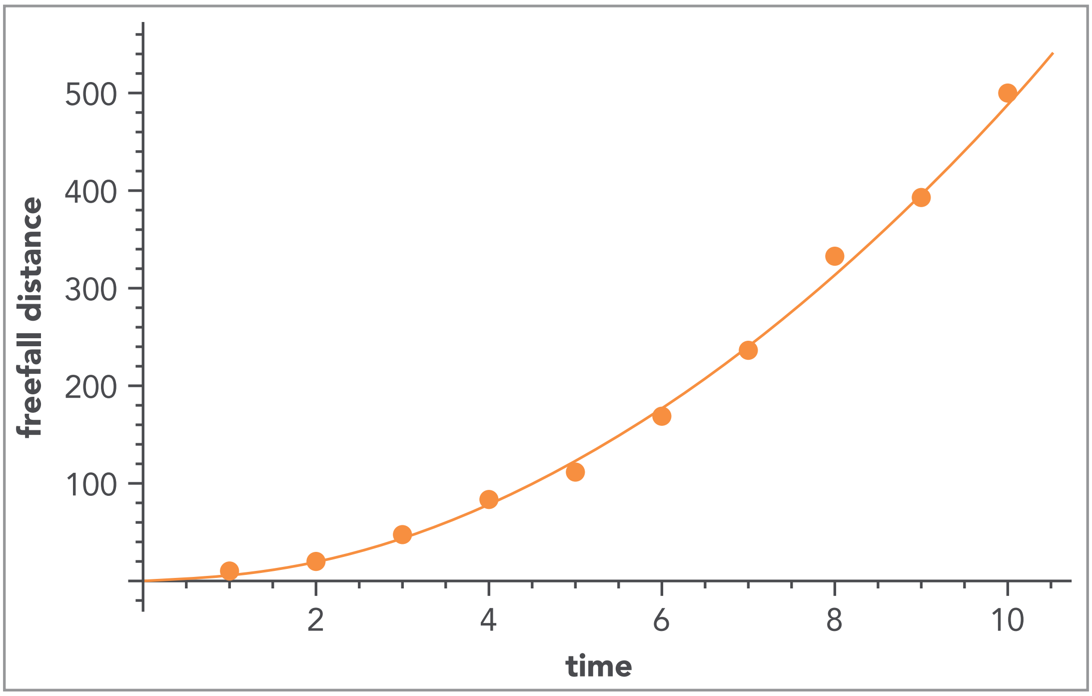


Figure 1.2

When such a relationship is found, it is repeatedly tested against physical data to determine the conditions under which it does or does not remain valid. If it proves a consistent description for a certain range of experiences, then it can be called a physical law.

Physical laws are quantitative relationships between or among physical quantities. The algebraic equation alone is not a complete statement of the law. We must know the meaning of each symbol, and we must know the conditions under which the law is valid. Galileo's relationship applies to some objects allowed to fall freely after being released from rest near the earth's surface. However, it is not valid as a description of the motion of a feather or a helium-filled balloon. It does, very closely, describe most heavy objects dropped from rest. Surprisingly, it applies exactly for any freely falling object released from rest near the surface of the earth if the object falls through a vacuum. That is, in the absence of air resistance, all objects fall at the same rate.

Over the years, Galileo and other physicists found a number of physical laws to describe the motions of falling objects, swinging pendulums, planets, and other moving objects. Isaac Newton went a giant step farther and found a more general relationship among all these physical laws. A physical theory is a more general structure that organizes and correlates various physical laws. Physical theories are built around a set of postulates. From the few basic assumptions and mathematical relationships that make up the postulates of a theory, we can deduce by purely logical and/or mathematical operations a whole group of physical laws. In this sense, we can then say we "understand" the physical laws. However, the basic assumptions and relationships themselves (the postulates) cannot be "understood" in this sense; they cannot be derived, but they must simply be accepted as the way the universe behaves -- the irreducible facts of nature.

Most of the known physical laws can be understood in terms of five major physical theories: classical mechanics, thermodynamics, electromagnetism, relativity, and quantum mechanics. In this book, we deal primarily with theories that have been created in the twentieth century, but a word or two about Newton's theory of classical mechanics will illustrate the relationship between the postulates of a physical theory and the physical laws. The two fundamental postulates of classical mechanics are:

1. **The acceleration of an object is equal to the force exerted on the object, divided by the mass of the object.**
2. **Objects having mass attract each other with a gravitational force that is proportional to the product of their masses and is inversely proportional to the square of the distance between them.**

Note that these postulates involve a number of physical quantities (acceleration, force, mass, distance) and mathematical relationships. From these postulates, we can deduce all the laws of classical mechanics, including the equation of free fall discovered by Galileo.

However, the postulates do not predict what value the constant in the derived relationships will have. Thus the strength of the gravitational force is not explained by the postulates of classical mechanics, but must be experimentally determined and added to the theory. The constants are simply "the way things are."

Also the postulates do not explain why the postulates are true. They are assumptions that must be accepted. The "truth" will be judged by the validity of the laws derived from them.

Newton's postulates "explain" both the motions of the planets around the sun and the motions of objects under applied forces. It is a powerful theory that uses just two simple assumptions to derive a great many physical laws.

1.2 WHICH CAME FIRST, THE DATA OR THE THEORY?

The preceding discussion seems to suggest that physicists collect quantitative data, analyze them to find physical laws, and then deduce the physical theories that correlate the laws. In practice, it is not that simple. Many physical laws have indeed been discovered through patient collection and analysis of physical data. However, the formulation of physical theories is a far more creative process, one in which the physicist's powers of imagination and sense of esthetics play as important a role as his or her powers of reason.

Physical theories are products of the human intellect and as such are more properly described as inventions rather than discoveries. The need for creativity in the formulation of physical theories arises because there is no unique physical theory associated with a given set of physical laws. The data do not uniquely determine the theory. There may be several different physical theories, each equally capable of providing a basis for the deduction of the known set of laws. However, there are criteria useful in choosing among several possible theories. One criterion is Occam's razor:

> **Among competing hypotheses, the one with the fewest assumptions (postulates) should be selected.**

Another criterion is aesthetic; which theory is, in Einstein's use of the word, the more beautiful. The word 'beauty' is often used by Einstein when speaking of physics. Among other things, it means simplicity, comprehensiveness, elegance. In his biography of Einstein, Jeremy Bernstein put it this way:

> What the scientist hopes and indeed what he must assume is the case in order to motivate his work is that there exists a theory ....

> which because of [its] inner harmony and the compelling nature of [its] underlying assumptions brings one deeper into the workings of the universe – "closer to the Secrets of the Old One," in Einstein's phrase.

Physical theories must be capable of deriving known physical laws. However, a good theory should also predict new, previously unsuspected, laws that follow directly from the postulates of the theory. The testing of such a new law (by comparison with physical data) is often among the most exciting and important events in physics If the data confirm the validity of the new law this provides very strong evidence for the validity of the theory. If the law turns out to be incorrect, then the validity of the theory has been "disproved" and the theory must be either modified or discarded However, in practice an attractive theory is seldom discarded as a result of a single (or even a few) experimental "disproof." In the long run, however, a theory must prove consistent with the physical data or it will have to be discarded, no matter no matter how esthetically appealing it may be. Experiment is the ultimate arbiter of any physical theory.

It is important to note that a physical theory can never be proven correct. If every physical datum is consistent with the predictions of the theory, the possibility (in fact, the probability) still exists that the theory will fail when the range of observations is extended. This has happened time and again, and there is no reason to believe it will not continue to happen. For a time, such inconsistent data can be explained by modifications of the theory or they can simply be regarded as puzzles yet to be solved. Eventually, however, the outstanding puzzles or the modifications to the theory become so numerous that some physicists are led to the more drastic step of seeking a new theory built on a different set of postulates.

The formulation of a new theory involves changes in our basic ideas about the fundamental relationships among our experiences, so it is an uncomfortable step and one that is not undertaken lightly. The goal is to find a new theory that will encompass both the old and the new observations in one theoretical structure of appealing simplicity. In this way, the physicist (and the philosopher) hopes to come successively closer to the most elegant and concise set of postulates that will describe the full range of quantifiable human experiences. Most physicists, however, believe this goal will never be fully achieved. Probably there will always be a frontier of physical data that are not explained by current physical theory. This is one of the things that makes the life of a theoretical physicist so exciting.

There is a constant interplay among data, laws, and theory. The experimental physicist must have some theory in mind (even if only a very vague one) when deciding which quantities to measure and how to analyze the data in seeking physical laws. The theoretical physicist must always have some physical laws in mind when seeking an elegant theory, even if the immediate goal seems more to find a pleasing abstract system of postulates rather than to explain any particular set of physical laws.

In some sense, the goal of physics is to approach a theory that describes the nature of the underlying reality we call the universe. Most contemporary physicists would agree with the following statement of the physicist and philosopher, Werner Heisenberg.

> I believe that the simplicity of natural laws has an objective character, that it is not just the result of thought economy If nature leads us to mathematical forms of great simplicity and beauty -- by forms I am referring to coherent systems of hypotheses, axioms, etc. -- to forms that no one else has previously encountered, we cannot help thinking that they are "true," that they reveal a genuine feature of nature. It may be that these forms also cover our subjective relationship to nature, that they reflect elements of our own thought economy. But the mere fact that we could never have arrived at these forms by ourselves, that they were revealed to us by nature, suggests strongly that they must be part of reality itself, not just of our thoughts about reality.
> Quoted in Howard Gardner's *Frames of Mind: The Theory of Multiple Intelligences*.

1.3 CAN'T YOU SPEAK PLAIN ENGLISH?

Physical laws are most concisely expressed in the form of algebraic relationships. More complex physical laws are best expressed in more complex mathematical languages, including those of vectors, matrices, complex numbers, and the calculus. Some people find these mathematical languages easy to learn and fun to use. Other people find them almost incomprehensible. If you are a person who finds mathematics confusing and forbidding, you have probably wondered why physicists cannot simply express their ideas in plain English.

It is possible (though not always easy) to talk about some of the concepts of physics in everyday language, but mathematical concepts lie at the heart of all physical theories. To learn about physics without reading any mathematical expressions is somewhat like learning about poetry without reading any poems. Physics took its modern form when physicists began to make quantitative measurements and to summarize their data in mathematical laws. Although these laws may have a simple and elegant expression in mathematical symbols, many of them have a very confusing and unreadable form when put into everyday language.

This book uses only simple algebraic expressions. You can follow all the mathematical logic and language we use here if you are familiar with basic high-school algebra. Although most of the content of this book is in plain English, we will have occasions in some of the chapters to rely on formulas. Don't be intimidated by these algebraic "sentences," but do take your time in reading them. An incredible amount of information is packed into even a simple algebraic equation. It may take some thought to understand the expression, but the equation serves as a convenient shorthand for remembering and using the relationship.

The language of physics differs from everyday language, even when ideas are written in words. In defining physical quantities and theoretical concepts, the physicist uses words in a very careful manner. Some words from everyday language are used with a meaning that is more precise than the usual meaning. New words are created to describe some quantities or concepts that have no everyday names. The resulting density of information may be confusing. Several readings may be required to extract all the information packed into a sentence or a paragraph. You may have to pause frequently and ponder the implications and meanings of some statements. It will take you a while to become accustomed to this style of reading, but you will find that scientific language suddenly becomes much more comprehensible when you discover the ability to "shift gears" into this more intense style of reading.

1.4 FUNDAMENTAL PHYSICAL QUANTITIES

A physical quantity is anything that can be measured. The laws of physics are expressed in terms of physical quantities such as energy, length, time, density, temperature, speed, acceleration, force, and so on. When a physicist performs an experiment, he or she makes careful measurements of the physical quantities that seem pertinent (on the basis of some theory, even if only a tentative one). These measurements are expressed as multiples of well-defined units. For example, the experimenter may measure the time between two events as 37.2 seconds. The unit "second" has a precise meaning, so the time interval of 37.2 seconds is unambiguous -- that is, other physicists will know exactly how much time 37.2 seconds represents.

Obviously, there are a great number of different physical quantities, and a great variety of units in which each quantity can be expressed. However, it is possible to define a small set of fundamental physical quantities, with all the other physical quantities expressed in terms of these fundamental quantities. For example, if length and time are defined as fundamental quantities, then speed can be defined as the distance traveled (length) divided by time. A set of four fundamental physical quantities proves to be sufficient for almost all measurements. (Other fundamental quantities are needed for measurements of certain properties of subatomic particles, but we need not worry about that now.) If we define a unit for each of the four fundamental quantities, we then obtain a unique system of units that can be used to describe any physical measurement. The choice of which quantities to regard as fundamental is somewhat arbitrary. For our purposes, we take length, time, mass, and charge as the four fundamental physical quantities.

Length

The standard unit of length is the meter. For many years, the meter was defined as the distance between two engraved marks on a particular bar made of platinum-iridium alloy, as measured at a temperature of 0°C. This bar is kept at the International Bureau of Weights and Measures near Paris. There is an obvious disadvantage to defining a fundamental unit in such a way. Something might happen to the bar, leaving future physicist with no way to check measurements against the fundamental definition of the unit.

The meter was redefined in 1960 in terms of the wavelength of light from a source containing krypton gas. One meter is now defined as the length that includes 1,650,763.73 wavelengths of the orange light emitted by a light source containing the gas krypton-86. To an incredible degree of accuracy, this length is equal to the distance between the two marks on the platinum-iridium bar. Thus the definition of the meter was changed without changing the length of the meter. However, it is now possible for any experimenter anywhere in the world to check measurements against the basic definition of the meter

Time

The standard unit of time is the second. The second was once defined as 1/86,499 of the mean solar day. However, the motion of the earth does change slightly over the centuries, so that the mean solar day is not quite constant. In order to make the definition of the second a true standard (that is, invariant), the second was redefined in 1960 as the fraction 1/3l,5S6,925.97474 of the tropical year 1900. In order to obtain a more universal and reproducible standard, the second was again redefined in 1967 in terms of the frequency of radiation from a cesium atom. One second is now 9,192,631,770 periods of the radiation associated with a specific change in the atom of cesium-133.

Mass

The standard unit of mass is the kilogram. The kilogram is defined as the mass of a particular platinum-iridium cylinder kept at the International Bureau of Weights and Measures. It would be desirable to have a universally reproducible atomic standard for mass, similar to the standards for length and time. There is a well-defined unit of mass called the atomic mass unit. It is defined as 1/12 of the mass of an atom of carbon-12. This is useful for comparing the masses of various atoms and molecules. However, we do not yet have the technology to define the standard kilogram in terms of this atomic unit with sufficient accuracy. At least for now, the mass of the platinum-iridium cylinder near Paris remains as the definition of the kilogram.

We have not discussed the meanings of the physical quantities length and time because we all have a good intuitive feeling for their meaning (although we will soon see those intuitive feelings can be misleading). Most of us also have some intuitive feeling for the meaning of mass, but it will be worthwhile to discuss it briefly. The mass of a given object has to do with the amount of matter present in it and is a measure of the resistance of the object to a change in its motion. The more mass an object has, the more difficult it is to change either its speed or the direction in which it is moving. Mass is also related to weight; the greater the mass of an object, the more it weighs. However, mass is a property of the object itself, whereas weight is a physical quantity associated with the interaction between the object and some other object such as the earth.

Consider a block of lead whose mass is 2.77 kilograms. Near the surface of the earth, this lead block has a certain weight (a little over 6 pounds in the English system, a system we will not be using in this course. Science uses the metric system.) That is, the earth pulls on the lead block with a certain force. As the lead block is moved farther and farther away from the earth, the gravitational interaction between the earth and the lead block decreases, and therefore the weight of the lead block decreases. However, the amount of matter in the lead block is unaffected, and therefore its mass remains unchanged. On the surface of the moon, the lead block has a definite weight that is considerably less than its weight on earth. Its mass, however, remains 2.77 kilograms. The experiences of astronauts confirm that mass (resistance to change of motion) is unaffected by such changes in weight. If we imagine the lead block sufficiently far away from all other matter in the universe, its weight would be zero but it would still have a mass of 2.77 kilograms. An astronaut would still have to exert just as much force to change its motion.

The interaction that creates attractive forces between two objects having mass is called gravity. Weight is a measure of in the strength of the gravitational force. You will recall that the existence of such a force is one of the postulates of the theory of classical mechanics. Thus we can explain apples falling from trees or planets orbiting the sun in terms of gravitational forces, but we cannot explain the existence of the gravitational force itself in terms of this theory. (We will see later that the theory of general relativity does provide an explanation of gravity in terms of other postulates.)

Charge

The standard unit of charge is the Coulomb. The Coulomb can be specified in terms of the charge on the electron, which is a subatomic particle having the smallest amount of charge that can readily be isolated. One Coulomb is equal to 6,24l,460,000,000,000,000 times the charge on an electron. Because very large and very small numbers are cumbersome to write and difficult to read, we follow the common practice of writing such numbers in scientific notation (see Appendix A). Thus we say that one Coulomb is equal to 6.24146×10^{18} times the charge on an electron. Charge, like mass, is a property of matter. Matter, as we are familiar with it in our daily lives, is made up of atoms. Atoms, however, are not indivisible; they in turn are made up of more elementary particles known as electrons, protons, and neutrons. Each of these elementary particles has associated with it a specific mass and a specific charge. Whereas there is only one kind of mass, there are two kinds of charge. They are called positive and negative, because their effects can balance one another -- for example, one unit of positive charge and one unit of negative charge combine to produce a net charge of zero. By arbitrary definition, the charge on the electron is called negative. The proton has an equal amount of positive charge. As its name implies, the neutron is electrically neutral -- that is, its charge is zero.

Consider an atom (hydrogen) composed of one proton, one neutron, and one electron. The proton, neutron, and electron combine together to give the atom a certain mass. Because there is only one kind of mass, any combination of particles that individually have mass will produce a combination that also has mass. However, the combination of charges need not itself have charge. Specifically, the negative charge on the electron exactly balances the positive charge on the proton, and the neutron has a charge of zero, so the atom has a net charge of zero. Atoms, as normally found on earth have equal numbers of protons and electrons, and therefore are uncharged. An object formed of such atoms is also uncharged. It is not very difficult, however, for an atom to gain or lose an electron. Ordinary chemical reactions involve the interchange of electrons among atoms. On a larger scale, an object may under some conditions acquire an excess of electrons and become negatively charged. For example, a plastic comb drawn through dry hair will pick up electrons from the hair and thus become negatively charged. An object that has a deficit of electrons (or, to look at it the other way, an excess of protons) is positively charged. For example, the hair is positively charged after the comb has been drawn through it because electrons have been removed.

A basic element of reality involving charge is that objects having like charges (both objects positive, or both negative) will repel each other, whereas objects having unlike charges (one positive and the other negative) will attract each other. You may have noticed that if the comb is brought near the hair in the previous example, the hair will stand up in the direction of the comb. That is, there will be an attractive force between the comb and the hair These interactions between charges play an important role in the structure of matter.

Systems of Units

A wide variety of units have been used at various times and places to measure physical quantities. The units we have just defined for the four fundamental physical quantities form the basis of a system of units called the metric system. Science almost exclusive uses the metric system. All of the units for derived physical quantities used in this book can be expressed as combinations of two or more of the four fundamental units: meters (m), seconds (s), kilograms (kg), and Coulombs (C).

In the realm of atomic and subatomic phenomena, the basic units of the metric system become cumbersome. It is not convenient to describe the dimensions of an atom in terms of meters, nor to discuss atomic or nuclear masses in terms of kilograms. Other more appropriate units are occasionally defined as needed for specific purposes. However, the metric system does include a handy set of prefixes that makes it easy to create larger or smaller units. For example, one centimeter (1 cm) equals 1/100 meter, and one microsecond (1 ms) equals 0.000001 seconds. Note that the fundamental unit of mass, the kilogram, is equal to 1000 grams (or 1 kg = 1000 g). I

In most of the world, the metric system of units is now becoming standard for everyday use. In the United States, however, the English system of units is still commonly used, with length measured in feet, force in pounds, and so forth. This system is poorly suited to scientific work, and we will generally not use it

> *The scientist does not study nature because it is useful; he studies it because he delights in it, and he delights in it because it is beautiful. If nature were not beautiful, it would not be worth knowing, and if nature were not worth knowing, life would not be worth living.*
>
> – Henri Poincare

Summary

The purpose of physics is the coherent description of our experiences of the physical world. Quantitative descriptions of phenomena are obtained by measuring physical quantities. Physical laws are relationships between or among physical quantities and are usually expressed in algebraic equations. A physical theory correlates a number of physical laws in the sense that the laws can be derived from the fundamental postulates of the physical theory. In that sense, the phenomena described by the laws can be understood. The ultimate criteria by which a physical theory is judged is its comprehensiveness and the correspondence of its logical deductions with the physical data. However, such criteria as simplicity, elegance, and beauty play important roles when physicists must choose between two theories that are approximately equal in their abilities to correlate known physical laws. A physical quantity is anything that can be measured. A measurement describes the quantity in terms of a number and a well-defined unit. Any measurement can be expressed in terms of some combination of one or more of the four fundamental units defined for four fundamental physical quantities. For our purposes, the four fundamental physical quantities and their units are length expressed in meters, time in seconds, mass in kilograms, and charge in Coulombs.

Important concepts

Physical quantity; physical data; physical law; physical theory; postulates; fundamental physical quantity; metric system of units; mass; charge.

Questions

1. Distinguish between quantitative and qualitative properties. List ten properties of objects, and tell which are quantitative and which are qualitative.
2. Distinguish among physical quantities, physical data, physical laws, and physical theories.
3. Explain why a physical theory can be disproved but never can be proved. How is a scientific theory disproved?
4. What is meant by the predictive power of a physical theory?
5. What constitutes an explanation of a physical event? Explain how our understanding of the cause of a physical event is fundamentally limited.
6. What is meant by a physical quantity? What is meant by derived and fundamental physical quantities?
7. Galileo, in carrying out some of his experiments, used his own pulse as a way of measuring time. What are some of the advantages and disadvantages of the pulse-beat as a unit of time compared to the SI time unit, the second?
8. What do you consider the most important characteristic of a standard used to fix the value of a unit? What are some other important characteristics?

   The following questions are of a more general nature. They have no single correct answer and are just something for you to think about. When possible, they are best answered in conversation with others.
9. What might Albert Einstein have meant when he said that "the eternal mystery of the universe is its comprehensibility"?
10. Heinrich Hertz, speaking of the mathematical formulas of the theory of electromagnetism, said, "They know more than we do and more than those who discovered them; they will give out more information than has ever been put into them." What do you think he meant?
11. Look up the definitions of inductive and deductive logic. What roles do these two types of reasoning play in the formulation of physical laws and physical theories?
12. What is meant when a theory, an opinion, or a procedure is criticized as being unscientific?
13. Discuss the meaning of the following statement by Galileo "In the end our controversies concern the explorable world and not what is written on paper. Let us proceed to demonstration, to observations, and to experiments."
14. If physical theories are someday able to "explain" the totality of physical data, then we can say that we have arrived at the Truth in nature. Discuss the validity of this statement in terms of (a) limitations on our ability to gather data, (b) the existence of non-quantitative reality, and c) the basic postulates of physical theories.
15. A physical theory explains how things happen, not why they happen. Discuss this statement.

2 CONSERVATION LAWS AND ENERGY

Among the most powerful physical laws created by physicists are those called conservation laws. In physics, it is often convenient to imagine an isolated system, a restricted part of the universe that has no interaction with any other part of the universe. Events within the isolated system neither influence the rest of the universe nor are influenced by it. Obviously, a perfectly isolated system can never be created in practice, but many actual systems are sufficiently isolated to permit experimental tests of our theoretical conclusions about isolated systems. A conservation law states that some particular physical quantity associated with an isolated system does not change its value over time. This physical quantity always has the same total or net value for the isolated system, regardless of what is going on within the system. In this case, the quantity is said to be conserved.

Conservation of mass seems an obvious conservation law. If you cut an object into pieces, the sum of the masses of the pieces is the same as the original mass of the object. If no mass is allowed to enter or leave the system under observation, then the total mass of the system remains constant. Early chemists were very puzzled by certain apparent exceptions to this law. Some substances gain or lose mass when heated or treated in other ways. Eventually, the chemists learned about gases and realized that the substances were gaining mass from the air or losing mass to the air in the form of gases. When they isolated their experimental systems so that gases could neither enter nor leave, they found that their results were indeed consistent with the law of conservation of mass and in the nineteenth century this law was thought to be correct. However, in 1905, Einstein's theory of relativity showed that it is not true. Mass can be created or destroyed. We will come back to this when we discuss relativity.

The discussion of conservation laws leads to another physical quantity -- energy. Energy serves as the central unifying concept of physics, and it links physics to all other fields of science. In a very real sense, modern physical theories explain all experiences in terms of just two basic concepts, energy and space-time. The universe seems to be an intricate pattern of energy distributed through the interlinked dimensions of space and time. Before discussing energy, we consider some of the basic conservation laws well known to physicists at the beginning of the twentieth century.

2.1 CONSERVATION OF CHARGE

In an electrically isolated system, charge is conserved. That is, the net charge of the system is not altered as a result of the interactions within the system. However, charge can be exchanged between objects within the system.

The net charge of an isolated system is constant.

Consider as an isolated system the plastic comb and dry hair mentioned in Chapter l. Initially, the comb and the hair each have a net charge of zero. The net charge of the system must remain zero after the comb is drawn through the hair. During the interaction, negatively charged electrons are transferred from the hair to the comb, giving the comb a net negative charge and leaving the hair with a net positive charge. The negative charge of the comb is exactly the same amount as the positive charge of the hair, so that the net charge of the system remains zero.

In general, conservation laws are useful because they express regularities that can be detected even when we know very little about the details of what is happening during an interaction. Without knowing anything about the process by which a net positive charge was created on some object, we can state with great confidence that the process must have involved the creation of an equal quantity of net negative charge somewhere else in the system.

The law of conservation of charge often is stated in another (equivalent) form:

Charge can neither be created nor destroyed.

However, it is important to remember that the law applies to the net charge of an isolated system. Under certain circumstances, a neutron (with net charge of zero) can change into a pair of particles: a positively charged proton and a negatively charged electron. Although negative and positive charges seem to be created during this reaction, no net charge is created. The proton has a positive charge equal to the negative charge on the electron. Therefore, the net charge of the system remains zero.

2.2 CONSERVATION OF MOMENTUM

A physical quantity is anything that can be measured. If a physical quantity can be described by a magnitude (number) and a unit it is called a scaler. For example, the mass of a particular object is given as 2.31 kilograms. This completely describes the mass of the object. Mass is a scalar physical quantity as are the other fundamental physical quantities – time, length, and charge.

However there is another class of physical quantities called vectors. To completely describe a vector, a direction must be specified in addition to a magnitude and a unit. An example of a vector physical quantity is velocity. An automobile traveling at a speed of 93 km/s (speed is a scalar physical quantity) is headed due north. Its velocity is 93 km/s due north. Velocity is a vector physical quantity.

Momentum is defined as the mass of an object times its velocity. The product of a scalar and a vector is a vector and therefore momentum is a vector physical quantity, $\mathbf{p} = m\mathbf{v}$. (Note that in equations, the symbol for a vector is indicated by using bold type. Thus $\mathbf{p}$ and $\mathbf{v}$ are bold while m is not.) The fact that the momentum of an isolated system is constant can be derived from the postulates of Newtonian mechanics.

The total momentum of an isolated system is constant in both magnitude and direction.

The law of conservation of momentum is one of the important laws in physics: The law applies only to the total momentum of the system. As a result of interactions within the system (for example, collisions between objects), the momentum of an individual object may change. However, the sum of the momenta of all the objects in the system remains unchanged. In applying this law, it is essential to remember that momentum is a vector quantity. Therefore, the direction as well as the magnitude of the momentum must be taken into account. This can best be understood by looking at some examples.

Example 2.1

A ball of clay with a mass of l kg is moving left to right with a speed of 10 m/s. It collides with another ball of clay having a mass of 3 kg that is at rest before the collision. The two balls stick together after the collision (see Figure 2.1). What is the velocity of the stuck-together balls after the collision?

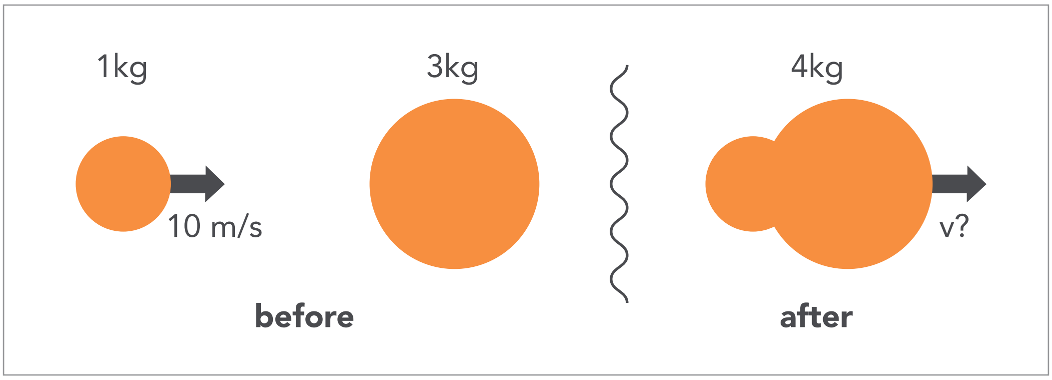


Figure 2.1

Solution

The momentum **p** of an object is defined by the relation **p** = m**v**: If we use the subscript l to represent quantities of the first ball before the collision, we can write

$$\mathbf{p}_1 = m_1 \mathbf{v}_1 = (1 \text{ kg}) \times (10 \text{ m/s left to right}) \text{ or } \mathbf{p}_1 = 10 \text{ kg-m/s left to right}$$

Similarly, using the subscript 2 to represent the second ball before the collision, we have

$$\mathbf{p}_2 = (3 \text{ kg}) \times (0 \text{ m/s}) = 0.$$

The total momentum of the system before the collision is

$$\mathbf{p}_1 + \mathbf{p}_2 = (10 \text{ kg m/s left to right}) + 0 = 10 \text{ kg m/s left to right}$$

From the law of conservation of momentum, we know that the total momentum of the system after the collision must be the same. Using the subscript 3 to represent the combined balls after the collision, we can write

$$\mathbf{p}_3 = 10 \text{ kg-m/s left to right} = m_3 \mathbf{v}_3 \text{ where } m_3 = (1 \text{ kg} + 3 \text{ kg}) = 4 \text{ kg}.$$

$$\mathbf{v}_3 = \mathbf{p}_3 / m_3 = (10 \text{ kg m/s left to right}) / 4 \text{ kg, or, cancelling the unit kg,}$$

$$\mathbf{v}_3 = 2.5 \text{ m/s left to right.}$$

Thus, without knowing anything about the details of the collision except the fact that the balls stick together, we are able to conclude from the conservation of momentum law that the combined balls after the collision must be moving with a velocity of 2.5 m/s left to right. Our confidence in the conservation law comes from a vast number of experimental data. Actually, we have made one other assumption about the collision, the assumption that none of the balls has a spin before or after the collision.

Momentum is also associated with spinning as well as linear motions, but that is far more complex, so we will avoid this complication. The conservation law does hold for interactions involving spin when the angular (rotational) momentum is properly accounted for.

Example 2.2

Suppose that the 3-kg clay ball of Example 2.1 is not at rest initially but instead has a velocity of 2 m/s right to left before the collision (see Figure 2.2). How will this change the result? Will the stuck-together ball move to the right or to the left after the collision? (You should be able to guess the answer to that question before working through the solution in detail.) What is the velocity of the ball after the collision?

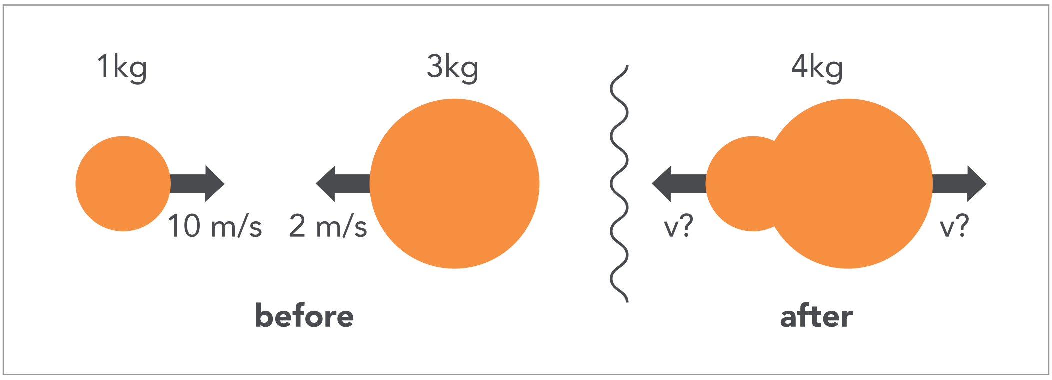


Figure 2.2

Solution

As in Example 2.1, we have $\mathbf{p}_1 = m_1\ \mathbf{v}_1 = 10$ kg m/s left to right.

But now we also have

$$\mathbf{p}_2 = m_2\ \mathbf{v}_2 = (3 \text{ kg}) \text{ x } (2 \text{ m/s right to left}) = 6 \text{ kg m/s right to left.}$$

How should we add these two vector quantities to find the total momentum before the collision? Because the directions of the vector quantities are opposite to each other, we can simply regard one direction as the negative of the other. Choosing left to right as positive, we can apply conservation of momentum to get

$$\mathbf{p}_3 = \mathbf{p}_1 + \mathbf{p}_2 = +\ 10 \text{ kg m/s} - 6 \text{ kg m/s} = +\ 4 \text{ kg m/s or } 4 \text{ kg m/s left to right.}$$

$$\mathbf{p}_3 = m_3\ \mathbf{v}_3 \text{ or } \mathbf{v}_3 = \mathbf{p}_{3/}\ m_3 = (4 \text{ kg m/s left to right})/4\text{kg} = 1 \text{ kg m/s left to right.}$$

Note that the higher speed of the smaller ball gives it the larger momentum, so that the combined balls after the collision move left to right.

Example 2.3

A steel ball with a mass of 1 kg is moving with a velocity of 2 m/s left to right. It strikes another 1-kg steel ball that is initially at rest. The two balls bounce apart (see Figure 2 3). What are the velocities of the balls after the collision?

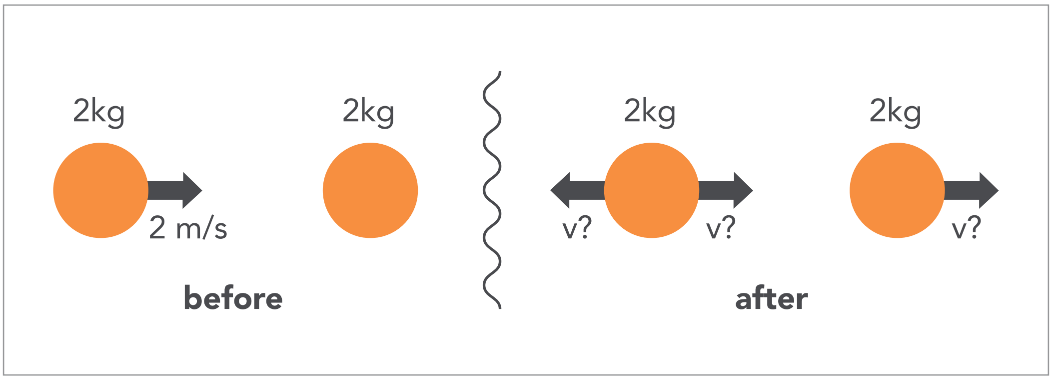


Figure 2.3

Solution

The second ball is motionless before the collision, so its momentum is zero The momentum of the first ball before the collision is 4 kg m/s left to right and this is the total momentum of the system After the collision, the sum of the momenta of the two balls must also be 4 kg m/s left to right

> 4 kg m/s = $\mathbf{p}_3 + \mathbf{p}_4 = m_3\, \mathbf{v}_3 + m_4\, \mathbf{v}_4$, where 3 and 4 represent the two balls after the collision.

We have one equation with two unknowns and therefore cannot solve for the final velocities. There are an infinite number of solutions that will conserve momentum. However, when this experiment is done repeatedly, making sure that the collision is perfectly head-on, the result is always the same: the originally at-rest ball moves with a velocity equal to the striking ball and the striking ball is at rest after the collision (see Figure 2.4).Thus there must be some other regularity (some other physical law) in addition to conservation of momentum that applies here. That other law is conservation of energy.

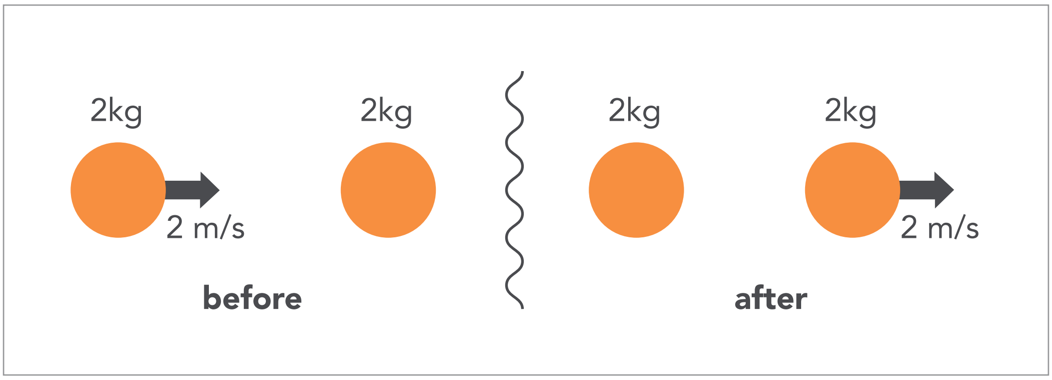


Figure 2.4

2.3 ENERGY

Energy is a word we all use. It is a common word as well as a basic term in physics. We all have an intuitive feeling for its meaning, but our use of the term in physics requires a proper definition. Energy is often defined as "the capacity to do work," with work being associated with the movement of a mass upon which a force is exerted. Ultimately, a variety of different definitions must be used to describe how energy is measured in its various forms. The usefulness of a general concept of energy lies in the fact that a given quantity of energy in one form does prove to be equivalent to the same quantity of energy (computed from a different formula) in a different form. The following discussion should help to clarify the nature of the modern concept of energy.

Energy appears in many different forms such as electrical energy, potential energy, kinetic energy, solar energy, heat energy, chemical energy, radiant energy, nuclear energy, and mass energy. These various forms of energy are not really as distinct from one another as the variety of names implies. In fact, in modern physics, it often proves convenient to group all these forms into three basic categories: kinetic energy, rest-mass energy, and radiant energy.

Kinetic Energy

Kinetic energy is the energy an object has because of its motion.

If an object of mass m is moving with a speed that is small compared to the speed of light, then the kinetic energy of the object is equal to one-half its mass multiplied by the square of its speed: Thus energy has the unit kg m^2/s^2. (Note that kinetic energy depends on speed rather than velocity. Kinetic energy, and energy in general, is a scalar physical quantity.)

$$E_k = \frac{1}{2} mv^2$$

The theory of relativity predicts (and experiment confirms) that this equation is valid only when v is much smaller than the speed of light (which is about 3×10^8 m/s or 186,000 miles per second). Because energy is such an important physical quantity, the derived unit, kg m^2/s^2 is given a special name, the Joule (abbreviated J and pronounced 'jool').

Example 2.4

An object of mass 3 kg is traveling with a speed of 5 m/s. what is its kinetic energy?

Solution

$$E_k = \frac{1}{2} mv^2 = \frac{1}{2} (3 \text{ kg}) (5 \text{ m/s})^2 = = \frac{1}{2} (3 \text{ kg})(25 \text{ m}^2/\text{s}2) = 37.5 \text{ kg m}^2/\text{s}2 = 37.5 \text{ J}$$

Using the concept of kinetic energy, go back and look at our earlier collision examples again In addition to the kinetic energy of the moving objects, kinetic energy appears in another form as well. We can regard matter as composed of great numbers of tiny atoms. Those atoms are in motion even when the object is at rest. As the temperature of the object increases, the speed of its atoms increases. In fact, the temperature is nothing more than a measure of the average kinetic energy of its individual components – its atoms or molecules.

In example 2.2, the initial speed of the 1-kg object was 10 m/s and the final speed of the combined 4-kg object was 1 m/s. Thus the initial kinetic energy of the objects is

$$E_k = \frac{1}{2}mv^2 = \frac{1}{2}(1\ \text{kg})\ (10\ \text{m/s})^2 = 50\ \text{J}$$

The final kinetic energy is

$$E_k = \frac{1}{2}mv^2 = \frac{1}{2}(3\ \text{kg})\ (5\ \text{m/s})^2 = \frac{1}{2}(4\ \text{kg})\ (1\ \text{m/s})^2 = 2\ \text{J}.$$

Kinetic energy does not appear to be conserved in this collision. Collisions of this type are called inelastic collisions. The kinetic energy that appears to have been lost has gone into the internal kinetic energy of the atoms or molecules of the object. That is, the temperature (or heat energy) increases as a result of the collision. In addition to momentum, energy is conserved in this collision.

In example 2.3, the initial kinetic of the object is $\frac{1}{2}(1\ \text{kg})\ (2\ \text{m/s})^2 = 2$ J. In this collision between steel balls, the colliding objects are not compressed, and no kinetic energy of motion is converted into heat energy. The temperature of the objects is not affected by the collision. In this case the kinetic energy of motion is conserved. Such collisions are called elastic collisions. (See Figure 2.4.) It is clear that both momentum and energy is conserved in this interaction.

The physical theory that describes the relationship between heat energy and the other properties of large objects is called thermodynamics. This theory was developed slowly over a long period of time, with contributions from many physicists, engineers, and chemists. Newton and many of his contemporaries believed heat to be due to the motions of atoms. This concept was later rejected However, as the concept of energy was gradually developed in the middle of the nineteenth century, physicists returned to the view that heat energy is associated with the kinetic energy of the individual atoms. Thus the theory of thermodynamics represents an extension of the theory of classical mechanics to provide an explanation of phenomena involving heat.

Rest-Mass Energy .

Rest-mass energy is the energy that an object has because of its mass.

In the late nineteenth century, the theories of classical mechanics and thermodynamics provided physicists with a view of a universe in which matter (made up of atoms) occupies space and has mass. This matter and the space in which it is distributed could be called the "real substance" of the universe. In addition, a physical quantity called energy can be associated with any moving object (including an atom), but this energy seems to be more of an abstract concept – an invention of the physicist's mind – rather than a part of the material reality of the universe. Energy was seen as a sort of process or condition that neither takes up space nor has mass.

When Albert Einstein (1879 – 1955) proposed his first theory of relativity in 1905, he showed that his theory leads to a surprising relation between mass and energy. Even when an object is at rest, it can be regarded as having an energy that is related to its mass alone. They are related by the famous equation

$$E = mc^2$$

where c represents the speed of light in a vacuum:

$$c = 3 \times 10^8 \text{ m/s, so } c^2 = 9 \times 10^{16} \text{ m}^2/\text{s}^2.$$

This simple formula has very important implications about the nature of the universe. It says that mass and energy are really the same thing and the two quantities are related to each other by a conversion factor that is equal to the speed of light squared. Clearly, the speed of light is not just a particular speed, but it is somehow a very important and basic part of the workings of the universe. Furthermore, mass is simply a quantity that we can measure when a lot of energy is compressed into a small space. Any form of energy has an associated mass, but only the very concentrated form of energy that we call matter has enough mass to be readily measured.

As we will see in chapter 5, the theory of relativity shows that the measured mass of an object depends on its speed relative to the observer. This is because of the mass associated with the kinetic energy of the moving object. Therefore, we will use the term rest mass to mean the mass of an object at rest with respect to the observer. The mass is measurably different from the rest mass only if the object is moving a speed of about l0 percent of the speed of light, or greater -- that is about 3×10^7 m/s or 20,000 miles per second or more.) Just to keep things straight, we shall use a subscript o to indicate rest mass, so we write the rest-mass energy of an object as

$$E_o = m_o c^2.$$

In a sense, rest mass represents dormant energy inherent in an object -- energy that could be converted under appropriate conditions to any of the other forms of energy. For this reason, rest-mass energy can also be called potential energy. However, the conversion factor c^2 is so large that a very large change in rest-mass energy corresponds to a very small change in rest mass. Until physicists began working with the huge amounts of energy per unit mass involved in nuclear reactions, the changes in rest mass were such a tiny fraction of the mass involved they could not be measured. Therefore, in most problems it is more convenient to speak of changes in the potential energy of the system rather than speaking of changes in the rest mass of the system.

To understand the role of c^2 in the equation $E_o = m_o c^2$., consider the way in which we convert a length measurement from one unit to another. Suppose a distance is measured as 5 yards and we wish to know the distance in feet. We use the conversion factor to make the conversion as follows:

5 yards = 5 yards x 3 feet/1 yard = 15 feet.

We have not changed the length because 3 feet/1 yard =1 because 3 feet and 1 yard is the same length. In much the same way, c^2 serves as a conversion factor to change the mass unit kilograms to the energy unit joules.

Example 2.5

A proton has a rest mass of 1.66×10^{-27} kg. What is its rest-mass energy?

Solution

$$E_o = m_o c^2. = (1.66 \times 10^{-27} \text{ kg}) \times (9 \times 10^{16} \text{ m}^2/\text{s}^2) = 1.49 \times 10^{-10} \text{ J}.$$

If you have trouble doing arithmetic with scientific notation, review Appendix A.

It is important to recognize that the conversion factor c^2 is an extremely large number, so that a very small amount of rest mass corresponds to a relatively large amount of rest-mass energy. This will be more obvious if we consider an example on a larger scale.

Example 2.6

The mass of a dime is about 0.002 kg. What is the rest-mass energy of a dime?

Solution

$$E_o = m_o c^2. = (0.002 \text{ kg}) \times (9 \times 10^{16} \text{ m}^2/\text{s2}) = 1.8 \times 10^{14} \text{ J}.$$

This is approximately the amount of heat energy that can be obtained by burning 30,000 barrels of oil! You can see why nineteenth-century physicists failed to detect the change in rest mass that occurs when energy is released in ordinary chemical reactions and mechanical events.

Radiant Energy

Radiant energy is the energy associated with electromagnetic radiation.

Both kinetic energy and rest-mass energy are associated with matter. That is, we can view these two forms of energy as quantities that are properties of objects (even though the objects may be as tiny as individual atoms or subatomic particles). However, there is a third form of energy that cannot readily be associated with objects. Radiant energy is more like energy transferred between objects. There are many forms of electromagnetic radiation, including radio waves, microwaves, infrared radiation, visible light, ultraviolet radiation, X rays, and gamma rays.

The theory of electromagnetism was created by James Clerk Maxwell (1831 – 1879) in the 1860s. It provided an explanation of such phenomena as electricity, magnetism, and light, something that Newton's classical mechanics could not do. Maxwell regarded radiant energy as a form of kinetic energy due to wavelike motions of an all-pervading, universe-filling substance called the "ether." As we shall see, twentieth-century physicists built their theories upon the postulate that no such ether exists. Although we do sometimes regard light as being composed of particles, we find that light particles have a rest mass of zero In other words, all of the energy in electromagnetic radiation is this special form of radiant energy, and none of is associated with a rest mass. The important factor "c" turns up here again as the speed with which electromagnetic radiation travels through empty space. We will discuss electromagnetic radiation in chapter 8 and postpone the development of quantitative expressions for radiant energy until that discussion. However, the unit of radiant energy is the joule, the same unit that is used to express kinetic energy and rest-mass energy.

2.4 CONSERVATION OF ENERGY

The concept of energy is useful because it makes possible the statement of a very powerful conservation law – the conservation of energy.

The total energy of an isolated system is constant.

As in the other conservation laws we have discussed, it is total energy of the system that is conserved. Energy may change from one form to another and the energy associated with a particular object in the system can change. But if we sum the kinetic energy, the rest-mass energy, and the radiant energy of the system at any given moment, we find that, if the system is completely isolated, the sum is constant.

The law of conservation of energy is often expressed in the following form: Energy can be neither created nor destroyed. It is the great success of physicists in finding a general concept of energy that makes possible this very fundamental conservation law. And it is the great generality of this law that makes energy the central unifying concept of physics.

Notice that the law of conservation of mass is simply a specialized form of the law of conservation of energy. In most situations, the amount of energy in the form of kinetic energy and/or radiant energy is infinitesimal compared to the rest-mass energy of the system, changes in these forms of energy produce immeasurable changes in rest mass and mass seems to be conserved. It is only when the masses involved are very small (atoms and elementary particles) and kinetic energy and/or radiant energy are large (as in nuclear physics) that measurable changes in rest mass occur.

The law of conservation of energy is also called the first law of thermodynamics. Although this law states that energy cannot be created or destroyed, energy can be converted from a form that is useful for doing work to a form that is not useful. For instance, when gasoline is burned, the total energy of the system is conserved. However, the amount of energy available to do work decreases. Another basic law of nature states that changes in an isolated system always occur in such a way that energy is converted to forms with less capacity to do work. In an isolated system, the exhaust gases will never reassemble to form gasoline, even though such a process would not violate the conservation of energy. That is no spontaneous event is ever observed in which the system's ability to do work increases. This additional constraint observed in all phenomena is called the second law of thermodynamics, and it is a very important law in physics. For one thing, it is the only basic physical law that gives a preferred direction to time. All of the other basic laws imply that movement backward through time should be equivalent to movement forward through time. The second law of thermodynamics will not play a major role in our discussion of twentieth-century physics.

It is easier to apply the law of conservation of energy than it is to apply the law of conservation of momentum, because energy is a scalar quantity rather than a vector quantity. Energy has only magnitude; it has no direction in space. (This is obvious in the case of rest-mass energy.) Therefore, we need only add up the individual energies to use the conservation law. We need not worry about how to add vectors pointing in different directions, as would be the case for collisions in two-dimensions.

Example 2.7

In an explosion, rest-mass energy is converted into kinetic and radiant energy at such a high rate that a shock wave is formed in the air. When 1 ton of TNT explodes, approximately 4.2×10^9 J of rest-mass energy is converted to kinetic energy (associated with gases and particles of the exploding material) and radiant energy. What is the resulting decrease in the rest mass of the matter in the system?

Solution

From the law of conservation of energy, we know that the lost rest-mass energy must equal the sum of the kinetic and radiant energy that has appeared. Using the equation for rest-mass energy, $E_o = m_o c^2$. Physicists use the Greek letter delta, Δ, to indicate a change in some quantity. Thus we indicate the change in rest-mass energy by ΔE_o and the change in rest mass by Dm_o.

$$\Delta E_o = \Delta m_o c^2 \text{ or } \Delta m_o = \Delta E_o / c^2 = (4.2 \times 10^9 \text{ J}) / (9 \times 10^{16} \text{ m}^2/\text{s}^2)$$
$$\Delta m_o = 4.67 \times 10^{-8} \text{ kg}$$

This change in mass is about equal to the mass of a grain of salt! It represents a loss of only 0.000000005 percent of the total rest mass of the system. Again we can see why such changes in rest mass were not detected by nineteenth-century chemists or physicists. It wasn't until nuclear reactions were discovered in the twentieth century that measurable rest-mass could be produced.

Example 2.8

A berilium-8 nucleus is unstable. After a tiny fraction of a second, it will decay to two helium-4 nuclei. Consider this reaction in the light of the three conservation laws discussed in this chapter.

Solution

First, consider the law of conservation of charge. The berilium-8 nucleus has 4 protons and 4 neutrons and therefore has a charge of +4. A helium-4 nucleus consists of 2 protons and 2 neutrons. Four protons before the reaction and four after. Charge is conserved.

Next consider the law of conservation of momentum. The berilium-8 nucleus was at rest before the reaction, so the total momentum of the system was zero before the reaction. After the reaction, both the helium-4 nuclei have nonzero velocities, and so each has momentum. However, momentum is a vector quantity. The only way the total momentum can be zero after the reaction is for each helium-4 nuclei are moving in opposite directions with the same magnitude of momentum. Since they have the same mass, they must also have the same speeds. The experimental observations are consistent with this prediction,.

Finally, consider the law of conservation of energy. Because there was no kinetic energy or radiant energy before the reaction, the total energy of the system is rest-mass energy. The kinetic energy of the two helium-4 nuclei must come from lost rest-mass energy. Precise measurements of the rest-masses and the kinetic energy are consistent with the predictions of conservation of energy.

Example 2.9

Our sun gives off 3.9×10^{26} J of energy every second. The source of this energy is nuclear reactions deep in the interior of the sun. In these nuclear reactions, rest-mass energy is converted into kinetic energy and radiant energy. As this energy is transferred from the interior to the surface of the sun, the kinetic energy is converted to radiant energy. What is the decrease in the rest mass of the sun each second? The sun has been shining for about 5×10^9 years (5 billion years) and scientists estimate it will continue to shine for another 5 billion years or so. If the sun had a rest mass of 2×10^{30} kg when it began to shine, what fraction of its original mass will be lost over its lifetime? (1 year = 3.15×10^7 s.)

Solution

From the law of conservation of energy, the amount of radiant energy given off in one second must be exactly equal to the decrease in rest-mass energy during one second.

$$\Delta E_o = 3.9 \times 10^{26} \text{ J in one second.}$$

$$\Delta m_o = \Delta E_o / c^2 = (3.9 \times 10^{26} \text{ J}) / (9 \times 10^{16} \text{ m}^2/\text{s}^2) = 4.34 \times 10^9 \text{ kg in one second.}$$

This is a tremendous amount of mass; it is the mass of 4.75 million tons of matter!

To determine how much rest mass will be lost over the lifetime of the sun, we must first convert 10 billion years to seconds.

$$10^{10} \text{ years} = 10^{10} \text{ years} \times 3.15 \times 10^7 \text{ s/yr} = 3.15 \times 10^{17} \text{ s.}$$

The total decrease in rest mass over this interval of time is

$$\Delta m_o = (4.34 \times 10^9 \text{ kg/s}) \times (3.15 \times 10^{17} \text{ s}) = 1.37 \times 10^{27} \text{ kg} = 0.00137 \times 10^{30} \text{ kg.}$$

Thus the fractional change in the sun's rest mass over its lifetime is

$$(0.00137 \times 10^{30} \text{ kg}) / (2 \times 10^{30} \text{ kg}) = 0.00069$$

Thus, although the sun will lose a tremendous amount of mass over its lifetime due to nuclear reactions in its interior (a mass equal to more than 200 earths), the mass loss will be less than 0.1 percent of its original mass.

Summary

In an isolated system, the total amount of certain physical quantities within the system remains constant in time. Such physical quantities are said to be conserved. Examples of conserved physical quantities are charge, momentum, and energy. Some physical quantities have only magnitude (mass and charge, for example) and are called scalars. Others have both magnitude and a direction in space (velocity and momentum, for example) and are called vectors. Because momentum is a vector rather than a scalar physical quantity, the direction as well as the magnitude of the momentum must be considered in applying the principle of conservation of momentum.

In the most general usage, energy is the capacity for producing an effect and is a scalar physical quantity. In twentieth-century physics, it is useful to discuss energy in terms of three general categories. Kinetic energy is the energy an object or particle possesses because of its motion. Rest-mass energy is the energy an object or particle possesses when it is at rest. The rest-mass energy of the object or particle is a direct measure of its rest mass according to the relationship $E_o = m_o c^2$, where c is the speed of the light in a vacuum. Changes in the potential energy of a system are changes in the rest-mass energy of the system. Radiant energy is the energy associated with electromagnetic radiation (radio, infrared, visible, ultraviolet, X-ray, or gamma-ray). Kinetic and rest-mass energies are properties of material objects, whereas radiant energy is energy transferred between material objects.

Energy may be transferred between objects or converted from one form to another, but the total energy in an isolated system is constant. Any amount of energy may be expressed as an equivalent mass, but the conversion factor c^2 is so large that a very large amount of energy corresponds to a very tiny amount of mass. Thus, for most normal interactions there is no detectable change in the total rest mass of an isolated system. The change in rest mass is only measurable in the case of nuclear reactions, the subject matter of chapters 14 through 16. However any change in the total kinetic energy and/or total radiant energy must, by conservation of energy, result in a change in rest-mass energy even if the change in mass may be undetectable.

Important concepts:

Isolated system; conservation law; conserved physical quantity; conservation of mass; conservation of charge; conservation of momentum; energy; work; kinetic energy; joule; thermodynamics; rest-mass energy; potential energy; radiant energy; conservation of energy.

Questions

1. Explain in general terms what a conservation law is.
2. State and explain the principle of conservation of charge and the principle of conservation of momentum.
3. A high-powered rifle has a mass of 5 kg. It fires a bullet having a mass of 0.015 kg. The bullet leaves the rifle with a speed of 300 m/s. With what speed does the rifle recoil?
4. A rocket is coasting at a constant velocity v in outer space. When the rocket engine is turned on, the rocket increases its speed. Explain how this is possible when there is nothing for the rocket to push against. Is this consistent with conservation of momentum?
5. Explain what is meant by kinetic energy; rest-mass energy; radiant energy. In what unit is energy measured?
6. By what factor does an object's kinetic energy increase if its speed is tripled?
7. How fast must a proton move to have the same kinetic energy as an electron traveling at 2×10^6 m/s? (The mass of a proton is 1.67×10^{-27} kg and the mass of an electron is 9.11×10^{-31} kg.) How fast must the proton move to have the same momentum as the electron?

8. Calculate the kinetic energies both before and after the collisions in examples 2.1 and 2.2. What has happened to the kinetic energy that appears to have been lost?
9. Explain the role of the factor c^2 in the equation $E = m c^2$.
10. The rest-mass energy transformed into kinetic energy in the chemical explosion of one ton of TNT is 4.2×10^9 J. This amount of energy is now commonly called a "ton." The kinetic energy produced by nuclear bombs is measured in kilotons or megatons (the prefix kilo means "thousand," and the prefix mega means "mi1lion"). What is the rest-mass energy of l kg of mass measured in tons?
11. Can an object have energy without having momentum? Can an object have momentum without having energy? Explain your answers.
12. In a typical chemical reaction, about 10^5 J of rest-mass energy is converted to kinetic energy for each l kg of mass of the substances involved in the reaction. What percentage of the rest-mass energy is lost during the reaction?
13. When a nucleus of uranium-238 undergoes radioactive decay, it ejects an alpha particle and becomes a nucleus of thorium-234. The rest mass of the uranium nucleus is 3.9508×10^{-25} kg, the rest mass of the alpha particle is 0.0664×10^{-25} kg, and the rest mass of the thorium nucleus is 3.8843×10^{-25} kg. Assume that the uranium nucleus is at rest before the reaction. How much rest mass is lost in this reaction? If no radiant energy is involved, what must be the total kinetic energy after the reaction?
14. Apply the laws of conservation of momentum and energy to the previous question to determine the speeds of the alpha and thorium after the reaction.
15. A car of mass 1500 kg collides head-on with a truck of mass 5000 kg. The truck is initially at rest before the collision and the initial velocity of the car is 20 m/s east. The car and truck stick together after the collision. What is the velocity of the wreckage immediately after the collision?
16. In question 15, what is the kinetic energy before and after thy collision? Where did the energy go?
17. In Question 15, assume that the truck has an initial velocity of l0 m/s west and that all else remains the same. What is the velocity of the wreckage immediately after the collision?
18. Explain why the fact that energy must be added to break something apart implies that the rest mass after the interaction must be greater than the rest mas of the original object. Two billiard balls are glued together with a powerful glue. The total rest mass is 0.2 kg. If 3 J of energy is required to break the balls apart, what is the change in rest mass of the system?
19. A certain atomic nucleus is initially at rest, with a rest mass of 3.3626×10^{-27} kg. It breaks into two parts when 3.52×10^{-13} J of radiant energy is added to the

system. The two parts are approximately at rest after the interaction. That is, the final kinetic energy is negligible. What is the total rest mass of the two parts?

The following questions are of a more general nature. They have no single correct answer and are just something for you to think about. When possible, they are best answered in conversation with others.

20. "The events, processes, and interactions by which reality manifests itself and from which reality receives meaning, consist of the transfer and/or transformation of energy." What do you think this statement means? Do you agree?
21. Discuss your intuitive feelings about the distinction between matter and energy.
22. An atom initially at rest emits electromagnetic radiation and recoils as a result. Discuss this interaction qualitatively in terms of radiant energy, kinetic energy, and rest-mass energy, and in terms of conservation of charge, momentum, and energy.

3 MODERN PHYSICS: A BRIEF HISTORY

With Newtonian physics, electromagnetism, the kinetic theory of heat, and thermodynamics (what we now refer to as classical physics) well established, it appeared to physicists at the start of the 20th century that it was only a matter of time until all phenomena in the physical universe would be described in complete detail. True, certain perplexing experimental results had come to light in the last decade of the 19th century, but physicists had no doubt classical physics could eventually explain these phenomena. A leading physicist was advising promising students not to go into physics as he felt all of the fundamental work had been done and all that remained for 20th century physicists was the rather mundane task of applying the theory and determining additional decimal places for the physical constants.

3.1 OBSERVATIONS THAT APPEARED TO DEFY ANALYSIS USING CLASSICAL PHYSICS

Thermal Radiation

It is a well-known fact that hot objects give off electromagnetic radiation. It is most obvious when the object is hot enough to emit visible light, as is the case with the sun or a light bulb filament. The fact is that all objects emit electromagnetic radiation, but cooler objects emit only in the radio and infrared regions. The radiation emitted by an object because of its temperature is called thermal or black-body radiation,

Observed Distribution of Thermal Radiation

As the temperature increases, the radiated energy increases at all wavelengths (height of curve at a particular wavelength), the total energy radiated increases (area under the curve), and more energy is emitted at shorter wavelengths (peak of the distribution shifts left).

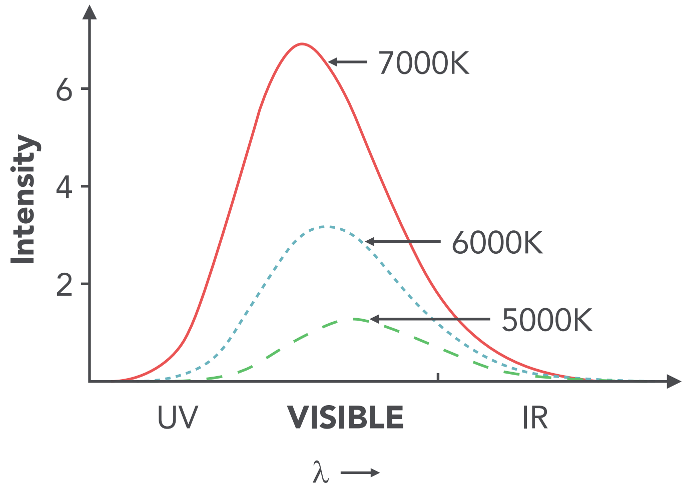


Figure 3.1 Thermal Radiation

In the late 1800s, Lord Rayleigh rigorously applied the principles of classical physics to the problem of thermal radiation. The results were blatantly wrong. They indicated any object, regardless of its temperature, will radiate more electromagnetic energy in the ultraviolet region (the shortest wavelength electromagnetic radiation known at that time) than it will in the visible or infrared. This classical treatment is aptly known as the 'ultraviolet catastrophe'. In fact, the actual distribution of thermal radiation had been accurately measured by that time and low temperature objects emit virtually no energy in the ultraviolet region.

Ultraviolet Catastrophe

Classical theory predicts a continuous increase in energy emitted as the wavelength decreases toward the ultraviolet (shorter wavelengths) region of the electromagnetic spectrum. A wavelength of 500 nanometers (nm) lies in the visible region of the electromagnetic spectrum.

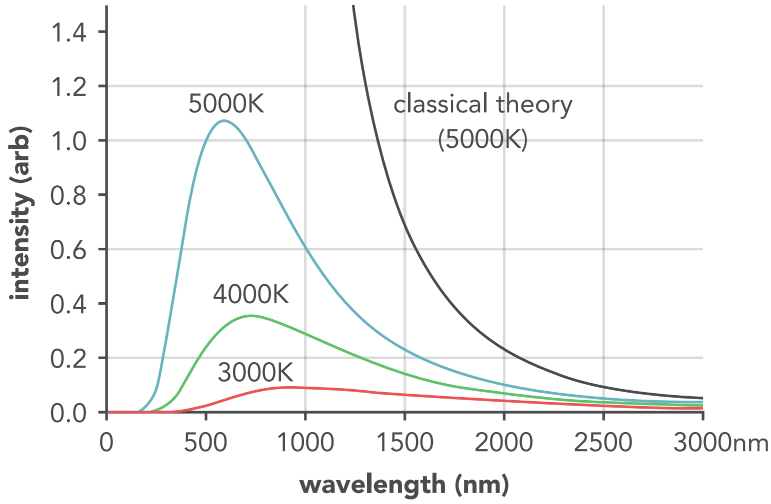


Figure 3.2 Ultraviolet Catastrophe

The Ether

An integral part of classical electromagnetic theory was the ether. The ether was supposed to be the invisible, weightless, substance which permeated the entire universe and served as the medium through which electromagnetic waves were propagated in much the same way that sound is propagated through the air. A wave by its very nature is an oscillating disturbance in some medium and ether was postulated as the medium that oscillated to propagate electromagnetic waves. Repeated efforts to detect the ether failed.

Atomic Spectra

In the last few decades of the 19th century, much experimental work was done on the electromagnetic radiation emitted by a low-density atomic gas. The distribution of the wavelengths emitted in this case differed considerably from that emitted in the case of thermal radiation. The radiation from an atomic gas did not form the continuous spectrum of thermal radiation, but rather a discrete spectrum. Only a limited number of wavelengths are emitted. Further, the discrete spectrum is characteristic of the emitting gas. An element could be identified simply by analyzing its emission spectrum. Classical physics could provide no explanation to account for this.

Atomic Emission Spectra

As opposed to continuous thermal emission spectra, emission from individual atoms is discrete. Only certain wavelengths are emitted and the wavelengths and intensities are characteristic of the chemical element. The figure show emission spectra for hydrogen, gold (AU), aluminum (Al), and silver (Ag).

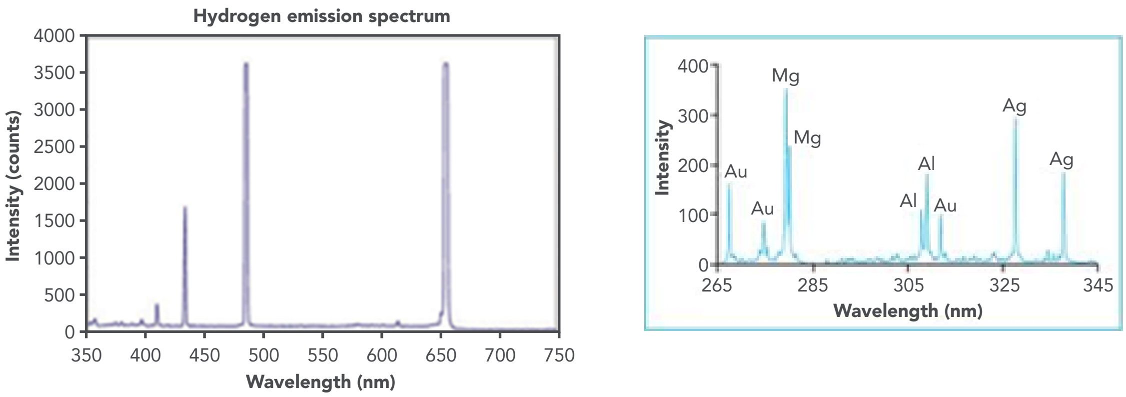


Figure 3.3 Atomic Spectra

Orbit of Mercury

Newton's law of gravity and his laws of motion could be used to calculate the orbits of the planets to an incredibly high degree of accuracy. The observed orbits of all the planets except Mercury correspond to the predictions based on classical Newtonian mechanics. For Mercury, the discrepancy was very, very slight, but undeniable. Newton's laws required that the axes of the elliptical orbit precess; that is, swing through a certain angle in a certain amount of time. However, the amount of precession calculated was slightly different from that observed. Some astronomers attributed the discrepancy to the perturbing effects of an undiscovered planet located between Mercury and the sun. They even went so far as to name the planet Vulcan in anticipation of its discovery. This same strategy had been used successfully to account for discrepancies in the orbit of Uranus, resulting in the discovery of Neptune. It soon became clear, however, that Vulcan did not exist.

Precession of Mercury's Orbit

Newtonian physics had successfully accounted for the orbits of the other planet but could not exactly match the observed orbit of Mercury. The discrepancy was very slight, but was undeniable.

Observed precession (in 100 yr)	= 574.2 ± 0.4″
Precession predicted by Newtonian theory	= 531.1 ± 0.2″
Observed discrepancy	= 43.1 ± 0.5″

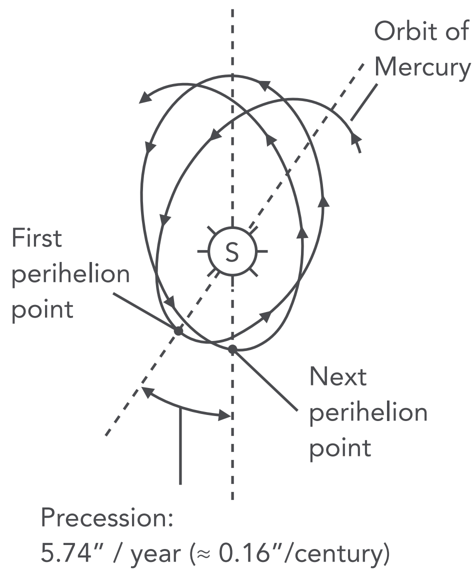


Figure 3.4 Mercury's Orbit

The Electron

If a high voltage is placed across the two ends of an evacuated glass tube, experiments demonstrate the existence of invisible rays emanating from the negative electrode or cathode. In 1869 these cathode rays were shown to travel in straight lines, and in 1870 they were shown to have both energy and momentum. In 1895 Jean Perrin discovered that the cathode rays carry negative charge by deflecting them in a magnetic field. J. J. Thomson (1856 – 1940) was able to show, in 1897, that cathode rays are steams of a single type of negatively charged particle whose properties are independent of the material from which they are emitted. Thomson's experiments indicated that the mass of these particles is much, much less than the mass of even the lightest atom. These particles were recognized as being responsible for electricity and became known as electrons.

By 1900 it was well established that all atoms contain electrons as part of their internal structure and the race was on to develop an atomic model into which electrons could be incorporated in a logical way that would be consistent with the laws of classical physics.

Cathode Rays

Cathode rays (electrons) are emitted by the negatively charged cathode and attracted toward the positively charged anode. Electrons missing the anode cross will fluoresce a screen at the end of the tube, leaving a shadow because the electrons could not pass through the cross.

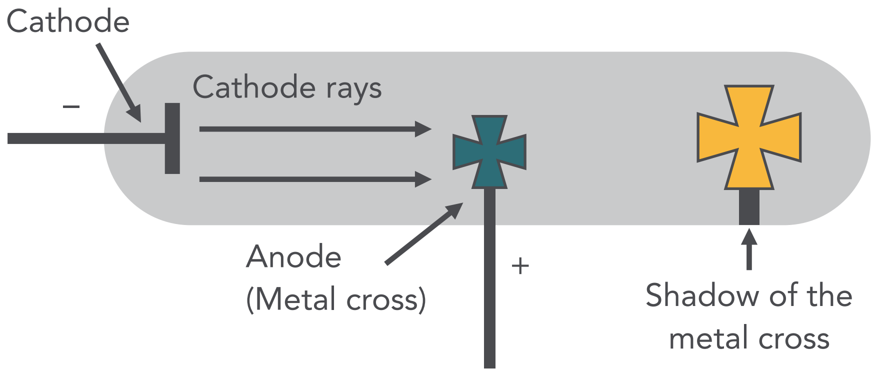


Figure 3.5 Cathode Rays

The Photoelectric Effect

In 1887 Heinrich Hertz (1857 – 1894) discovered that it was possible for ultraviolet, or in some cases visible light, to free electrons from certain metallic surfaces. The photoelectric effect itself is not too surprising. Ultraviolet and visible light possess energy and therefore should have the ability to free electrons from the metal by transferring energy to the electrons. The difficulty comes when the photoelectric effect is studied in detail. For each type of metal, there is a minimum wavelength of radiation that will free electrons regardless of how much energy is transferred to the metal. On the other hand, below this minimum wavelength, electrons will be released, even for very low amounts of energy transferred. This behavior is incompatible with the classical wave model of electromagnetic radiation.

Radioactivity

The phenomenon of radioactivity was discovered in 1896 by the French physicist Henri Becequerel (1852 – 1908). In January of that year, Becquerel learned of an amazing discovery made by the German physicist Wilhelm Roentgen (1845 – 1923). When cathodes rays strike glass, they cause the glass to emit visible light. This phenomenon was well known and is called fluorescence. What Roentgen discovered is that in addition to the visible light, the fluorescent areas of the glass also emit an extremely penetrating radiation. Because the nature of this unexpected radiation was unknown, Roentgen simply called them X-rays. (Later X-rays were shown to be electromagnetic radiation with wavelengths shorter than ultraviolet light.) News of this mysterious radiation spread rapidly, and physicists all over the world began to study the properties of X-rays. Because X-rays could be used to produce dramatic photographs of bones inside a living body, the popular press splashed the story over the front pages.

When Becquerel learned about Roentgen's discovery, he immediately set out to try to discover whether the X-rays are simply a peculiar feature of cathode-ray tubes or whether they are associated with fluorescence in general. Becquerel knew that certain minerals will fluoresce when irradiated with ultraviolet light, so he set out to discover if X-rays are also associated with this fluorescence.

Becquerel carefully wrapped a photographic plate with black paper to block visible and ultraviolet light and placed the fluorescent mineral on top of the wrapped photographic plate. After irradiating the mineral with ultraviolet light to cause fluorescence, he developed the photographic plate to see if X-rays had penetrated the black paper to expose the film. His early experiments produced no exposure of the film. Then he used some fluorescent uranium minerals, which did cause exposure of the film, indicating that a very penetrating radiation was emitted by the fluorescing uranium minerals. Naturally Becquerel assumed that this radiation was X-rays.

One day Becquerel happened to develop some photographic plates that had been left in a drawer with samples of the uranium minerals. The minerals had not been exposed to ultraviolet light and had therefore not fluoresced. The plates had been wrapped in black paper, so there was no reason to expect any exposure of the plates, but they were exposed. Subsequent experiments showed that the uranium minerals spontaneously emit the penetrating radiation, even if they are not fluorescing. Becquerel was able to show that it was specifically the uranium atoms in the minerals that were emitting the radiation. Any sample of uranium spontaneously emits this radiation without any external energy supply. This phenomenon is quite different from that observed by Roentgen where X-rays are emitted only when glass is bombarded by cathode rays. Becquerel's new phenomenon was called radioactivity. A substance that emits this spontaneous radiation is said to be radioactive. Classical physics was at a loss to explain the nature of this radiation.

Henri Becquerel
(1852 – 1908 * France)

1896 – discoverd that uranium spontaneously emitted radition, later identifies as alpha radiation.
Thesis advisor to Marie Curie.
1903 – Nobel Prize in Physics.

Figure 3.6 Henri Becquerel

The Scientific Worldview at the Start of the 20th Century

In spite of these perplexing experimental results, at the dawn of the 20th century, physicists believed that their theoretical understanding of the physical universe was complete. Newton had provided a description of gravity, the only force acting on a universal scale, as well as the laws governing the motion of objects acted upon by forces. Maxwell had provided a complete description of electromagnetic phenomena, including electromagnetic radiation. Heat was now understood as energy of motion and the laws of thermodynamics had been established. Atoms were understood as tiny, submicroscopic objects obeying the same laws of motion as ordinary-sized macroscopic objects. Space and time were understood as absolute and independent structures within which the phenomena of the physical universe unfolded.

This mechanistic-deterministic universe seemed so obvious, it must be true. There was little doubt that soon some clever physicist would figure out how to account for all of the strange experimental results mentioned above, and that the explanations would lie completely within the theoretical framework of classical physics. Such hubris seldom goes unpunished.

3.2 TWENTIETH-CENTURY PHYSICS

Most physicists in 1900 believed they were nearing their goal of explaining all physical data in terms of the Newtonian mechanics, thermodynamics, and electromagnetism. However, within the first decade of the new century, attempts to solve the remaining puzzles outlined above led to the proposal of radically new physical theories.

What follows is a brief outline of the major developments of twentieth-century physics. Do not get bogged down in the details of these developments; we will discuss the physical content of the new theories more thoroughly in the later chapters.

The Quantum Hypothesis

In 1900 the German physicist Max Planck (1858 – 1947) was able to derive a black-body radiation formula that closely agreed with the observational data. He derived his formula from theories of classical physics, closely following an argument that other physicists had used earlier to explain the distribution of kinetic energy among the molecules of a gas. However, other physicists soon pointed out that Planck's successful derivation involved a hidden assumption that was not justified by classical physical theories. Hidden in his argument was the assumption that a vibrating atom can vibrate only with certain allowed levels of kinetic energy. Radiant energy then is emitted as an atom jumps from one allowed level to the next lower one, so that energy can be emitted only in certain allowed amounts. By this time it was quite clear that a classical treatment cannot explain the observed radiation law, so physicists were forced reluctantly to consider the possible validity of Planck's disturbing assumption.

The new assumption is called the quantum hypothesis, in reference to the fact that the energy of the vibrating atoms (or charged particles within the atoms) can assume only certain allowed values. Any physical quantity whose possible values are restricted in this way is said to be quantized. The idea that the energy of a vibrating particle is quantized is completely foreign to classical theories of physics. It was not that quantization was unknown in classical physics – charge and the frequency of a vibrating guitar string were clearly quantized.

However, kinetic energy is proportional to the square of an object's speed, and classical mechanics does not provide any justification for the idea that speed should be quantized. In fact, it states clearly and without any ambiguity that a vibrating particle should be able to have any amount of energy whatsoever. The hypothesis that most values of energy are 'forbidden' seems just as absurd as saying that an object can move with a speed of l m/s or a speed of 3 m/s, but cannot move with any speed in between these values. How can the vibrating particle get from one allowed energy to another without having any of the forbidden energies in between? No physicist in 1900 took the quantum hypothesis seriously. In fact, even Planck himself resisted the idea and suggested that some more reasonable theoretical explanation would soon be found to account for this absurd appearance of quantized energies.

We now know that the quantum hypothesis marked the beginning of a revolution that was to shake the very foundations of our worldview. Planck's presentation of his radiation formula before a meeting the German Physical Society on December l4, 1900, is usually considered the birthday of modern physics. In 1918 Planck was awarded the Nobel Prize in physics for his work.

The Photon Hypothesis

In 1905 Albert Einstein (1879 – 1955) was able to explain the details of the photoelectric effect by extending Planck's quantum hypothesis. In Planck's theory only kinetic energy of the vibrating particles is quantized. No assumptions are made about the nature of the emitted radiant energy itself. Planck did not for a moment doubt that the classical wave description of electromagnetic radiation was correct. Einstein, however, had an unusual ability to consider nature without being unduly limited by existing theories. Recognizing the implication of Planck's hypothesis, Einstein suggested the possibility that radiant energy itself is quantized; that is, it can exist only in discrete packets of energy. Such a packet is called a quantum. A quantum of radiant energy later came to be called a photon.

Einstein pointed out that this quantized model of electromagnetic energy is equivalent to describing light as a stream of particles (photons) rather than a pattern of waves. This is a direct contradiction to all the experimental evidence obtained the nineteenth century for the wave nature of light. Nonetheless, Einstein showed that his quantum hypothesis completely accounts for all the results of the experiments on the photoelectric effect.

Like Planck, Einstein was dubious about the actual significance of his idea. He felt that some later explanation would account for the apparent quantization of radiant energy without requiring the discarding of the very successful model of electromagnetic waves. Einstein was awarded the 1921 Nobel prize in physics for his theoretical work in general, with particular emphasis on his explanation of the photoelectric effect.

The Special Theory of Relativity

In that astonishing year of 1905, Einstein published another paper that was to have far - reaching effects on physical theories. With his rare ability to take a fresh look at reality, Einstein was not particularly disturbed by the failure to detect the ether. He simply used as one of the basic postulates of his theory of relativity the fact that the measured speed of light is the same for all observers, regardless of their relative motions with respect to the source of the light. As Einstein pointed out, this assumption means that the ether can never be detected, and therefore that the concept of the ether is superfluous and can simply be discarded. This is an extremely radical notion. Not only is it contrary to the principles of classical physics, but worse yet for most of us, it seems completely contrary to commonsense thinking. How can an observer measure the speed of light as a constant value, no matter whether the observer is moving toward the light or away from it? For example, suppose that the observer is moving away from the light source at a speed nearly as great as the speed of light. How can it be possible that the observer will still measure the light moving past at the speed of 3×10^8 m/s, exactly the same speed as determined by an observer at rest with respect to the source?

Even more upsetting were the conclusions that Einstein showed to follow from his basic postulates.

1. Space is not absolute, as Newton postulated. The measured distance between two objects in space cannot be absolutely defined. Instead, the measured distance depends on the relative motion of the observer with respect to the two objects.
2. Time is not absolute either. The measured time interval between two events cannot be absolutely defined, but depends on the relative motion of the observer with respect to the two events.
3. The measured mass of an object is not an absolute property of the object but rather depends on the relative velocity of the observer with respect to the object.
4. Mass and energy are equivalent. That is, mass and energy are merely two different ways of measuring the same physical quantity.
5. If an object has a nonzero rest mass, then it can never move at a speed equal to or greater than the speed of light. If a particle has zero rest mass, then it can move only at the speed of light.

These results seem counter to commonsense, but commonsense notions are based on personal experiences. Under normal circumstances, the predictions of Einstein's theory are almost identical to those of the classical physical theories. The differences between Einstein's predictions and the classical ones become significant only when we deal with speeds comparable to the speed of light. We have no experiences with this range of phenomena, so it is not too surprising that our commonsense expectations prove unsuitable when extended into this range. The predictions of Einstein's special theory of relativity have been repeatedly tested, and every test has confirmed the validity of the theory.

The General Theory of Relativity

In 1915 Einstein proposed his general theory of relativity, which is essentially a theory of gravitation. In classical theories, gravity is regarded as a force exerted at a distance by one mass on another. In Einstein's theory, gravity becomes a property of space-time. The presence of a large mass (a large concentration of energy) is associated with a curvature of the space-time in its vicinity. The motion of objects in space-time is altered by this curvature. No gravitational force need be postulated; each object simply moves along the shortest possible path through the curved space-time

This theory provided an exact explanation of the orbit of Mercury. Newtonian predictions are inexact in the curved space-time near the very large mass of the sun. Einstein predicted some other observations that could be used to test his theory, including the prediction that the path of light would be curved when passing very near the sun and that time would slow down in the vicinity of a very large mass. Both of these predictions and others have subsequently been confirmed experimentally. The general theory of relativity also provides the basis for our current understanding of the structure of the universe and of such exotic objects as black holes in space. The special theory of relativity is completely consistent with the general theory, and the combined theories are commonly known as the theory of relativity

The Quantized Atom

In 1913 the Danish physicist Niels Bohr (1885 – 1962) was able to account for the discrete emission spectrum of hydrogen gas by assuming that electrons in the hydrogen atom exist only in certain allowed orbits such that the kinetic energy of the electron is quantized. Like Planck before him, Bohr based this assumption upon the need to account for the known physical data, and he admitted that he could not explain the quantized orbits in terms of existing physical theories. Although it proved very difficult to develop similar models for more complex atoms or for molecules, the great success of Bohr's rather simple model in explaining the hydrogen spectrum led many chemists and physicists to adopt a quantized model of atomic structure. This approach did provide qualitative explanations for many features of other atoms and molecules, even though detailed quantitative models for atoms other than hydrogen could not be developed at the time. Bohr was awarded the l922 Nobel prize in physics for his work.

The Wave Nature of Matter

As physicists were forced to admit the validity of the puzzling notion that electromagnetic radiation has both wave and particle aspects to its nature, they were soon forced to confront an even more surprising proposition. In 1924 a young French physicist of noble birth set forth in his doctoral thesis the hypothesis that such particles as electrons and protons should also have wave aspects to their nature. Prince Louis de Broglie (1892 – 1987) proposed an equation predicting the wave nature of these particles. Experimental verification of this prediction was obtained in 1927 by the American physicist Clinton Davisson who quite unexpectedly noticed wavelike patterns in the behavior of electrons reflected from a crystal surface. De Broglie was awarded the 1929 Nobel Prize in physics for the theoretical work he had set forth in his thesis.

Quantum Mechanics

In 1926 the Austrian physicist Erwin Schrodinger (1887 – 1961) provided a theoretical foundation for the quantum hypotheses of Planck and Bohr. In his theory of wave mechanics, Schrodinger showed that these quantum properties follow logically from the assumption that electrons behave under some conditions as "matter waves" of the type predicted by de Broglie.

At the same time, another German physicist Werner Heisenberg (1901 – 1976) proposed a theory called matrix mechanics, based upon an advanced mathematical technique called matrix algebra. Very shortly thereafter, Schrodinger was able to show that the two theories were equivalent, differing only in their mathematical form and in the particular choice of postulates. Both theories are now referred to as quantum mechanics. It is no longer necessary to make arbitrary assumptions about quantized energy; the quantized energy states follow logically from the much more basic postulates of quantum mechanics.

The theory of quantum mechanics also leads to the important result that physical laws on the atomic scale must be expressed in terms of probabilities. That is, the quantum mechanical laws say that, given a certain set of initial conditions, various possible outcomes can be assigned various probabilities of occurrence. In contrast, classical mechanical laws say that, given a certain set of initial conditions, a particular event either will occur or else cannot possibly occur. Thus the laws of classical mechanics are deterministic; a certain cause must always lead to a certain effect. According to the theory of quantum mechanics, physical laws on the atomic level become nondeterministic; we cannot predict a specific result but can only predict the probability of any possible result.

Quantum mechanics proved extremely successful in explaining and predicting the results of experiments involving events on the atomic level. None could deny its usefulness as a tool in physics. However, a heated controversy soon developed over the philosophical implications of the new theory. Bohr and Heisenberg were among a group who regarded quantum mechanics as a fully satisfying description of reality at the atomic level.

Einstein, de Broglie, and Schrodinger were among many prominent physicists who felt that quantum mechanics is an incomplete theory and there must be some underlying, fully deterministic description of the phenomena of atomic physics. They felt that the probabilistic nature of quantum mechanical laws must simply reflect our failure to understand fully interactions at the atomic level, and a more complete theory with completely deterministic laws would eventually be created to replace the theory of quantum mechanics. This belief prompted Einstein's famous remark, "God does not play dice with the universe." (Einstein often referred to God when discussing physics, but this is a little misleading. Einstein did not believe in a personal God. "I believe in Spinoza's god, who revealed himself in the harmony of all being, not in the God who concerns himself with the fate and actions of men." For him, God meant an underlying beauty and order in the universe.)

Most physicists today have come to accept the viewpoint of Bohr and Heisenberg. The younger generations of physicists, those who learned quantum mechanics as a basic part of their original training in physics, accept it as a natural and quite satisfying representation of reality.

3.3 THE CORRESPONDENCE PRINCIPLE

A physical theory is created to summarize the relationships that exist among a number of physical laws, which in turn summarize the relationships among great numbers of physical data. The validity of the theory can be tested by using it to predict the outcome of experiments falling within the same range of general phenomena as those used to create the theory. If the theory fails such tests, it must either be modified or discarded. The theory can also be used to predict the outcome of experiments that lie outside the known range of its applicability. For example, the theories of classical mechanics and electromagnetism were used to predict the structure of the atom. If experimental results had confirmed these predictions, then the range of the theory would have been extended. However, in this case, the experimental results did not match the predictions. Therefore, it became necessary to seek a new theory that would more adequately describe the physical laws in the new regions of investigation.

It is not a very satisfactory situation to have two incompatible theories that each applies only to one region of phenomena. The physicist's goal is always to find a completely general theory that seems to "explain" all the available physical data. One reason the new theories of relativity and quantum mechanics were accepted as quickly as they were is was the fact that they are compatible with the classical theories. At speeds much less than the speed of light, the predictions of the theory of relativity are quantitatively identical to the predictions of classical mechanics. At the scale of objects visible to the naked eye, the predictions of quantum mechanics are quantitatively identical to the predictions of the classical theories. Looking back at this situation, Niels Bohr proposed a principle that might be useful in judging the acceptability of new theories. It is called the correspondence principle:

> **A new and more general physical theory must yield the same quantitative predictions as the older and more restricted theory when applied in the range where the older theory is known to give accurate predictions.**

The correspondence principle will be an important criterion used in judging new physical theories that may emerge in coming decades. Even though the theories of relativity and quantum mechanics may eventually be supplanted by new and more general theories, we may confidently expect the laws included in the current theories will also emerge as predictions of the new theories within the appropriate range of conditions. Just as the theories of classical mechanics, thermodynamics, and electromagnetism remain useful today for predicting the outcomes of experiments within the ranges of their applicability, so we can expect that the theories of the twentieth century will remain in use in the future even if they are eventually shown to be just special cases of more general theories.

Summary

Classical physics includes the theory of classical mechanics formulated by Isaac Newton in the last part of the seventeenth century, the theory of thermodynamics formulated by a number of physicists and chemists in the middle of the nineteenth century, and the theory of electromagnetism formulated by James Clerk Maxwell at about the same time. The theory of classical mechanics explains the behavior of objects and particles under the influence of forces. The theory of thermodynamics explains the interactions of heat with the phenomena of mechanics, utilizing the important unifying concept of energy. Maxwell's theory of electromagnetism explains the forces that charged and magnetized particles exert on each other and the behavior of electromagnetic waves.

In the classical world view, the universe consists of particles whose positions in space and motion through space and time are determined by the forces they exert on each other. Thus, if the forces and the positions of the particles are known, then their motions at all future times can (at least in principle) be calculated. This is called the mechanistic-deterministic worldview.

At the end of the nineteenth century, physicists had become aware of a number of phenomena in nature that could not be explained in terms of the classical physical theories. Although few physicists at the time perceived a crisis in our understanding of the physical universe, the new physical theories of relativity and quantum mechanics proposed to solve these puzzles soon led to revolutionary changes in the basic worldview of physicists. The new theories account for a much wider range of phenomena than do the theories of classical physics. These advances in our understanding, however, have not been without a price. The mechanistic/deterministic worldview, along with much of our commonsense intuitive understanding of the universe were destroyed.

When new, more general physical theory are proposed, the correspondence principle requires that the new theory yield the same quantitative predictions as does the older theory when applied to the phenomena for which the older theory is known to be adequate. Hence, although the theories of quantum mechanics and relativity may eventually be explained in terms of more general theories, they will remain useful tools for the study of the phenomena to which they apply.

Important concepts

Mechanistic-deterministic world view; classical mechanics; thermodynamics; electromagnetism; thermal radiation; photoelectric effect; the ether; quantum hypothesis; quantized quantity; photon; theory of relativity; theory of quantum mechanics; correspondence principle.

Questions

1. What is the mechanistic-deterministic world view?
2. Rayleigh's formula for thermal radiation is often said to have failed because it predicted an "ultraviolet catastrophe." What is meant by this term?
3. What does the word "quantized" mean when applied to a physical quantity such as energy? Are there any physical quantities that are considered to be quantized in classical physical theories? Name some.
4. What is a photon?
5. The theories of modem physics are often said to involve a "wave-particle duality." What does this phrase mean?
6. Observer B is moving from left to right at a speed v with respect to observer A. Observer C is moving from right to left at a speed v with respect to observer A. All three measure the speed of a light beam that is traveling from left to right. Observer A measures the speed of the light beam to be c. According to the theories of classical physics (and commonsense reasoning), what would be the predicted result of B's measurement? of C's measurement?

7. According to the special theory of relativity, what would be the predicted results of B's and C's measurements in the previous question? If the experiment is performed, what will be the result?
8. Discuss the relationship between the theory of relativity and the theories of classical physics in terms of the correspondence principle.

The following questions are of a more general nature. They have no single correct answer and are just something for you to think about. When possible, they are best answered in conversation with others.

9. What is meant when it is said that classical mechanics is a deterministic theory, whereas quantum mechanics is a nondeterministic theory?
10. In your own words, explain the difference of opinion between Einstein and Bohr about the interpretation of the theory of quantum mechanics.
11. How did most physicists in 1900 view the state of physics? In your opinion, how is the attitude of physicists today similar to the attitude of the physicists in 1900? How is it dissimilar?
12. Discuss the values and limitations of commonsense reasoning.
13. Jacques Merleau-Ponty said, "The success of classical physics rests on the fact that there is often a certain numerical continuity between its results and those of modem physics, but this numerical continuity conceals a logical discontinuity that is abruptness itself." Discuss this statement in light of what you have read in this chapter.
14. John Dewey said, "Physical science makes claim to disclose not the inner nature of things, but only those connections of things with one another that determine outcomes." Discuss this statement. Do you think that statements concerning the "inner nature of things" may be hidden in the underlying assumptions and postulates of a physical theory,
15. Newton's theory of classical mechanics led to predictions that were in many cases quantitatively different from the predictions of the Aristotelian theory of physics. Would the correspondence principle imply that the Newtonian theory should have been rejected?
16. Is it possible for a physical theory to be useful even though the physicist using it does not believe that it represents the best possible explanation of physical reality? If your answer is yes, give a couple of examples.

4 PHYSICAL LAWS IN DIFFERENT FRAMES OF REFERENCE

What does it mean to say an object is moving or is at rest? Consider a person sitting quietly in a seat in a railroad car. The person is at rest with respect to the train. But if the train is moving, the person is also moving with respect to the earth. Even if the train is at rest in a station, the earth itself is moving in its orbit around the sun, so the person also is moving with respect to the sun. Clearly, we can describe the motion of an object only in terms of some particular reference object, or frame of reference.

Relativity of Speed

In A's frame of reference, the train is moving left to right at 60 mph. In B's frame of reference, the train is moving left of right at a speed of 20 mph. In C's frame of reference, the train is moving left to right at a speed on 100 mph. Speed, in general, is a relative physical quantity.

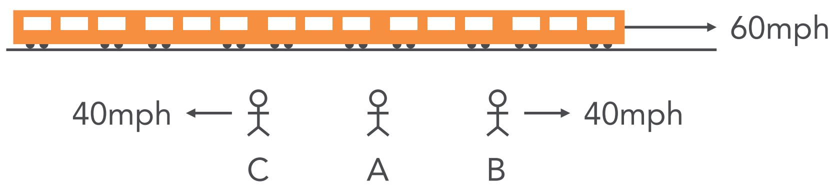


Figure 4.1 The relativity of speed

4.1 FRAMES OF REFERENCE

Imagine the situation illustrated in Figure 4.2. A boy with a baseball in his hand is on a train traveling at a speed of 30 m/s from left to right with respect to the ground. A girl is standing on the ground and also has a baseball in her hand. Using the train as his frame of reference, the boy says his baseball is at rest, whereas the girl's baseball is moving right to left with a speed of 30 m/s. Using the ground as her frame of reference, the girl says her baseball is at rest, whereas the boy's baseball is moving left to right at a speed of 30 m/s. Who is right? Which baseball is actually moving and which is actually at rest?

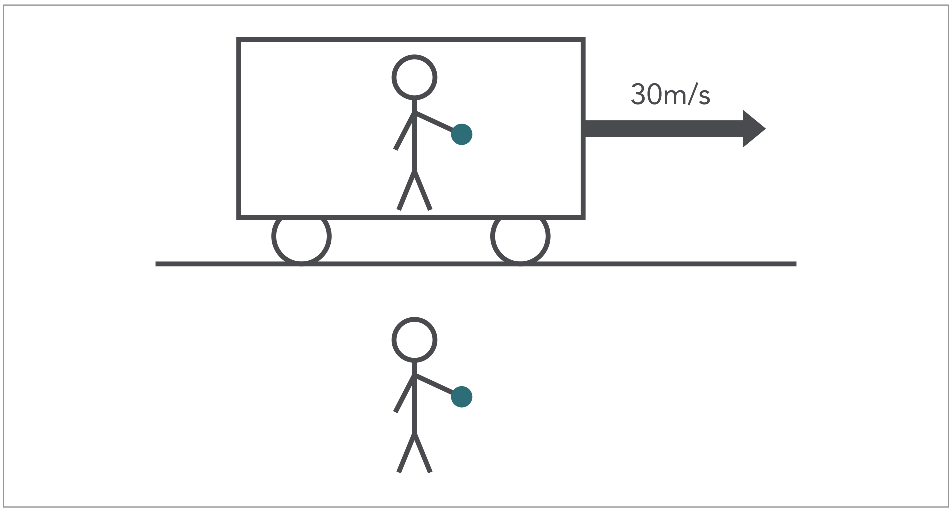


Figure 4.2

You may be inclined to say the girl is correct, whereas the boy only thinks his baseball is at rest because he does not take into account the movement of the train. But this is only because we are in the habit of using the earth as our frame of reference. We tend to feel an object moving with respect to the earth is actually moving, and an object at rest with respect to the earth is actually at rest. However, an observer hovering in a spaceship near the solar system notes the earth is moving in its orbit around the sun at a speed of nearly 30,000 m/s. This observer concludes both the boy and girl are wrong and both baseballs are moving with very large and almost identical speeds. It is meaningless to talk about whether an object is actually in motion or actually at rest. The terms "at rest" and "in motion" have no physical meaning when used alone. Any motion must be described in relation to some frame of reference. We can choose the train or the earth or the sun as the reference object for our frame of reference, and there is no reason to say any one of these is more valid than the others (although one particular reference frame may be more useful in finding a simple solution to a particular problem).

As a further example, suppose the boy and the girl are to measure the velocity of a thrown baseball (see Figure 4.3). The girl on the ground measures the velocity of the baseball as 40 m/s right to left. What velocity will the boy on the train measure, using the train as his frame of reference?

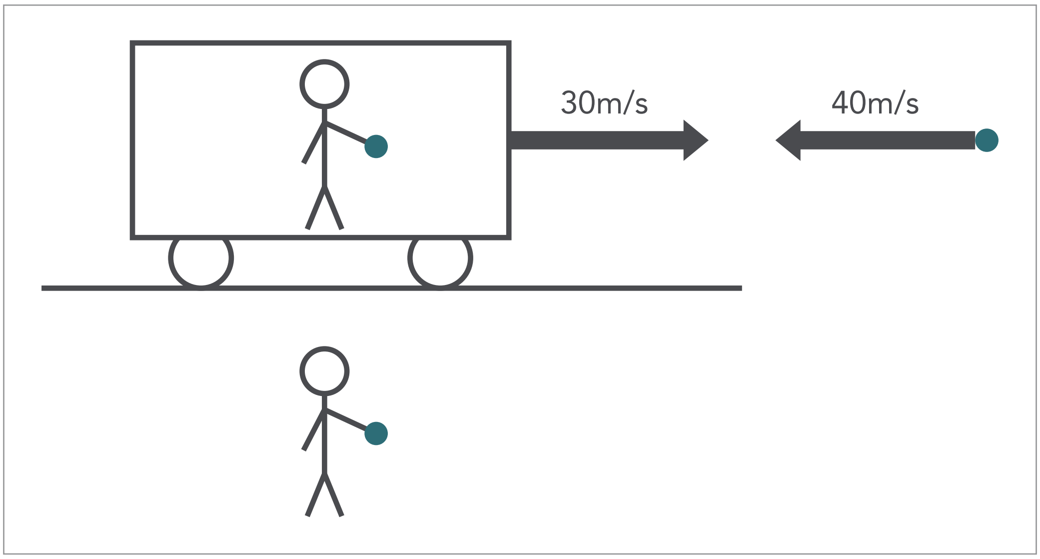


Figure 4.3

It should be clear, the boy, who is moving toward the oncoming ball, will see it moving past him with a speed much greater than the speed measured by the girl on the ground. During one second with respect to the earth, the ball travels 40 m toward the boy and the boy travels 30 m toward the ball. If the boy regards himself as standing still, then he sees the ball travel 70 m toward him during that second. Therefore, the boy measures the speed of the ball as 70 m/s. The boy concludes the ball has a velocity of 70 m/s right to left. (Be sure this conclusion makes sense to you. Later we will discuss a similar example of fundamental importance in the understanding of modern physics.)

What is the velocity of the ball? Again, both observers are correct. The velocity of the ball is 40 m/s right to left with respect to the ground, and it is 70 m/s right to left with respect to the train. There is no reasonable criterion that can be used to give one of these observers precedence over the other in quantitatively describing the velocity of any object. In general, any quantitative description of physical quantities must be expressed in relation to the frame of reference of the observer. We have just illustrated this principle with the physical quantity velocity. The same conclusion can be drawn about measurements of the quantity kinetic energy. A physical quantity whose quantitative description is different in different frames of reference is called a relative quantity. Thus, velocity and kinetic energy are relative physical quantities.

There are other physical quantities that seem to have the same quantitative description no matter what frame of reference is used by the observer. In the case we have just discussed, both commonsense and classical theories of physics tell us the boy and the girl would obtain identical measurements of the dimensions and the mass of the ball. We are also confident that, if the boy on the train bounces the ball, both observers will obtain the same value for the time it takes the ball to reach the floor and return to his hand. This is what Newton meant by 'absolute time.' Physical quantities that do not depend on the frame of reference of the observer are said to be absolute. Thus, according to the theories of classical physics, quantities such as length, mass, charge, and time are absolute physical quantities.

A physical law is a quantitative relationship between physical quantities. If the quantitative description of some physical quantities depends on the frame of reference of the observer, then does the form of the physical laws also depend on the frame of reference? The answer to this question is yes. In general, the form of a physical law does depend on the frame of reference in which the relevant physical quantities are measured. However, there does exist a class of frames of reference in which the laws of physics are the same and in which they have their simplest physical and mathematical descriptions. The laws of physics are always stated in the form that is valid in such frames of reference.

Returning to the train example, if the boy on the train bounces the ball on the floor, the motion he observes is identical to the motion the girl observes if she bounces her ball on the ground. The physical laws each observer would deduce to summarize the motion of the ball are identical in these two frames of reference. However, there are some frames of reference in which the motion of the ball is different, and therefore different physical laws are needed to describe it.

Example 4.1

Observer A is at rest with respect to the earth. Observer B is in a train moving at constant speed with respect to the earth. Observer C is in a train slowing down with respect to the earth. Each observer drops a ball and observes its motion as it falls to the floor. (See Figure 4.4.) How does each observer describe the motion of the ball with respect to the observer's frame of reference?

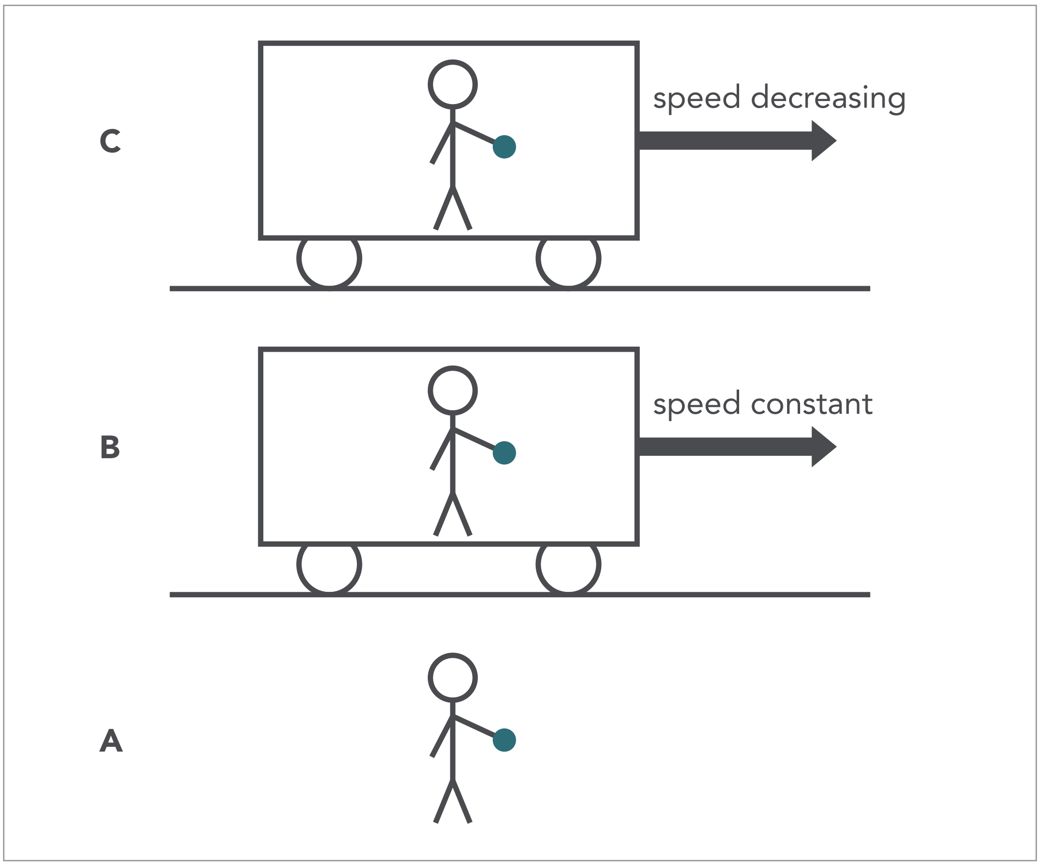


Figure 4.4

Solution

Observer A sees the ball move straight downward with ever increasing speed until it strikes the floor directly below the point at which it was released. Observer B also sees the ball accelerate downward until it strikes the floor directly under the point at which it was released. Thus observer B and observer A will formulate identical physical laws to describe the motion of a dropped ball in their respective frames of reference. Both summarize their observations by a physical law stating dropped objects are accelerated straight downward.

Now consider the experiment of observer C. When observer C releases the ball, both the frame of reference (the train slowing down) and the ball are moving sideways at some speed v relative to the ground; let's assume for the sake of argument this speed is v = 30 m/s. Once the ball leaves the observer's hand, the motion of the train no longer affects its motion – it is free falling, effected only by gravity. Therefore, it continues to move sideways at 30 m/s relative to the ground. Meanwhile, however, the train is slowing down, so the frame of reference of observer C is slowing down with respect to the ground. Thus, observer C sees the ball strike the floor at a point somewhat farther toward the front of the train than the point on the floor just below where it was released. Observer C concludes the ball has been pulled (or pushed) toward the front of the train as it falls down to the floor. Observer C must use a physical law more complicated than the law used by observers A and B. Observer C's law must include both a force pulling the ball down and some other force pulling the ball toward the front of the train.

4.2 INERTIAL FRAMES OF REFERENCE

An inertial frame of reference is one in which the laws of classical mechanics have their simplest physical and mathematical form. Once any inertial frame of reference is found, then any other frame of reference moving in a straight line at a constant speed with respect to the inertial frame of reference is also an inertial frame of reference.

The laws of classical mechanics have exactly the same form in all inertial frames of reference. Any frame of reference moving in a curved path or speeding up or slowing down with respect to an inertial frame of reference is not an inertial frame of reference The laws of classical mechanics are more complicated physically and mathematically in such a non-inertial frames of reference. That is why observer C in the previous example found a more complicated law to describe free fall in his frame of reference.

In our examples, we treated the earth as an inertial frame of reference. (More on this later.) A frame of reference moving at constant speed in a straight line with respect to the earth is also, by definition, an inertial frame of reference. All of the physical laws of classical mechanics are identical in these two frames of reference. However, a frame of reference speeding up or slowing down with respect to the earth (the frame of reference of observer C in the previous example, for instance) is, by definition, a non-inertial frame of reference. The physical laws of classical mechanics in a non-inertial frame of reference are different from (and more complex than) those in an inertial frame of reference.

In fact, a frame of reference attached to the earth is not quite an inertial frame of reference. The earth is rotating on its axis and moving in a curved path about the sun. However, these motions produce only very small complications in the exact expressions of the physical laws. For all practical purposes, a frame of reference attached to the earth is an inertial frame of reference. A better definition of an inertial frame of reference is one at rest or moving in a straight line with constant speed with respect to the very distant stars,

4.3 THE PRINCIPLE OF RELATIVITY

The laws of classical mechanics are the same in every inertial frame of reference. Therefore, we may suspect all physical laws (and specifically those of Maxwell's theory of electromagnetism) may be identical in every inertial frame of reference. This postulate is called the principle of relativity:

The laws of physics are the same (that is, they have the same mathematical forms) in every inertial frame of reference.

In example 2.1 in chapter 2, we illustrated conservation of momentum in a frame of reference in which one of the two balls before the collision was at rest. Imagine that example in a frame of reference moving with a velocity of 2 m/s left to right in a straight line with respect to the original frame. That would change all of the speeds in the example by 2 m/s and also change all of the momenta in the problem. The total momentum of the system would be different. However, in both cases, the momentum before and after the collision will be the same. Even though the two observers disagree on the value of the total momentum, they both agree momentum is conserved. Therefore the law of conservation of momentum satisfies the principle of relativity.

It is possible to show each of the laws of classical mechanics is consistent with the principle of relativity. But what about the laws of Maxwell's theory of electromagnetism? Maxwell's theory leads inescapably to the prediction that the speed of light in a vacuum is a constant – that is, it has a particular value given in terms of certain other constants determined in electrical and magnetic measurements. The speed of light in a vacuum is equal to a constant written as c, which is approximately equal to 3×10^8 m/s or 186,000 miles per second. Is this law consistent with the principle of relativity?

We've already shown speed is a relativity physical quantity – that is its value is different in different frames of reference. Because the law assigns a specific value to the speed of light, it would seem this law (and in fact all the laws of Maxwell's theory) cannot satisfy the principle of relativity. Either we should modify Maxwell's theory of electromagnetism in such a way it does not predict a unique value for the speed of light in a vacuum or we must conclude the principle of relativity is not valid when applied to Maxwell's theory. By the late nineteenth century, most physicists had concluded the principle of relativity must be abandoned as a general principle of physics. Apparently all inertial frames of reference are equivalent as far as the laws of mechanics are concerned but Maxwell's laws are valid in their simplest forms in only one special inertial frame and have different mathematical forms in all other inertial frames of reference.

Physicists were not surprised to conclude the speed of light must be different in different inertial frames of reference. After all, electromagnetic theory describes light as a wave and physicists were accustomed to the idea a wave has a certain speed with respect to the medium in which it travels. Because speed is a relative physical quantity, the speed will then have other values in other inertial frames of reference, frames moving with respect to the medium. Physicists assumed light travels through a medium called the ether in much the same way sound travels through air. Thus they assumed Maxwell's laws of electromagnetism are valid only in the special frame of reference at rest with respect to the ether. The speed of light would thus be a constant 3×10^8 m/s in any direction in the ether, and an observer at rest with respect to the ether would obtain this value in any measurement. However, any other inertial observer (not at rest with respect to the ether) would obtain a different value for the measurement of the speed of light, depending on the observer's relative motion with respect to the ether.

There is a flaw in our discussion of the principle of relativity and the law of propagation of light. We deduced the two are inconsistent on the basis of a thought experiment. If we wish to demonstrate with certainty that the principle of relativity and the law of propagation of light are inconsistent, then we must perform an actual experiment. Of course, common sense tells us what the result of that experiment will be, but we must be careful. Aristotle's common sense told him heavier objects fall to the ground more quickly than light objects do. Because this result was "obvious," it was accepted by Aristotle and by most who studied his teachings for nearly 2000 years. Galileo, however, easily proved Aristotle wrong by performing simple experiments.

An experiment must be designed to measure the speed of light with respect to two inertial frames of reference moving with respect to one another. It cannot be a simple experiment because the two inertial frames of reference must move at a measurable fraction of the speed of light with respect to one another. Albert Michelson (1852 – 1931) designed a clever experiment to get around this problem.

In 1881, Michelson performed the experiment with equipment that should just barely have detected such a difference, but the speeds seemed to be the same. There remained a good possibility the failure to detect any difference was due to experimental errors. Michelson designed more elaborate and exact measuring devices with the help of Edward Morley and they repeated the experiment in 1887. This time they expected to detect any difference in speeds even if it were as little as one-fortieth of the predicted difference. No difference was observed. To the astonishment of the physics community, the results of the Michelson-Morley experiment seemed to indicate the speed of a light beam measured in two inertial frames of reference moving as a high speed relative to each other, is exactly the same in both frames of reference!

Summary

The quantitative value measured for certain physical quantities depends on the frame of reference of the observer. For instance, observers in motion relative to each other will obtain different values for the velocity of an object. Physical quantities that depend on the frame of reference of the observer are called relative physical quantities; those that do not are called absolute physical quantities. In general, the mathematical form of a physical law depends on the frame of reference of the observer. However, there does exist a special class of frames of reference in which the laws of mechanics (the laws describing the motion of an object under the influence of a force) are the same and have their simplest forms. These frames of reference are called inertial frames of reference. The laws of mechanics are the same in every inertial frame of reference.

The principle of relativity extends this observation to state every physical law has the same form in any inertial frame of reference. One of the laws from Maxwell's theory of electromagnetism states light travels through a vacuum at a speed of approximately 3×10^8 m/s. If this law is to have the same form in all inertial frames of reference, as required by the principle of relativity, then the speed of propagation of light must be the same in every inertial frame of reference. That is, the speed of light must be an absolute physical quantity. However, it is clear speed is a relative physical quantity and the speed of light should have the value 3×10^8 m/s in only one inertial frame of reference. There seems to be a contradiction between the law of propagation of light and the principle of relativity. Near the end of the nineteenth-century, most physicists had reached a seemingly inescapable conclusion: the principle of relativity does not apply to the laws of electromagnetism. They concluded the laws of electromagnetism are valid only in the inertial frame of reference at rest with respect to the ether, and these laws have different forms in all other inertial frames of reference. This obvious conclusion was thrown into doubt by the startling results of the Michelson-Morley experiment in 1887, which failed to reveal the expected variation of the speed of light in different inertial frames of reference.

Important concepts

Frame of reference; absolute physical quantity; relative physical quantity; inertial frame of reference; principle of relativity; Michelson-Morley experiment.

Questions

1. What is an absolute physical quantity? a relative physical quantity? Give some examples of each.
2. The relative speed of an object with respect to some other object is an absolute physical quantity. Verify this statement in the cases of the examples discussed in this chapter.
3. What is the principle of relativity?
4. How is an inertial frame of reference defined?
5. List some physical properties of a moving body that have the same quantitative value in all inertial frames of reference according to classical physics. List some properties that have different values in different inertial frames of reference according to classical physics.
6. Explain why Maxwell's theory of electromagnetism makes the ether a special inertial frame of reference.
7. What is the significance of the Michelson-Morley experiment?
8. According to the theory of electromagnetism, light travels through the ether with a speed equal to the constant c. Explain why classical physics leads to the conclusion observers in different inertial frames of reference should measure different values for the speed of light.
9. A light source is stationary in the ether. Spaceship A is moving directly toward the light source with a speed of half the speed of light. Spaceship B is moving directly away from the light source with a speed of half the speed of light. According to the classical theories of mechanics, what is the speed of the light as measured on each spaceship?
10. A spaceship moving through the ether at a speed c is traveling parallel to a light beam going in the same direction. According to classical theories, what is the speed of the light as measured on the spaceship?
11. Explain the distinction between the time of observation of an event and the time of occurrence of the event. How are these two quantities related?
12. Explain in your own words why the results of the Michelson-Morley experiment imply the principle of relativity is valid for at least some of the laws of electromagnetism.

13. Using the conversion factor 1 mile = 1609 m, express the best value of c in terms of miles per second.

    The following question is of a more general nature. It has no single correct answer and is just something for you to think about. When possible, a question like this is best answered in conversation with others.

14. Thought experiments such as the one in which observers in two different inertial frames of reference measure the speed of a light beam are often used in physics. Discuss the value and limitations of such experiments.

5 THE SPECIAL THEORY OF RELATIVITY: TIME DILATION

At the end of the last chapter, we found that the Michelson-Morley experiment of 1887 produced results that posed a real puzzle for physicists. The classical theory of electromagnetism predicts that the speed of light (or any other electromagnetic radiation) must have the value 3×10^8 m/s in a frame of reference fixed to the ether. Commonsense reasoning (classical physics) predicts that the measured speed of light will be different in other inertial frames of reference, because speed is a relative physical quantity. The results of the Michelson-Morley experiment showed beyond reasonable doubt that the speed of light is the same in all inertial frames of reference. How can this experimental result be reconciled with physical theory?

In l905, Albert Einstein published his special theory of relativity. (We use the qualifier "special" because in 1915 Einstein extended the theory of relativity in what is now known as the general theory of relativity, the subject of chapter seven.) The special theory of relativity provides an explanation of the Michelson-Morley results. However, it is clear that the Michelson-Morley experiment was not an important factor in Einstein's thinking; in fact, he may not even have been aware of the results of that experiment at the time he worked out his theory. Einstein simply regarded it as logical and necessary that such results would be obtained.

5.1 THE SPECIAL THEORY OF RELATIVITY

The year of 1905 was a remarkable one in the history of science. A young physicist working as a patent examiner in Bern, Switzerland, published five rather short papers. The least important of these papers earned Albert Einstein his Ph.D. from the University of Zurich. The others led to revolutionary changes in three separate areas of physical theory.

The first of Einstein's papers revolutionized our concept of light; we will discuss this paper in detail in Chapter 10. The second was the Ph.D. paper. The third was a paper on the motion of tiny particles such as pollen grains suspended in a fluid. This phenomenon is called Brownian motion after its discoverer. The experimental verification of the predictions of this paper (performed in 1908) represented the first direct experimental evidence that atoms exist, convincing the few remaining skeptics. The fourth paper, "On the Electrodynamics of Moving Bodies," introduced the theory of relativity; perhaps the most famous scientific paper ever written. The fifth paper extended the ideas of relativity and introduced the relationship $E = mc^2$, perhaps the most famous equation in science.

Although Einstein corresponded with some leading physicists, he did most of the work contained in these papers during his spare time, without the benefit of an adequate library or graduate training in physics, and without much discussion of his own ideas with other physicists. Yet the five papers published in 1905 completely revolutionized three distinct fields of study in physics! It is impossible to convey the magnitude of this achievement. If any one of the three revolutionary ideas had been Einstein's only contribution to physics, he almost certainly would have received its highest, the Nobel Prize in Physics. In fact, he was awarded that prize in 1921 for his general contributions to theoretical physics, with special emphasis upon the first of these papers; the one on the nature of light. In this chapter and the one that follows, we'll examine the implications of the special theory of relativity. Einstein's other contributions to physical theory will be covered in subsequent chapters

Philosophy played an important role in Einstein's thinking about physics. He was particularly confident in the validity of the principle of relativity "because it is so natural and simple." He also had no doubts about the validity of Maxwell's law of propagation of light, the prediction that the speed of light has a value equal to the constant c. How then did Einstein deal with the contradictions that arise when we compare the predictions of these two theoretical concepts? In a 1931 book called *Relativity*, Einstein wrote: "As a result of an analysis of the physical concepts of time and space, it became evident [perhaps only to Einstein] that in reality there is not the least incompatibility between the principle of relativity and the law of propagation of light, and that by systematically holding fast to both these laws a logical rigid theory could be arrived at."

Einstein based his special theory of relativity upon the following two postulates:

> **Postulate I. The laws of physics are the same, that is, they have the same mathematical form in any inertial frame of reference.**
>
> **Postulate II. The speed of light in a vacuum is an absolute physical constant.**

These two postulates appear to be contradictory, but Einstein simply accepted the validity of both and "by systematically holding fast" to them, logically derived the physical consequences. This led to a host of very peculiar predictions. For instance, this "logically rigid theory" of Einstein's demanded radical changes in our commonsense and classical notions of time, space, mass, and energy. First, the implications for the physical quantity time.

5.2 TIME IS A RELATIVE PHYSICAL QUANTITY

A relative physical quantity is one whose quantitative description depends on the frame of reference of the observer. In classical physics, velocity and kinetic energy are examples. An absolute physical quantity is one whose quantitative description is the same for all observers in inertial frames of reference. In classical physics, length, mass, and time are examples. In particular, a boy on a train bouncing a ball and an observer watching from the ground would disagree on the velocity of the ball and the distance it had traveled, but surely they would agree on the time interval for the movement of the ball from the boy's hand to the floor and back to his hand. We now reexamine the concept of absolute time in light of the postulates of the special theory of relativity.

Einstein proposed the following thought experiment to illustrate the implications of his theory for the concept of time. It is often called Einstein's train paradox.

A train moves with a speed nearly as large as the speed of light, as measured in a frame of reference fixed to the ground. Three observers are on the train: observer A at the front, observer B at the rear, and observer O exactly in the middle of the train. A fourth observer 0* stands on the ground beside the tracks. Observers A and B fire flashbulbs in such a way that the two light flashes from A and B reach O and O* at the very instant that O passes O* (see Figure 5.1). Which observer, A or B, fired a flashbulb first?

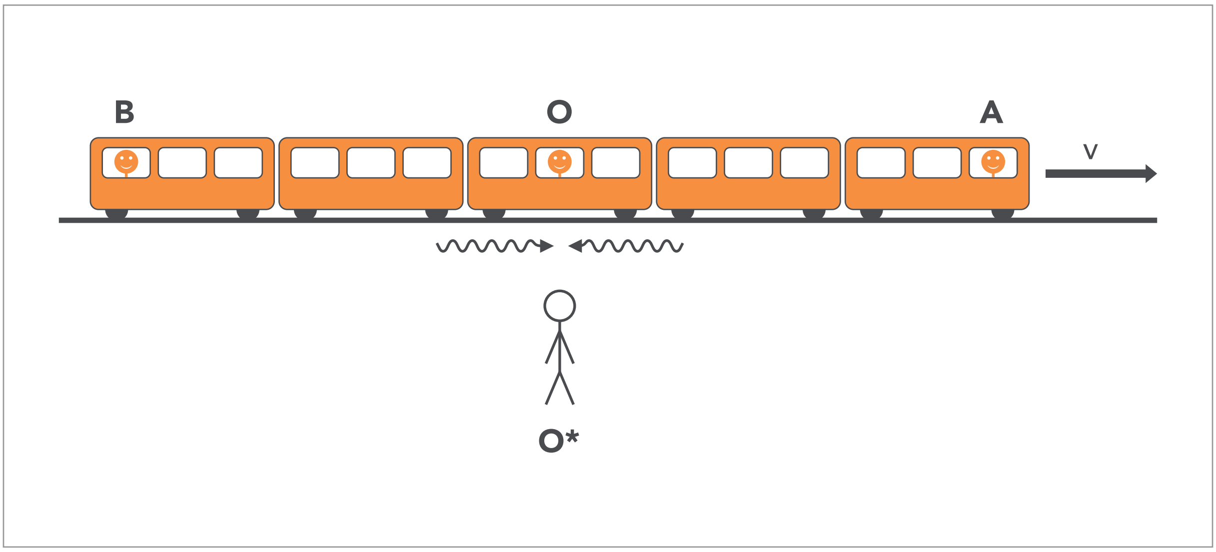


Figure 5.1 Flashes arrive at O and O*

Analyze this thought experiment according to Einstein's special theory of relativity. In particular, assume the validity of his second postulate. First, consider the situation from the viewpoint of observer O. Light travels at a constant speed represented by the symbol c. Observers A and B are at equal distances from O. The two light flashes arrived simultaneously at O after having traveled equal distances at the same speed. Observer O concludes they were emitted simultaneously. Be sure that you understand why this must be so.

Now consider the situation from the viewpoint of observer 0* (see Figure 5.2). The two flashes from A and B arrive simultaneously at the instant when observer O (in exactly the middle of the train) is exactly in front of observer 0*.

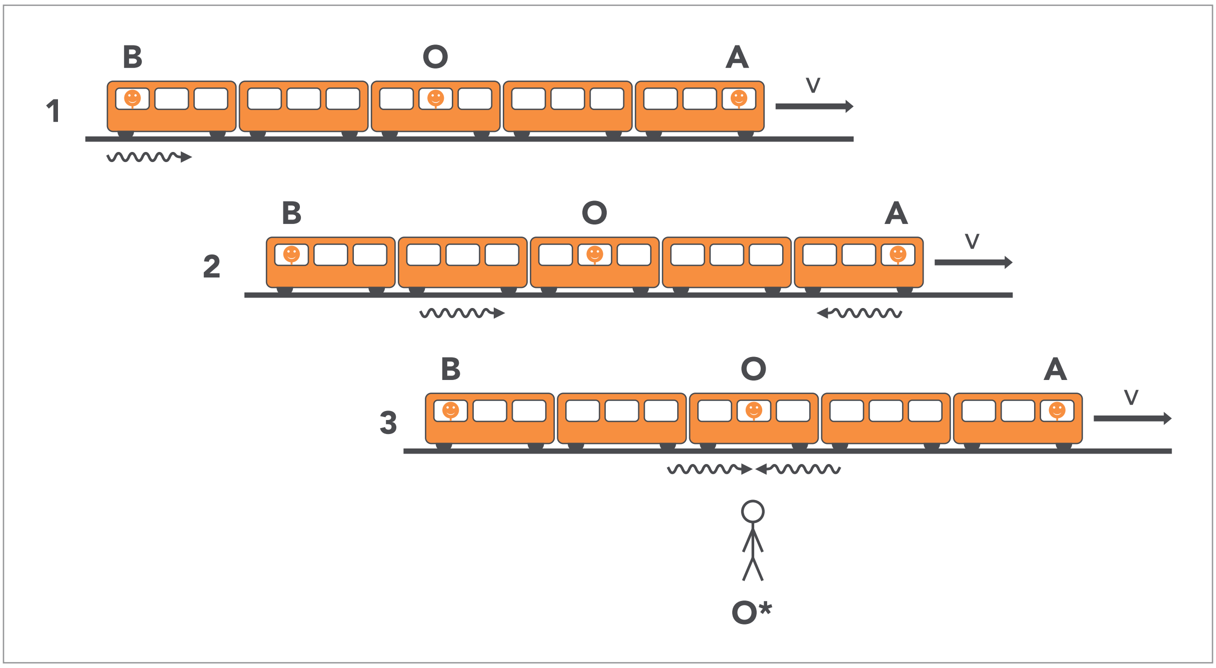


Figure 5.2 Relativity of Time

However, the speed of light is not infinite, and therefore it is clear that the flashes were set off at some time in the past, before the train reached its present location Thus observer O must have been located to the left (opposite to the direction the train is moving) of observer O* when the flashes were emitted. The light flash from B therefore must travel a greater distance to O* than does the light flash from A. Because light travels at a constant speed, it is clear that observer B must have set off his flash first (situation 1 in Figure 5.2). This is necessary in O*'s frame of reference in order the two flashes to arrive simultaneously. At a later time, observer A sets off his flash (situation 2 in Figure 5.2). Examine the line of reasoning carefully. Be sure you see that this conclusion is the only logical one. Observer 0* concludes, based on the assumption that the speed of light is constant in all inertial frames of reference, that B's flash was emitted before A's flash in order for both flashes to arrive simultaneously in O*'s frame of reference (situation 3 in Figure 5.2). The two observers, O and O*, do not agree about the time sequence in which the light flashes are emitted. Observer O concludes they are emitted simultaneously; observer 0* concludes that B's flash was emitted before A's. Who is right? Which observer is measuring the actual sequence of events?

Recall a similar dilemma in discussing velocities in Chapter 4. In that case, we resolved the dilemma by saying that velocity is a relative physical quantity. Its quantitative expression differs in different frames of reference. One measurement is not right and the other wrong because velocity is a relative physical quantity. We can resolve the present dilemma in exactly the same way. We conclude that, according to the special theory of relativity, time is a relative physical quantity. Events separated in space can be simultaneous for one observer but not for another observer who is in motion relative to the first observer. The answer to the question, which flash was fired first depends upon the frame of reference of the observer.

The question of why nineteenth-century physicists were so sure time was an absolute physical quantity is easily answered when the above experiment is treated quantitatively. It turns out that the difference in measured time intervals depends on the relative speed between different frames of reference. Quantitatively, the difference in measured time intervals is large enough to be detectable only if the relative speed of the frames is a significant fraction of the speed of light. Even today, we rarely experience relative speeds greater than about 300 m/s (except for elementary particles) or about 0.0001 percent of the speed of light. At such relative speeds, the two observers would agree that the flashes are simultaneous to within the accuracy of their measuring devices. The disagreement between the two observers would become detectable only when the train moves with a speed that is a significant fraction of the speed of light.

Not only is the time interval between the two events relative, but the order of certain events can also be relative. Consider an observer, O**, who is moving left to right (the same direction the train is moving with respect to the ground) but at a speed greater than the speed of the train. This observer would see the train moving right to left (the observer is outrunning the train). Again, assume that at the instant when O, O*, and O** are at the same location, the two flashes of light arrive simultaneously. Observer O** would conclude (correctly, in his frame of reference) that A's flash was emitted at an earlier time than B's flash. The logic is similar to that used for O* but in this case the front of the train is further away at the time the flashes are emitted. In Einstein's thought experiment, observer O concludes the two light flashes were emitted simultaneously, observer O* concludes that B's was emitted first, and observer O** concludes that A's was emitted first. And all of them are right!

It may have occurred to you that this relativity of the order of events could throw the entire concept of cause and effect into utter disarray. Does this mean that in some frames of reference you were born before your mother was born? The answer is no. If two are causally related – that is, if one event produces the other, then the order of the events is not relative. Analysis using the theory of relativity always yields the same order of causally related events for all observers, regardless of their relative motions. All observers will agree that the mother is born before the child; however, they will disagree on the time interval between the two events. It is also possible that the observers may disagree on which person dies first, assuming there is no causal relation between the two events. There is no causal relationship between the emission of the two flashes in the train paradox and the order of events in this case can be relative.

Comparing Clocks in Relative Motion with Respect to One Another

A clock is a devise for measuring time intervals. Suppose an observer "A" has a device for emitting and detecting light. At a certain distance above the observer, there is a mirror that reflects the light signal back to the device, which detects its return and immediately emit another signal. The repeated detection of a reflected light signal is the equivalent of the ticking of a mechanical clock.

Now suppose a second observer, "B" has an identical clock and moves left to right with respect to observer A. Each observer emits and receives one light signal. Figure 5.3 shows the situation in the frame of reference of observer A. In this frame of reference, how do the time intervals between emission and detection of the light signal compare for the two clocks?

Relativity of Time

Two observers with identical clocks are in different frames of reference. Observer A observes three ticks on B's clock to take longer than three ticks on his own clock, and concludes that B's clock runs slow.

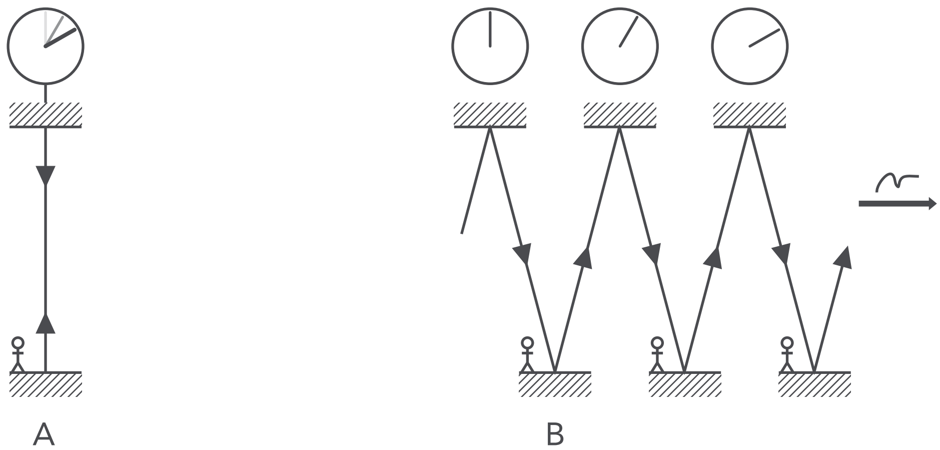


Figure 5.3 Light clocks

According to the second postulate of the special theory of relativity, A observes both light signals to travel at the same speed. It is clear from Figure 5.3 that A observes B's light signal to travel a greater distance than does the signal from A's own device. Thus, A concludes that the time interval between the emission and detection of the light signal by B's clock is greater than the time interval between these two events for A's own clock. In A's frame of reference, observer B's light signal must travel a greater distance at the same speed, and this must take a greater time. During an interval while A counts 60 "ticks" of A's own clock, A observes B's clock to "tick" fewer than 60 times. Thus, A concludes that B's clock runs more slowly than A's own clock.

From the viewpoint of observer B, observer A is moving right to left with speed v. It should be clear that we can apply the same arguments to this situation, obtaining the same final result with one important exception: now observer B concludes that A's clock runs more slowly than B's own clock. Observer A claims that B's clock is running slow, but observer B claims that A's clock is running slow. Who is right? Which clock is actually running slow? Again, we see that time is a relative physical quantity. Each observer is correct in that observer's own frame of reference. In general, a clock in motion relative to the observer runs slower than an identical clock at rest relative to the observer.

With a little geometry and algebra, a quantitative relationship between time intervals measured by the stationary and moving clocks in the above example can be derived. The result is the following:

$$T = \frac{T_0}{\sqrt{1 - (\mathrm{v}^2/\mathrm{c}^2)}}$$

Where "T" is the time interval measured by a clock at rest relative to the observer and "To" is the same interval measured by a clock moving with speed "v" relative to the observer.

5.2.1 TIME DILATION

A clock moving relative to the observer runs more slowly than a clock at rest with respect to the observer. This effect is called time dilation. Time dilation must apply to any kind of a clock, not just to clocks that use light flashes. Suppose that observer A and observer B, while at rest with respect to each other, adjust their clocks so that they both "tick" at a rate of one flash per second. Furthermore, both observers agree that their hearts are beating at a uniform rate of 80 beats per minute as measured by either clock at rest. In this case, the observers' heartbeats are also identical clocks. When the two observers are in motion with respect to each other, each observer can consider the other observer to be in motion. There is no reason to regard either observer as being the one actually at rest or actually in motion. Therefore, we must expect that each observer will continue to measure the rate of his or her own heart as 80 beats per minute according to his or her own clock.

However, each observer will claim that the other's clock is running slow. Therefore, each observer also must conclude that the other's heart is beating at a rate of less than 80 beats per minute. If observe A concludes that one minute on B's clock is longer than one minute on A's own clock, then observer A must also conclude that the 80 beats of B's heart take correspondingly longer than the 80 beats of A's own heart. That is, A must observe B's heart to beat more slowly than A's own. The same reasoning can be applied to any time-keeping device. No matter what the nature of the clock.

All biological processes, not just the heartbeat, are time-dependent. Hence all biological processes must slow down when observed from a frame of reference that is motion with respect to the organism. Time dilation implies that an organism in motion with respect to the observer will age more slowly than a similar organism at rest with respect to the observer. It is important to realize that this is not some sort of "optical illusion." Time itself is a relative physical quantity. From the point of view of an observer on earth, an astronaut moving away from earth at high speed ages more slowly than the observer. Similarly, however, from the point of view of the astronaut the observer back on earth ages more slowly than the astronaut. Although our common sense tells us that this is a contradictory statement, the theory of relativity tells us that our commonsense conclusions are wrong when we deal with very large speeds. Later we will discuss several real physical phenomena that make it clear that this effect actually exists. We will also deal with the seeming paradox that both the observer on earth and the observer on the rocket ship measure the other to be aging more slowly when we discuss the "twin paradox."

Example 5.1

A rocket ship passes the earth at a relative speed of 98 percent of speed of light. At the instant that the ship passes an observer on earth, identical clocks on the earth and on the ship are synchronized and each reads one o'clock. When the earth clock reads two o'clock, what will be the reading on the rocket ship clock in the earth's frame of reference? When the observer on the earth has aged 50 years since the passage of the ship, how much will an astronaut on the ship have aged in the frame of reference of the observer on earth?

Solution

The earth observer measures a time interval of T = 1 hour. We wish to find the time interval T_o for a clock moving with a speed v = 0.98 c. From time dilation, we have

$$T = \frac{T_0}{\sqrt{1-(v^2/c^2)}} \text{ so } T_o = T\sqrt{1-(v^2/c^2)}$$

$$T_o = T\sqrt{1-((0.98c)^2/c^2)} = T\sqrt{1-(0.9604\,c^2/c^2)} = T\sqrt{1-0.9604}$$

$$T_o = T\sqrt{0.04} = T\,(0.20) = (1\text{ hour})(0.2) = (60\text{ minutes})(0.2) = 12\text{ minutes}$$

When the earth clock reads 2:00, the rocket ship clock will read 1:12 according to the earth observer. So, if 50 years pass on the earth clocks, then 0.2 (50 years) = 10 years will pass on the rocket ship's clock -- the astronaut will have aged 10 years. It is very important to realize that in the rocket ship frame of reference, it is the earth observer that is aging more slowly.

Albert Einstein
(1879 – 1955 * Germany)

1905 – developed special relativity and the photon model of electromagnetic radiation.

1916 – developed the general theory of relativity.

1935 – with two colleagues, developed the EPR thought experiment which he believed showed quantum mechanics to be an incomplete description of physical reality.

1921 – Nobel Prize in Physics.

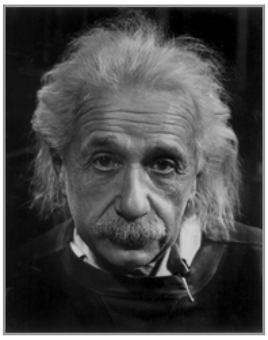


Figure 5.4 Albert Einstein

Summary

The special theory of relativity is based on two postulates: the principle of relativity, and the constant value of $c = 3 \times 10^8$ m/s for the speed of light in any inertial frame of reference. Accepting these two apparently incompatible postulates, Einstein showed they become compatible if we accept certain changes in our concepts of space and time. If these postulates are valid, then time is a relative rather than an absolute physical quantity. Observers in different inertial frames of reference obtain different quantitative measurements of the time intervals between two events. Events that are simultaneous for one observer will not be simultaneous for a second observer in a different inertial frame of reference.

The postulates of the special theory of relativity lead to the prediction of a phenomenon called time dilation: clocks in motion with respect to an observer will keep time more slowly than clocks at rest with respect to the observer. The quantitative relationship between the time intervals on the two clocks is

$$T = \frac{T_0}{\sqrt{1-(\mathrm{v}^2/\mathrm{c}^2)}}$$

where T is the time interval measured on a clock at rest with respect to the observer, and T_o is the time interval measured on a clock moving with speed v relative to the observer. This time-dilation effect becomes measurable only when the speed v is a significant fraction of the speed of light.

Important concepts

Special theory of relativity; time is a relative physical quantity; time dilation.

Questions

1. State and explain the two postulates of the special theory of relativity. Explain Einstein's reasons for adopting each of them.
2. State and explain the time-dilation equation. Which clocks appear to run more slowly to an observer, those at rest or those moving in the observer's frame of reference? Discuss whether this effect is a real or an apparent one. (Be careful, this last question is a tricky one.)

3. A person has a pulse rate of 90 beats per minute. An observer moves with a speed of 0.5 c relative to this person. What does this observer measure as the pulse rate of the person according to clocks in the observer's frame of reference?
4. An observer notes that a moving clock indicates a time interval of l5 minutes while a clock at rest with respect to the observer indicates that the same time interval is 20 minutes. What is the speed of the moving clock in the observer's frame of reference?
5. A person lives for 70 years according to clocks in the person's own frame of reference. How long does this person live according to the clocks of an observer moving at a speed of 80 percent of the speed of light relative to the person?
6. Calculate the value of $\sqrt{1-(v^2/c^2)}$ for the following values of v: 0.l c; 0.3 c; 0.5 c; 0.7 c; 0.9 c; 0.96 c; 0.99 c.
7. Can the time-dilation equation be applied when v = c? What value does To/T approach as v approaches c? What happens when v is greater than c?
8. A spaceship is traveling away from earth at a constant speed of one half the speed of light. A light flash is emitted on earth and travels toward the spaceship. What is the speed of the light according to observers on the earth? With what speed does the light pass the spaceship according to observers on the spaceship?
9. There exists a subatomic particle called the π^o meson (or neutral pi meson). Such a particle exists for only about 10^{-16} seconds. as measured by a clock at rest with respect to the particle. Suppose that a π^o meson is moving with a speed of 0.5 c relative to the observer? According to the observer's clocks will the lifetime of the particle be greater or less than 10^{-16} seconds? What will be the measured lifetime of the particle according to the observer's clocks?

   The following questions are of a more general nature. They have no single correct answer and are just something for you to think about. When possible, they are best answered in conversation with others.
10. According to the special theory of relativity, simultaneity is a relative concept. Do you think that the concept of simultaneity is still a useful one in physics? Explain your reasoning. Is it meaningful to say that an event in San Francisco occurred at the same instant as an event in New York?
11. The special theory of relativity is not difficult to understand; it is simply difficult to accept. Do you agree? Discuss your reasons.
12. In Einstein's train paradox, suppose that observers A and B had thrown baseballs rather than sending light flashes. Explain how that would affect the argument ..
13. Explain how both postulates of the special theory of relativity violate classical physics.
14. Explain in some detail why clocks consisting of a light bean reflected back through a fixed distance, together with the postulated of the special theory of relativity requires that a clock in motion with respect the observer must keep time slower than an identical clock at rest with respect to the observer.

Albert Einstein (1879 – 1955)

Albert Einstein was born in the Bavarian region of Germany. He was a quiet, dreamy child, late in learning to speak His parents feared he was retarded. Once in school, Einstein was quick in mathematics but slow in other subjects. Einstein thoroughly disliked school and applied himself little. When he was sixteen, one of his teachers told him he would never amount to anything and advised him to leave school. Einstein promptly took this advice.

In 1895, after a vacation in Italy to avoid military service, Einstein applied for admission to the Swiss Federal Polytechnic Institute in Zurich (the MIT of Central Europe). After failing the entrance exam, except for the mathematics part, Einstein took a year to complete the European equivalent of a high school diploma and was finally admitted to the Institute He spent most of his time there reading and doing experiments on his own, seldom attending classes. His fellow students found him charming and witty. They sensed that he had a peculiar sort of intelligence. With the aid of class-notes borrowed from a friend, Einstein was able to pass his final exams and graduate. He wanted to continue in school and work toward an advanced degree but no professor would take him on as a student.

In 1901, the father of that same friend whose class-notes he had borrowed, used his influence to obtain an appointment for Einstein as a patent examiner in the Swiss Patent Office in Bern, Switzerland. The Patent Office job was ideal for Einstein. He examined patents on the subject that interested him most, electromagnetism. He had plenty of spare time to read physics and engage in a favorite activity, the creation of thought experiments. For example, Einstein would imagine he was traveling away from the earth at the speed of light. What would he see? Einstein was a visual thinker. He often visualized the answer to such thought experiments before he worked out the mathematical solutions.

By 1909, Einstein's work was becoming recognized and he began to receive offers of academic positions. In 1914 he became Director of the Kaiser Wilhelm Physical Institute in Berlin with no specific duties. In 1915 he published the general theory of relativity, his most significant contribution to physics. In 1921 he was awarded the Nobel Prize for his work on the photoelectric effect, the subject of one of his 1905 papers.

As a socialist, a pacifist, and a Jew, Einstein came under increasing attack in Germany during the 1920s and early 1930s The Nazis burned his books and denounced his theories as "Jewish physics." Einstein was out of the country when Hitler came to power in 1933, and never returned. This almost certainly saved him from arrest and probably death.

Einstein accepted a position at the Institute for Advanced Studies at Princeton, where he continued to work until his death. He became an American citizen in 1940.

Einstein spent the majority of his adult life recognized as the most brilliant person in the world. His name became synonymous with genius. Yet no one could wear this mantle with more gentle humility. He greeted child and king alike with good-humored respect. His great passions were physics, the violin, and peace.

Without regard for his own well-being, Einstein consistently spoke out on the side of peace and justice. While still in Berlin during the First World War, he denounced German militarism. In the United States in the 1950s, he urged people to defy Senator McCarthy's anticommunist hearings. Although he felt obliged to join in calling President Roosevelt's attention to the Nazis' work on an atomic bomb, he was distraught when Hiroshima and Nagasaki were destroyed by nuclear weapons. He deeply regretted the small role he had played in their construction. His last public act was to join with Bertrand Russell and many other scientists and scholars in an unsuccessful attempt to bring about a ban on the further development of nuclear weapons.

6 THE SPECIAL THEORY OF RELATIVITY: LENGTH CONTRACTION AND VELOCITY ADDITION

In Chapter 5, we found that time is a relative rather than an absolute physical quantity. That is, the quantitative measurement of the time interval between two events depends on the relative motion of the observer. Specifically, a clock in motion is observed to keep time more slowly than a clock that is at rest with respect to the observer. Now consider the implications of the special theory of relativity for the physical quantity length.

6.1 LENGTH CONTRACTION

In classical physics, length was considered to be an absolute physical quantity. That is, the quantitative measurement of distance between two points in space was considered to be independent of the frame of reference of the observer. In Chapter 5, we explored some of the implications of the special theory of relativity and, in Example 5.3, we saw that an earth observer and a moving observer disagree about the time interval required for a trip from the earth to the moon. In the frame of reference of the earth observer, the rocket ship travels at a speed of 0.80c over the distance of 3.8×10^8 meters (or 240,000 miles) between earth and moon. The trip takes a time of 1.60 s on the earth clocks and 0.96 s on the rocket ship clock.

In the frame of reference of the rocket ship, the earth and moon are passing the stationary rocket ship at a speed of 0.8c or 2.4×10^8 m/s (see Figure 6.1). The time interval between the passage of the earth and the passage of the moon is 0.96 s as measured on the rocket ship clock.

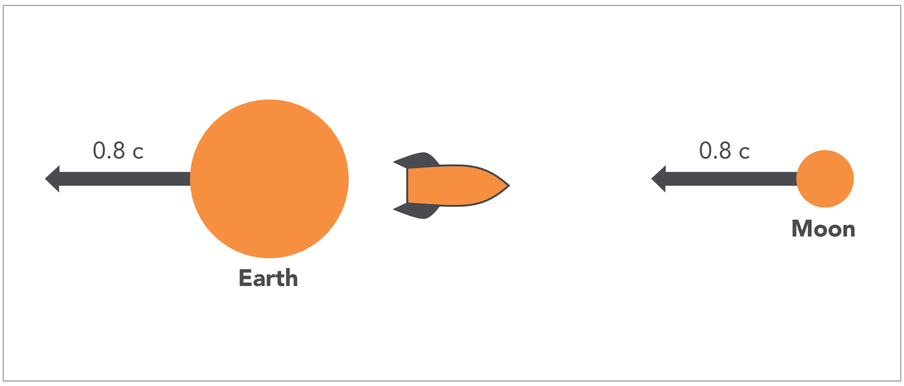


Figure 6.1

From the relationship d = vt, the rocket ship observer concludes that the distance between earth and moon is

$$d = (2.4 \times 10^8 \text{ m/s}) \times (0.96 \text{ s}) = 2.3 \times 10^8 \text{ m (or 144,000 miles).}$$

Thus, the rocket ship observer and the earth observer disagree about the distance between the earth and the moon.

In the frame of reference of the rocket ship, the distance between the earth and moon is 2.3×10^8 m; in the frame of reference of the earth and moon, the distance is 3.8×10^8 m. The separation in space between two objects is not an absolute physical quantity. Rather, it is relative to the frame of reference of the observer who measures that separation. The value obtained for the distance between two objects by an observer moving with respect to those objects will be smaller than the value measured by an observer at rest with respect to the objects. This shortening of distance due to relative motion is called length contraction

In Example 5.3, the relative speed of the two frames of reference is 0.8 c, so the factor $\sqrt{1-\left(\frac{v^2}{c^2}\right)}$ is equal to 0.6. Also, comparing the distance measured by the rocket ship observer, 2.3×10^8 m, to the distance measured by the earth observer, 3.8×10^8 m, shows the distance in the rocket ship frame is exactly 0.6 times to distance in the earth frame of reference. That is, $(3.8 \times 10^8 \text{ m}) \times 0.6 = 2.3 \times 10^8$ m.

This is not a coincidence; the relationship follows logically from the reasoning we used in finding the distance measured by the rocket ship observer (because we used time dilation in that reasoning). We can express the relationship as a length-contraction equation. Let L be the distance between two objects as measured by an observer at rest with respect to those objects and let L_o be the distance between the objects as measured by an observer moving in the direction between them with speed v relative to the objects. Then

$$L = L_o \sqrt{1-\left(\frac{v^2}{c^2}\right)}$$

Note that length contraction is observed only when the observer is moving along the direction between the two objects. In other words, measured distances contract only along the direction of the observer's motion; the moving observer will agree with a stationary observer on the value of distances measured at right angles to the motion of the moving observer.

The length-contraction phenomenon also leads to a distortion of the dimensions of an object that is in motion relative to the observer. A moving object will appear contracted in the direction of its motion by the factor $\sqrt{1-\left(\frac{v^2}{c^2}\right)}$ as compared to the corresponding measurement made by an observer at rest with respect to the object.

Example 6 . 1

A square object has sides of length 2 m as measured by an observer at rest with respect to the object. An observer passes the object from left to right with a speed of 0.98 c relative to the object. What dimensions does this moving observer measure for the object?

Solution

In the observer's frame of reference, the object is moving past at a speed of 0.98 c. Due to length contraction, the observer will measure the width of the square (in the direction along the square's path of motion) as

$$L = Lo\sqrt{1-\left(\frac{v^2}{c^2}\right)} = (2\text{ m})\sqrt{1-\left(\frac{(.98c)^2}{c^2}\right)} = (2\text{ m})\sqrt{1-0.96}$$

$$L = (2\text{ m})\sqrt{0.04} = (2\text{ m}) \times 0.2 = 0.4\text{ m}$$

The length of the sides perpendicular to the direction of motion are unaffected. Therefore, the moving observer sees a rectangle with dimensions of 2 m by 0.4 m. Again, it is important to realize that this is not an optical illusion; these are the dimensions of the object as (correctly) measured by the moving observer.

We can summarize time dilation and length contraction as follows.

1. **Every clock goes at its fastest rate when it is at rest relative to the observer. If it moves relative to the observer with a speed v, then its rate is slowed by the factor $\sqrt{1-\left(\frac{v^2}{c^2}\right)}$.**
2. **Every object is longest when it is at rest relative to the observer. If the object moves relative to the observer with speed v, then it is contracted along the direction of its relative motion by the factor $\sqrt{1-\left(\frac{v^2}{c^2}\right)}$. This relationship also applied to the distances between two objects.**

No doubt these two relativistic effects seem strange. They are not consistent with our commonsense, intuitive notion of the way things are, and it should be clear why this is so. Time dilation and length contraction produce detectable effects only when the factor $\sqrt{1-\left(\frac{v^2}{c^2}\right)}$ is measurably different from one. This occurs only when the speed v is an appreciable fraction of the speed of light and we do not normally experience such speeds. Therefore, it is not surprising that our commonsense ideas of how things behave fail when we extend them so far beyond the range of experiences. At speeds of the sort with which we normally deal, the effects of time dilation and length contraction are far too small to be detected with ordinary measuring devices. Thus the special theory of relativity predicts results that are quite consistent with our commonsense expectations for the range of experiences that form the foundation of our commonsense ideas

As an illustration of the inadequacy of our commonsense notions in dealing with very high speeds, consider the following situation. Two observers approach each other with a large relative speed (see Figure 6.2).

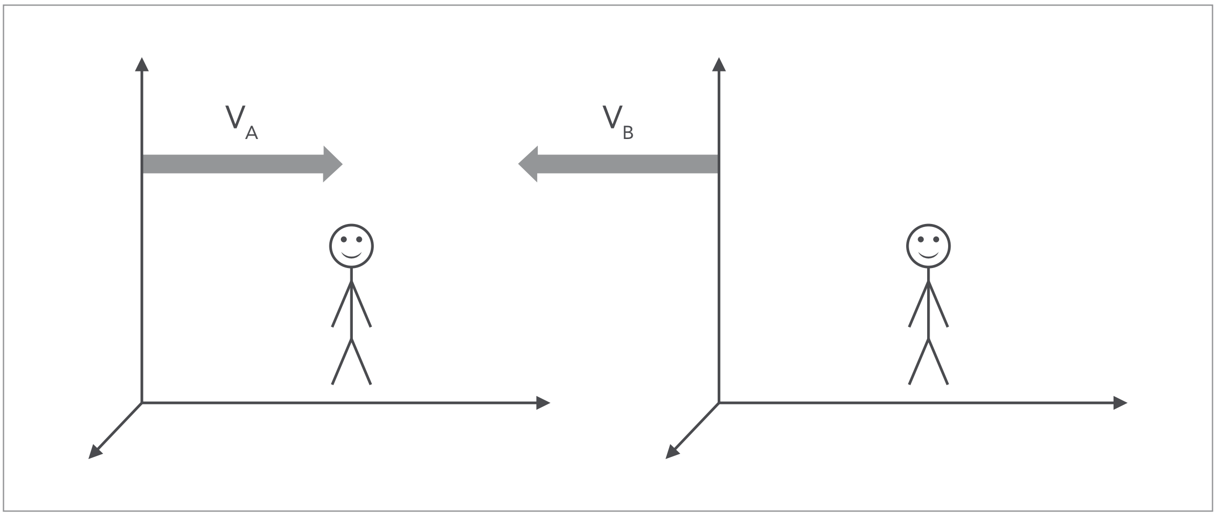


Figure 6.2

They approach each other in such a way that the origins of their two frames will pass through each other. That is, there is some instant of time when both origins occupy the same point in space. At this instant a flash of light is emitted from the common origin, and the light expands outward through space from this point as time passes.

Observer B continues with the uniform speed v relative to observer A, so that the position of B relative to A at some later time will be as shown in Figure 6.3.

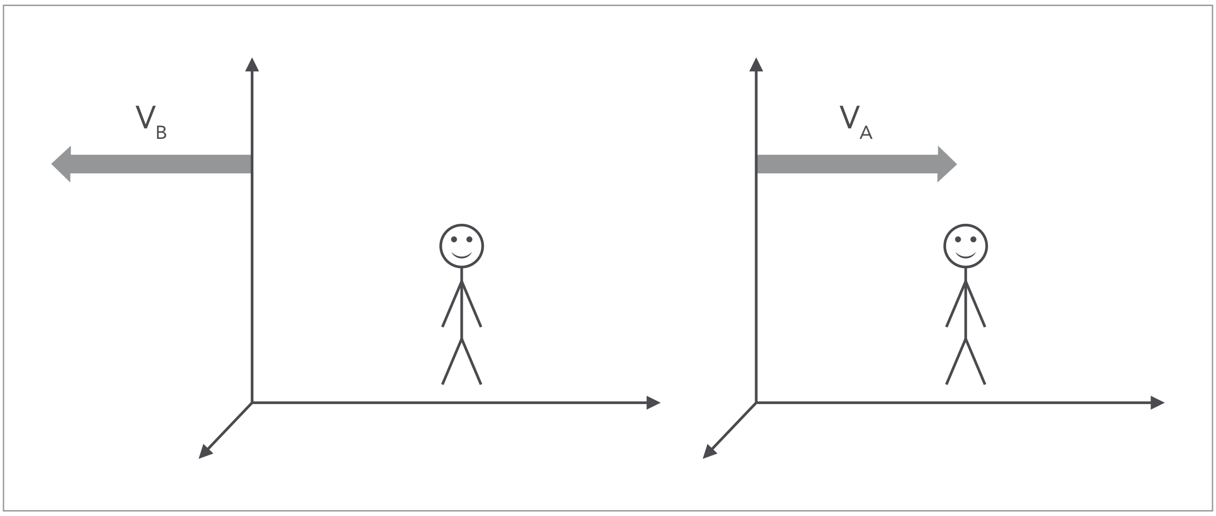


Figure 6.3

Now consider the expanding light signal. The special theory of relativity requires that the speed of light always be measured as the value c by any inertial observer, regardless of the relative motions of the observer and the light source. Thus, observer A measures the same value c for the speed of the light in the forward and backward directions, in the up and down directions, in the left and right directions -- in fact, in all directions Therefore, observer A must observe an expanding sphere of light centered at the origin of A's frame of reference.

However, we can apply exactly the same reasoning to observer B. We must conclude that observer B observes an expanding sphere of light with its center at the origin of B's frame of reference. But the two frames of reference are in motion relative to each other, so the two origins are not at the same point in space at times after the light flash. How then can they both be at the center of the same sphere of light? Using our commonsense notions of space and time, we perceive a paradox. We seem to be led to two conclusions that are mutually contradictory. However, when we apply the concepts of relative space and time according to the special theory of relativity, there is no paradox at all. Each observer concludes that the other's origin is not at the center of the sphere. The opinions of the two observers are consistent because both time and space are relative physical quantities.

The discussion so far has been based on thought experiments using imaginary observers using hypothetical devices and traveling at impossibly high speeds. Such thought experiments illustrate the predictions of relativity, but they cannot provide evidence that a theory is valid. Only actual physical experiments or measurements can provide physical data to test the validity of the physical theory. You might ask (in fact, you must ask) whether the results of our thought experiments have been confirmed by any actual physical experiments. They have. The special theory of relativity is neither hypothetical nor conjectural. All of its physical consequences, no matter how strange they may seem to us, have been confirmed to an incredibly high degree of accuracy by actual experiments. There can he no doubt that the laws of the special theory of relativity are valid over the entire range of experiences now available to us (so long as we restrict ourselves to the inertial frames of reference that are specified in the postulates of the theory).

It is important to remember that the factor $\sqrt{\mathbf{1}-\left(\frac{\mathbf{v}^2}{\mathbf{c}^2}\right)}$ is very nearly equal to one for speeds that are small compared to the speed of light. For example, when v = 3000 m/s = 6711 miles/hour, the factor is equal to 0.999995 Although 6711 miles per hour is a very great speed in our normal range of experiences, it is a very small speed in the equations of the special theory of relativity. For such speeds (and smaller speeds), time and length may be treated as absolute physical quantities because the effects of time dilation and length contraction are far too small to be detected by ordinary laboratory measurements. Thus, the special theory of relativity satisfies the correspondence principle (see Chapter 3); its results do agree with the predictions of the classical theory when applied to the range of experiences where the classical theory is known to be valid.

6.2 SPACE-TIME

For the first year after its publication, Einstein's special theory of relativity seems to have attracted little attention among physicists, although many of them were in fact studying it and exploring its implications. The first published response to Einstein's theory came about a year after his publication, and it was a claim by a prominent experimental physicist to have obtained experimental results inconsistent with Einstein's theory. (It was another decade before physicists were able to find the errors in this experimental work, but Einstein's faith in his theory was not shaken. He was confident -- correctly, as it turned out -- that his theory would prove to be consistent with experimental results.) It was not long, however, before a number of other physicists began to express their support for the theory.

One of the earliest supporters of the theory of relativity was Hermann Minkowski, who had been Einstein's mathematics teacher at Zurich Polytechnic In a 1908 address to the Congress of German Scientists and Physicians, Minkowski said, "From now on space by itself, and time by itself, are destined to sink completely into shadows, and only a kind of union of both [will] retain an independent existence." It is this union of the two concepts that is now indicated by the term space-time.

In classical Newtonian physics, space and time are considered two quite distinct physical quantities. Separation in space is regarded as completely independent of separation in time, and vice versa. In this chapter and the preceding one, however, we have seen this is not the case in relativistic physics. Einstein's train paradox suggests that for events separated in space can be simultaneous for one observer and not for another observer in a different inertial frame of reference. The greater the separation of the two events in space, the more the two observers disagree about their separation in time. The properties of space and those of time are closely interwoven. It is impossible to construct a relativistic physics in which time and length are treated as completely independent physical quantities. Therefore, Minkowski suggested a reformulation of the special theory of relativity in which the three dimensions of space and the one dimension of time are represented as a single four-dimensional space-time.

This new worldview has profound implications. When this version is further explored, momentum and energy are found to be similarly united concepts. The separate laws of conservation of momentum and conservation of energy become a single conservation law in terms of the four-dimensional space-time. The phenomena and laws of electricity and magnetism are similarly much more profoundly and simply united in this new worldview.

Space and time are relative, but space-time is not. In relativistic physics, the space-time interval between two events is an absolute physical quantity, independent of the relative motion of an observer. Although two observers in different inertial frames of reference disagree about which portion of the space-time interval should be regarded as separation in space and which should be regarded as separation in time, they both agree on the same value for the space-time interval itself.

The term space-time should not be taken to imply that time is simply a fourth spatial dimension. Time has properties that are different from those of space. For example, we can readily move back and forth in space, but we can move only one direction in time. This distinction also exists in space-time. For example, there is no way for you to move through space-time to affect an event that has already happened. However, it is theoretically possible for an astronaut to travel to a distant star and return only 50 years older to find that 1000 years have passed according to earth clocks. Although there is no way for you to meet your great-great-great-grandparent (who is already dead), it would be possible for you to meet your great-great-great-grandchild. (Is this a paradox? Think about it carefully.)

> *Now my suspicion is that the universe is not only queerer than we suppose, but queerer than we can suppose.*
>
> – J. B. S. Haldane

6.3 SOME RELATIVISTIC EXAMPLES

The ideas of length contraction and time dilation seem so strange that it will be useful to examine a few more examples; If we accumulate enough experiences with these ideas, they should eventually become more familiar and reasonable.

Example 6.2

A muon is an elementary particle whose mass is less than that of a proton but greater than that of an electron. Unlike these two particles, however, the muon is unstable. That is, a muon will not exist indefinitely as a muon. After some time interval, the muon will disintegrate to form other types of particles. Extensive laboratory studies of muons have been carried out. A muon created with low speed has an average lifetime of 2.2×10^{-6} s (or 2.2 microseconds). That is, when a muon is created with low speed, it will exist for about 2.2×10^{-6} s before it disintegrates into other types of particles. Many muons are produced by natural processes in the upper atmosphere at distances of about 3000 m (approximately 2 miles) above the earth's surface. These muons are created with speeds very near the speed of light. Is it likely that such muons will reach the surface of the earth?

Solution

First, we answer this question in terms of classical physical theories. A muon is created 3000 m above the surface of the earth. If it is moving straight downward with a speed of approximately 3×10^{8} m/s (the speed of light), how far can it travel during its lifetime of 2.2×10^{-6} s? We use the relationship d = vt to obtain

$$d = (3 \times 10^{8} \text{ m/s}) \times (2.2 \times 10^{-6} \text{ s}) = 6.6 \times 10^{2} \text{ m} = 660 \text{ m}$$

Therefore, we conclude that the average muon will travel less than one-quarter the distance from the upper atmosphere to the earth. We predict that the number of muons detected in the upper atmosphere will be much larger than the number detected at the earth's surface. Experiments show, however, that most of the muons do reach the earth's surface.

Do you see why our prediction is invalid? We are dealing with very large speeds (speeds approaching the speed of light). Therefore, we must use the special theory of relativity in our analysis of the problem. Consider the situation from the earth frame of reference. The muon must travel 3000 m from the upper atmosphere to the surface of the earth. Assume the speed of a typical muon is 98% of the speed of light; that is, v = 0.98 c. Because of time dilation, all processes occurring in the muon's frame of reference will appear to be slowed in the earth frame of reference. The average muon will disintegrate after 2.2×10^{-6} s in the muon frame of reference, but the earth observer will measure a longer time interval on the earth clock. According to the time-dilation formula, the time interval T indicated on the earth observer's clock is equal to

$$T = \frac{T_0}{\sqrt{1-(v^2/c^2)}} = \frac{2.2 \times 10{-}6\, s}{0.2} = 11 \times 10^{-6}\ \text{s}$$

To the earth observer, the high-speed muon has 11×10^{-6} s to make it to the surface of the earth. We use the relationship d = vt to obtain

$$d + (0.98 \times 3 \times 10^{8}\ \text{m/s}) \times (11 \times 10^{-6}\ \text{s}) = 32.34 \times 10^{2}\ \text{m} = 3234\ \text{m}$$

Using time dilation we predict that most of the muons will make it to the surface, consistent with observations.

Example 6.3

Show that the result is the same if we analyze Example 6.2 from the frame of reference of the muon.

Solution

In the muon's frame of reference, the muon is at rest and the atmosphere is moving at a speed of 0.98 c. Therefore, the distance from the upper atmosphere and the surface will be length contracted. L is the distance between the upper atmosphere and the surface in the muon's frame of reference and L_o is the distance in the frame of reference where the upper atmosphere and the surface are at rest (the earth frame of reference). Look back at the length-contraction equation and make sure you agree.

$$L = Lo \sqrt{1 - \left(\frac{v^2}{c^2}\right)} = (3000 \text{ m}) \text{ x } (0.2) = 600 \text{ m}$$

In the muon's frame of reference, the distance from the upper atmosphere where it is created to the surface of the earth is only 600 m. In its lifetime of 2.2 x 10^{-6} s, the distance of the muon can travel in its frame of reference is

$$d = v\ t = (0.98) \text{ x } (3 \text{ x } 10^8 \text{ m/s}) \text{ x } (2.2 \text{ x } 10^{-6} \text{ s}) = 647 \text{ m.}$$

Again, a relativistic treatment indicates most muons will make it to the surface, consistent with observations.

Example 6.4

The unit meter is inappropriately small for expressing distances in astronomy. A more convenient unit of distance is the light-year – the distance light travels in one year. (1 year = 3.16 x 10^7 seconds)

$$1 \text{ Ly} = (3 \text{ x } 10^8 \text{ m/s}) \text{ x } (3.16 \text{ x } 10^7 \text{ s}) = 9.48 \text{ x } 10^{15} \text{ m}$$

Note that the light-year is a unit of distance not a unit of time. It is an unimaginably large distance. The distance to the nearest star is 4.22 Ly away and the distance to the center of our Milky Galaxy is approximately 26, 000 Ly,

A star is determined to be 10 Ly distant from earth. A rocket ship leaves earth and travels to the star at a speed of 0.98 c. How long will the trip take according to earth clocks? In the rocket ship frame of reference, what is the distance between the earth and the star? How long will the trip take according to rocket ship clocks?

Solution

In the earth frame of reference, the rocket ship travels a distance of 10 Ly at a speed of 0.98 c, Distance equals velocity times time. So

t = d/v = 10 Ly/ 0.98 c

We could convert the distance to meters and the speed to m/s and solve for the time for the trip in seconds, but seconds is not a good unit for space travel. There is a way to avoid all that arithmetic Distance in light-years divided by speed as a fraction of the speed of light equals time in years.

t = 10 Ly/ 0.98 c = 10.2 years

This is a reasonable answer. The ship is traveling at almost the speed of light, so we would expect the trip to take a little more than 10 years.

In the rocket ships frame of reference, the distance to the star will be shortened by length contraction. Using the length-contraction equation, the distance to the star will be

$$L = Lo \sqrt{1-\left(\frac{v^2}{c^2}\right)} = (10\ \text{Ly}) \times \sqrt{1-\left(\frac{(.98\ c)^2}{c^2}\right)} = (10\ \text{Ly}) \times \sqrt{1-0{,}96}$$

$$L = (10\ \text{Ly}) \times \sqrt{0.04} = (10\ \text{Ly}) \times (0.2) = 2\ \text{Ly}$$

To determine how long the trip will take on the rocket ship clocks there are two ways to do this. Knowing how long the trip took on earth clocks, we could use time dilation to determine the time of rocket ship clocks. Or, knowing the distance in the rocket ship frame, we could determine the time from that.

First time dilation. The time measured for the trip on clocks at rest on earth was 10.2 years. The rocket ships were moving at a speed of 0.98 c with respect to the earth.

$$T = \frac{T_0}{\sqrt{1-(v^2/c^2)}} \text{ or } T_o = T \sqrt{1-(v^2/c^2)} = (10.2 \text{ years}) \times (0.2) = 2.04 \text{ years}$$

Now, using length-contraction, we know that the distance traveled in the rocket ship frame is 2 Ly. In that frame of reference, the earth is receding at 0.98 c and the star is approaching at 0.98 c.

d = v t or t = d/v – 2 Ly/ 0.98 c = 2.04 years.

If you get confused about how to apply the equation of time dilation and length contraction, just return to the basic definitions of the effects. The distance between two objects (or the length of an object) is greatest when measured by an observer at rest with respect to the objects. The time interval between two events measured on a clock moving relative to the observer will always be shorter than the time interval measured on a clock at rest with respect to the observer. The factor $\sqrt{1-(v^2/c^2)}$ must always be smaller than one, so you must multiply by this factor to obtain a smaller value or divide by this factor to obtain a larger value.

Length contraction and time dilation have interesting implications for space travel. In a space-time universe, space travel done at an appreciable fraction of the speed of light means time travel.

6.4 THE TWIN PARADOX

Because time passes at different rates in different frames of reference the theoretical possibility exists for twins in different frames of reference being reunited to find one twin older than the other. Suppose that one twin stays home and the other is in a rocket ship traveling at, say 98% of the speed of light. The time dilation factor is 0.20 as in the earlier calculations.

From the point of view of the stay-at-home twin, the traveling twin will only age 6 years in the same time interval he ages 30 years. If the traveling twin turns around and immediately returns home, again at 98% of the speed of light, the return trip will take 30 years on the stay-at-home twin's clock, while again only 6 years pass on the rocket ship clocks. The twins will be reunited with the stay-at-home 60 years older than when the trip started while the traveling twin will be only 12 years older.

Twin paradox - the rate at which time passes can be different for observers in different frames of reference. In this case it is the accelerated frame in which less time has passed.

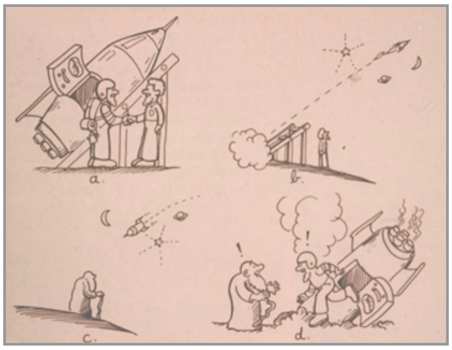


Figure 6.4 Twin paradox

This scenario is known as the Twin Paradox, although it is not a paradox at all. There is nothing paradoxical about the above calculation. The above situation was described from the point of view of the stay-at-home twin. This is because of the two, his frame of reference is the only inertial frame throughout the trip out and back. The rocket ship must turn around in order for the two to be reunited. That is, it must decelerate to a stop, then accelerate back up to 98% of the speed of light. During the time it takes to do this, the rocket ship frame is not an inertial frame, and the laws of the special theory do not apply.

However, the later general theory of relativity, as discussed in the next chapter, can be used in a non-inertial frame of reference. For the first half of the trip, both frames are inertial (neglecting the initial acceleration). The traveling twin agrees that he has aged 6 years, but because the stay-at-home twin is the moving frame, the rocket observer see his twin age only 1.2 years (6 years times 0.20). On the return half of the trip, both are again inertial observers and the rocket travelers again ages 6 years and sees his twin age another 1.2 years. However, during the turn around, the equations of special relativity do not apply, and the equations of the general theory must be used. According to the general theory, during the turn around, the rocket observer will see the stay-at-home clocks suddenly speed up dramatically. According to precise calculations, the stay-at-home clock will speed forward and tick off 57.6 years during the turn-around time, even if it takes only a few minutes on the rocket ship clocks. Thus each twin will be able to explain why the stay-at-home twin has aged 60 years while the traveling twin has only aged 12 years.

This is a thought experiment and cannot be fully accepted until an actual experiment has been done. In fact it has. Modern atomic clocks can measure time accurately to a very tiny fraction of a second. In 1971 four such clocks were synchronized carefully. One was flown around the world west to east, and one flown east to west. The other two remained on the ground. Although the resulting differences were only a few billionths of a second, comparison of the clocks after the experiment verified the predictions of relativity. Additional experiments since 1971 have also confirmed these results.

6.5 VELOCITY ADDITION

Velocity is a derived physical quantity that depends on the distance traveled and on the time required to travel that distance. In view of the relativistic effects of length contraction and time dilation, we may suspect that velocity also will display some peculiar relativistic effects. We now examine the question of how the velocity depends upon the frame of reference of the observer. Consider two observers whose frames of reference are in relative motion with a speed v with respect to each other, with B moving left to right away from A Suppose that observer B measures the velocity of an object and determines that it is moving left to right with speed u in B's frame of reference. If observer A measures the velocity of the object, what value will A obtain in A's frame of reference? According to the classical laws of physics, the answer is obvious. Observer A will determine that the object is moving left to right with speed $V = v + u$.

We used this commonsense reasoning in Chapter 4 when we began to discuss frames of reference. If a person standing in a train that is moving with a speed of 30 m/s relative to the ground throws a ball forward with a speed of l0 m/s relative to the train, then we expect naturally that an observer on the ground will see the ball moving with a speed of 40 m/s relative to the ground. If the person throws the ball toward the back of the train with a speed of 20 m/s, then the ball should appear to be moving forward with a speed of 10 m/s in the frame of reference of the ground observer.

This simple intuitive reasoning, however, leads to trouble when we deal with speeds that are significant fractions of the speed of light. Suppose observer A is at rest and measures the speed of a light beam to be $c = 3 \times 10^8$ m/s. Observer B is in a frame of reference moving with a speed v relative to observer A. The simple reasoning we have used before tells us that B will measure speed of the light beam as $V = c + v$. However, this prediction contradicts the second postulate of the special theory of relativity (as well as the results of the Michelson-Morley experiment). The speed of light must be c, no more and no less, when measured in any inertial frame of reference. The correct relativistic expression for velocity addition is

$$V = (u + v)/(1 + (uv/c^2)$$

where positive values for speeds u and v indicate velocity in one directions and negative values are used to indicate speeds in the opposite direction. We will not worry in this book about how to add velocities that do not lie along a single direction in space. Notice that, if both u and v are small compared to c, then the term uv/c^2 is negligibly small. In that case, the velocity-addition formula is approximately equal to the classical formula, $V = u + v$. This is required by the correspondence principle.

Is this expression for velocity addition consistent with the Michelson-Morley results and with the second postulate? Observer B determines the speed of a light beam to be $u = c$. If observer A measures the speed of the same light beam, what value will be obtained in A's frame of reference? From the velocity-addition formula,

$$V = (u + v)/(1 + (uv/c^2) = (c + v)/(1 + (cv/c^2)$$

Multiplying both numeration and denominator by c, we get

$$V = c(c + v)/ (c + c^2v/c^2) = c(c + v)/ (c + v) = c$$

Thus the velocity of light is independent of the frame of reference of the observer. The speed of light is equal to c in any inertial frame of reference. (Note that we could use any value for v so our result must be valid for any two frames of reference in motion with respect to each other.) Therefore, the relativistic velocity-addition formula is consistent with the second postulate of special relativity and the Michelson-Morley experiment.

Example 6.5

A spaceship moves way from the earth at a speed v = 0.75 c with respect to the earth (see Figure 6.5). The spaceship launches a missile in the direction away from the earth with a speed of 0.6 c with respect to the spaceship (see Figure 6.6). What is the speed V of the missile with respect to the earth?

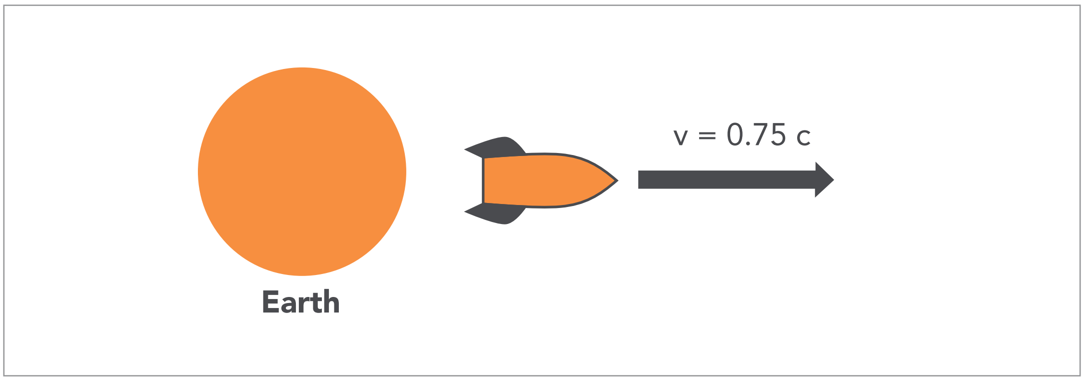


Figure 6.5

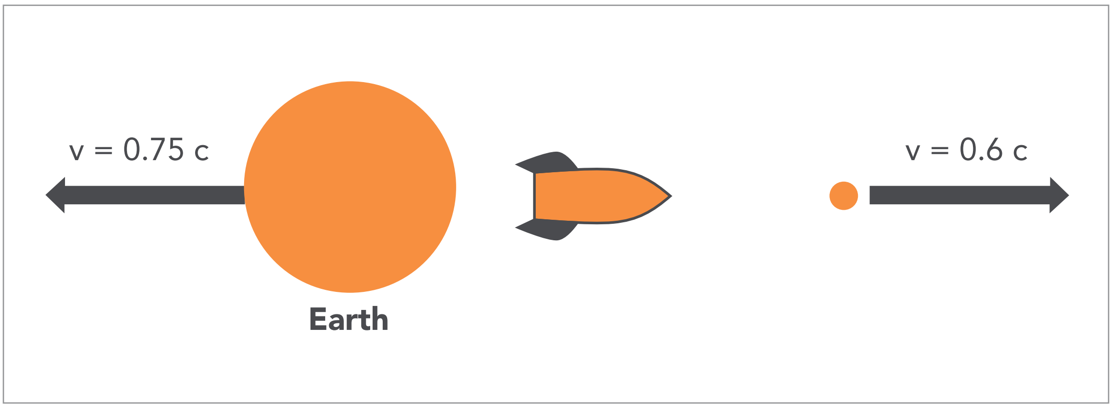


Figure 6.6

Solution

From the velocity-addition formula (because both speeds are in the same direction), we obtain

$$V = (u + v)/ (1 + uv/c^2) = (0.6\ c + 0.75\ c)/ (1 + (0.6\ c)(0.75\ c)/ c^2)$$
$$V = (1.35\ c)/(1 + 0.45) = 1.35\ c/1.45 = 0.93\ c$$

According to the special theory of relativity, the sum of any two velocities less than the speed of light (no matter how close they may be to the speed of light) always is a velocity less than the speed of light.

Units and Calculations

In practice, we seldom encounter relativistic speeds (speeds that are significant fractions of the speed of light) except when dealing with atomic and nuclear phenomena. The factor $\sqrt{1-\left(\frac{v^2}{c^2}\right)}$ is appreciably different from one only for speeds of about 0.5 c or greater. Objects of ordinary size never acquire speeds sufficiently great to require the use of the relativistic laws. For example, an artificial satellite circling the earth may move at 18,000 miles per hour with respect to the earth. This seems like a large speed in terms of our normal experiences, but in this case v = 0.000027 c and the classical mechanics can be applied for any kind of calculation.

The metric (SI) system of units described in Chapter l was developed for use in dealing with objects of the size normally encountered in everyday life or in the laboratory. These units are inconveniently large for calculations involving the small-scale phenomena of atomic and nuclear physics. Therefore, we now define a set of smaller, more appropriate units. If we use these units carefully, they will simplify our calculations. The unit that we will use for length in atomic and nuclear physics is the Angstrom (abbreviated Å). One angstrom is defined as 10^{-10} m.

Although one second is a very long time interval compared to the intervals involved in most atomic and nuclear interactions, the second is commonly used as the unit of time in describing such interactions. The unit that we will use for mass in atomic and nuclear physics is the atomic mass unit (abbreviated u). One atomic mass unit is defined as one-twelfth the mass of a neutral atom of carbon-12. In terms of our large-scale unit of mass,

$$1\ u = 1.66 \times 10^{-27}\ kg$$

The masses of the proton and neutron are approximately 1 u and the mass of an electron is 0.00055 u.

Although.one coulomb is a very large amount of charge on the atomic or the nuclear scale, charges in such calculations typically are expressed in terms of the coulomb.

The unit we use for energy in atomic physics is the electron-volt (abbreviated eV).

$$1 \text{ eV} = 1.6 \times 10^{-19} \text{ J}$$

For example, when one atom of carbon combines with one atom of oxygen to form a molecule of carbon monoxide, 11 eV of energy is released. (Where does this energy come from?) The electron-volt is a convenient unit of energy for use in describing phenomena at the atomic level.

When we deal with nuclear phenomena, we find that the energies involved are much greater. A more convenient unit for energy is the megaelectron-volt (abbreviated MeV), which is defined as 10^6 eV (the prefix mega means one million):

$$1 \text{ MeV} = 10^6 \text{ eV} = 1.6 \times 10^{-13} \text{ J}$$

When a neutron combines with a proton to form the nucleus of a heavy hydrogen atom, 2.2 MeV of energy is released. When a uranium nucleus splits into two smaller nuclei in a fission reaction, approximately 200 MeV of energy is released.

When measuring energy in eV or MeV, the unit m^2/s^2 is not a convenient unit for c^2.

$$c^2 = (9 \times 10^{16} \text{ m}^2/\text{s}^2) \times (1 \text{ kg}/1 \text{ kg}) = (9 \times 10^{16} \text{ kg m}^2/\text{s}^2) \;/\; (1 \text{ kg}) = 9 \times 10^{16} \text{ J/kg}$$

$$c^2 = (9 \times 10^{16} \text{ J/kg}) \times (1 \text{ MeV} / 1.6 \times 10^{-13} \text{ J}) \times (1.66 \times 10^{-27} \text{ kg} / 1 \text{ u}) = 931 \text{ MeV/u},$$

Thus, we see that 1 u of mass is equivalent to 931 MeV of energy.

Summary

The two postulates of the special theory of relativity lead to the prediction that an object in motion relative to an observer will have a shorter measured length in the direction of the motion than it will have for that same dimension when the object is at rest relative to the observer. Similarly, the measured distance between two objects in space depends upon the motion of the observer relative to the objects. An observer in motion with respect to the objects will determine their separation to be smaller than the separation measured by an observer at rest relative to the objects.

This phenomenon is called length contraction. The quantitative relationship between the two distance intervals is $L = L_o \sqrt{1-(v^2/c^2)}$ where L is the length in the direction of motion as determined by an observer moving at a speed v relative to the object (or the distance between two objects), and L_o is the corresponding length (or distance) measured by an observer at rest relative to the length (or distance) being measured. This prediction has been verified experimentally.

One of the important consequences of the special theory of relativity is that space and time, classically regarded as separate and independent concepts, are now seen to be closely interdependent. Space and time are united into a single concept called space-time.

The special theory of relativity also indicates that the classical treatment of the addition of two velocities must be incorrect for velocities that are significant fractions of the speed of light. With the relativistic equation for velocity addition, any two speeds less than c (no matter how close they may be to c) must add to a speed less than c. This is consistent with the prediction that no material object can travel at a speed equal to or greater than c relative to an observer.

There is a set of units when dealing with atomic or nuclear physics that are more convenient than SI units. For distance the Angstrom, Å, for mass the atomic mass unit, u, for energy the electron-volt, eV, and the megaelectron-volt, MeV, and for c^2, 931 MeV/u.

Important concepts

Length contraction; space-time; twin paradox; velocity addition.

Questions

1. An observer at rest with respect to a spaceship measures its length as 100 m. Another observer in a different frame of reference measures the length of the spaceship as 75 m. What is the relative speed of the two reference frames?
2. A billboard standing parallel to a highway is in the form of a square 3 m on a side. A traveler passes the billboard at a speed of 0.96 c with respect to the ground. In the traveler's frame of reference, what are the dimensions of the billboard?
3. A spaceship 30 m long is moving away from the earth at a speed of 0.90 c. What is the length of the spaceship as measured by an observer on earth?
4. An astronaut on a rocket ship is traveling toward a star that is 1.4 Ly distant from the earth in the rocket ship frame. The star is moving toward the rocket ship at a speed v. An observer on the earth agrees that the ship is traveling with speed v relative to the earth, but the earth observer says that the star is 10 Ly distant from the earth. What is the speed v?
5. The factor $\sqrt{1-(v^2/c^2)}$ appears frequently in the laws of relativistic physics. What happens to the value of this factor as v becomes a very small fraction of the speed of light? What happens to the value of the factor as v approaches the speed of light? What is the value of this factor when v equals the speed of light? What happens to the value of this factor when v is greater than the speed of light?
6. A rocket ship is traveling away from the earth at a constant speed 0.5 c. When it is at a distance of 7.3×10^{10} m from the earth (as measured by earth observers), it emits a light flash. According to earth clocks, how long will it take for the light to reach the earth? According to the rocket ship clocks, how long will it take for the light to reach the earth?
7. Alpha Centauri, the nearest star to our solar system, is a distance of 4.22 Ly from the earth. If a spaceship travels to Alpha Centauri with a constant speed of 0.8 c relative to the earth, how much will the astronauts age during the trip to the star?

8. A and B are twins. A stays home, while B leaves immediately after birth to a star located 10 Ly from earth in the earth frame of reference. The spaceship travels at a speed 0.8 c relative to the earth. Upon reaching the star, B reverses direction and immediately returns to the earth at the same relative speed. What is the distance from earth to the star as measured by B during the trip? According to earth clocks, how long does it take B to reach the star? According to spaceship clocks, how long does it take B to reach the star? If A and B are reunited after the trip, what are their respective ages?
9. A science-fiction story is set in the future near the center of our galaxy, about 26,000 Ly from the earth. A character in the story is a retired spaceship captain, 70 years old. She was born on the earth. Is such a situation possible according to the laws of relativistic physics? Explain.
10. According to astronomical observations made from the earth, the star Arcturus is about 40 Ly from the earth. A spaceship travels at a constant speed of 0.99 c relative to the earth. How long does the trip take according to earth clocks? How long does the trip take according to spaceship clocks?
11. An isolated neutron at rest has an average lifetime of about 1000 s before it disintegrates. Suppose that a neutron is created on the sun and travels away from the sun at a speed of 0.995 c, would such a neutron be able to reach the planet

Saturn, at a distance of about 1.5 x 10^{12} m from the sun (as measured by an observer at rest with respect to the sun)

12. A spaceship exploring the universe is damaged in an accident. The captain discovers the damaged life-support system can keep the crew alive for only another 6 hours (21,600 s). According to the star maps, the nearest base where supplies are available is 8 x 10^{12} m from the ship. If the crew is to survive, what is the minimum speed at which the ship must travel toward the base?
13. A particle called the K+ meson has an average lifetime of about 10^{-8} s before disintegration, as measured when the particle is at rest. How fast must such a particle be moving through a piece of experimental equipment if it is to travel the l5 m from one end to the other before disintegrating?
14. State the relativistic equation for the addition of velocities. Identify all terms in the equation and explain the conditions under which this equation is valid. Compare the equation with the Newtonian equation for velocity addition and explain the conditions under which the Newtonian equation is valid.
15. A fast train moves at a speed of 0.8 c relative to the earth. Inside the train a fast runner moves toward the front of the train at a speed of 0.7 c relative to the train. What is the speed of the runner relative to the earth?
16. Two spaceships leave the earth traveling in opposite directions with speeds of 0.9 c relative to the earth. An observer on one spaceship measures the speed of the other spaceship relative to the observer's ship. Is the measured speed equal to 1.8 c? If not, what is the measured speed?
17. A spaceship moves away from the earth at a constant speed of 0.9 c. A rocket is launched from the spaceship in the same direction with a speed of 0.8 c relative to the spaceship. What is the speed of this rocket relative to the earth?

The following questions are of a more general nature. They have no single correct answer and is just something for you to think about. When possible, questions like this are best answered in conversation with others.

18. Is it possible for a person to be older than his or her parent? Explain.
19. Because the speed of light is so great, many early physicists assumed it was infinite. Discuss the implications of time dilation, length contraction and velocity addition if this assumption was correct.
20. Discuss the validity of the following statement. "Einstein's special theory of relativity proves Newton's theory of classical mechanics is wrong."
21. Discuss the following statement. "The relativistic phenomena of length contraction, time dilation, and velocity addition represent actual differences in length, time, and velocity; they are not simply apparent changes." This is an important question. Think carefully about the exact meaning of a length, a time interval, or a velocity.

7 THE GENERAL THEORY OF RELATIVITY

In the preceding two chapters we have discussed the special theory of relativity that Einstein proposed in 1905. The qualifier "special" refers the fact that the theory is restricted in its application; it applies only to inertial frames of reference. Recall that an inertial frame of reference is one in which physical laws have their simplest physical and mathematical forms. Once any inertial frame of reference is found, all other frames of reference moving in a straight line at constant speed with respect to the inertial frame of reference are also inertial frames of reference. The special theory of relativity creates a special class of observers (those in inertial frames of reference) for whom the laws of the theory are valid. All observers whose frames of reference are non-inertial (changing their state of motion with respect to an inertial frame of reference) are excluded. For such observers, the physical laws do not have the simple forms given in the special theory of relativity.

This situation is an improvement over the situation with the classical theories, where the laws of electromagnetism were thought to have their simplest forms only for an observer in the single frame of reference that is at rest with respect to the ether. However, Einstein was not satisfied with this improvement. He set out almost immediately to find a more general theory of relativity in which the laws would be valid for any observer in any frame of reference. The task was formidable, and it required a degree of mathematical sophistication that had never before been used in physics. Late in 1915, Einstein published his general theory of relativity, in which the principle of relativity, in its most general form, is assumed as a postulate.

> **The laws of physics are the same (that is, they have the same mathematical forms) in all frames of reference.**

You will note that the restriction to inertial frames of reference used in the special theory has now been removed.

Recall that Einstein was led to develop the special theory of relativity because of his philosophical and esthetic conviction that the principle of relativity should apply to the laws of electromagnetism as well as to the laws of mechanics. To achieve his goal of a physical theory applicable to all inertial frames of reference, he assumed that the speed of light is the same for any inertial observer. A rigorous exploration of the consequences of the two postulates led to redefinition of the basic concepts of length, time, mass, and energy. Although most physicists were at first disturbed by the surprising implications of the special theory, they were soon led to consider it very seriously because it seemed to be consistent with some puzzling experimental results, such as those of the Michelson-Morley experiment, that could not be explained by the classical theories. As physicists began to experiment with subatomic particles, they began to measure the properties of particles traveling at speeds that are significant fractions of the speed of light. In every case, the results were quantitatively consistent with the predictions special theory of relativity.

Einstein was confident that his physical intuition would again be justified when the implications of the new theory were explored. Other physicists were less willing to plunge into immediate acceptance of the new theory. For one thing, few physicists possessed the mathematical sophistication needed for a full understanding of the general theory. Although the laws of the special theory were a bit more complicated than the classical laws, they required no new mathematical techniques.

7.1 THE PRINCIPLE OF EQUIVALENCE

The concept of mass is familiar. In most people's mind it is associated with weight. More massive objects weigh more. This property of mass is called gravitational mass because it determines the force that gravity will exert on the object. There is another, seemingly unrelated, property of mass. The more massive an object is, the harder is it to change either the direction or the speed of its motion: that is, the more difficult it is to accelerate the object. This property of mass is called inertial mass. In classical physics, inertial effects are not related by theory to gravitational effects. In developing his theories of motion and gravity, Newton assumed that they were the same property, although there was no theoretical justification for this assumption. It just seemed to work. By the end of the 19^{th} century it had been empirically shown that, to a high degree of accuracy, inertial and gravitational mass can be considered numerically equal, although there was still no theoretical connection between the two.

Having two completely independent properties that are accidentally equal to each other is, philosophically, a very unsatisfactory situation. Einstein reasoned that there must be some underlying physical significance to this equality which would constitute a single interpretation for mass.

As with the special theory, Einstein utilized thought experiments to guide his thinking. Imagine a person in a small, windowless compartment with two objects of different mass. Suppose this compartment is located on the surface of the earth. If the two objects are dropped simultaneously from the same height, they will hit the floor at the same time. Galileo was the first to show this and to measure the acceleration due to gravity as 9.8 m/sec^2. That is, near the surface of the earth a free-falling object will change its speed at a rate of 9.8 m/sec each second.

Suppose now that the compartment is in space far removed from all massive objects. There will be no gravitational forces acting on the compartment and its contents. If there are no forces acting on the compartment, the person and the two objects will float around in the compartment much as you see the astronauts doing when in orbit around the earth. This is often referred to as weightlessness, but that is not strictly correct. The astronauts are not weightless; the earth is still pulling them and the space craft downward. That weight is what is keeping them in orbit rather than from flying off into space. However, with the rockets turned off, the space craft is free falling around the earth and this is, as we will see, equivalent to being weightless.

Now imagine there are rockets attached to the bottom of the compartment causing it to accelerate, just as the astronauts are accelerated on takeoff. The inertial mass of the person, with its tendency to maintain its state of motion, will resist this change in velocity. The person will be pressed against the floor of the compartment just as the astronauts are pressed against their seats on takeoff. With effort, the person will be able to stand. If the two objects are dropped as before, their inertial mass will tend to keep them in a constant state of motion and the floor of the compartment will accelerate up and overtake them at an ever-increasing speed.

In the frame of reference of the observer, the objects will 'fall' to the floor. In fact, if the acceleration of the compartment is 9.8 m/sec per second, they will 'fall' to the floor exactly as they would if the compartment was at rest on the surface of the earth. As Einstein thought about this, he realized that there is no possible measurement the person in the compartment could make that would distinguish between the two situations. In 1911, based only on his thought experiments, Einstein made the following bold statement. "It is impossible to distinguish, by any experiment whatsoever, between the effects of acceleration and the effects of gravity." This statement is known as the principle of equivalence.

Note that the principle of equivalence only applies in a small region of space. Otherwise non-uniformities in the gravitational field can be distinguished from the uniform effects produced by acceleration. The principle of equivalence should not be confused with the equivalence of mass and energy, which is an entirely different matter.

The thought experiment above is a mechanical one. What about experiments involving electromagnetic phenomena? Again, as was the case in 1905, Einstein turned to the propagation of light. Suppose that there is a source of light in an inertial frame of reference and that a beam of light from the source enters the compartment horizontally from the side. If the compartment is at rest with respect to the light source, it is also an inertial frame. The light beam will enter the compartment, travel horizontally across the compartment, and strike the wall the same distance above the floor as the place it entered. By considering this situation in the frame of reference of the accelerating compartment in outer space (a non-inertial frame of reference) Einstein was able to make an important prediction.

In an inertial frame the speed of light is an absolute constant. Thus, the beam of light in its inertial frame travels at a constant speed as it passes through the compartment. However, the vertical speed of the compartment is constantly increasing. Thus, in fixed time intervals the beam of light will always travel the same horizontal distance while the compartment will travel increasingly greater vertical distances.

In this case, the path of the beam of light is not a straight line in the frame of reference of the compartment. It will be bent toward the floor. If an accelerating frame of reference is in no way distinguishable from a frame of reference in a gravitational field, this thought experiments predicts that a beam of light will not be a straight line in a gravitational field.

Einstein concluded that the path of light is altered by gravity such that a light beam passing near a massive object ought to travel as if attracted toward the object.

The Principle of Equivalence

A beam of light is bent in exactly the same way in a frame of reference that is accelerating and in one in a gravitational field. The inability to distinguish between the two experimentally is known as the principle of equivalence.

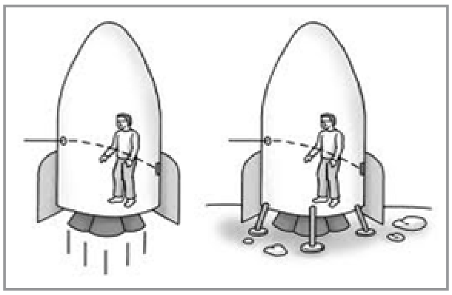


Figure 7.1 The Principle of Equivalence

Calculations showed that the amount of bending produced even by an object as massive as the sun is extremely small. However, in a 1911 paper, Einstein proposed an experiment that should in principle allow a measurement to be made. For reasons discussed later, this experiment was, fortunately, not carried out until 1919.

7.2 THE GENERAL THEORY OF RELATIVITY

Einstein's development of the General Theory of Relativity was strikingly similar to his development of the Special Theory of Relativity. In the special theory, Einstein took the law of propagation of light and the principle of relativity (for inertial frames of reference), which appeared to be incompatible and showed that they were both true. In the case of the general theory, he took the principle of equivalence and the general principle of relativity (that the laws of physics are the same in all frames of reference, both inertial and non-inertial) which appeared to be incompatible and showed that by assuming both were true, a coherent theory could be produced.

In the previous section it was shown that the principle of equivalence led to the fact that while light in an inertial frame of reference travels in a straight line, light in an accelerating frame (or in a gravitational field) travels in a curved path. It appears that the laws governing the motion of light in an inertial and in a non-inertial frame are different, violating the general principle of relativity. That Einstein was able to resolve this paradox with the General Theory of Relativity, in spite of unimaginable mathematical and conceptual difficulties, stands as a monument to the human intellect. As was the case with the special theory, the solution lies in our concept of space-time.

The main problem Einstein was having with developing his new theory was mathematical. Although by ordinary standards, Einstein might be considered a mathematical genius, by the standards of theoretical physicists, Einstein was not exceptional. It was his unprecedented physical intuition (with the possible exception of Isaac Newton) rather that his mathematical ability that made him the greatest physicist of his time. It was only in the process of work on his theory that Einstein gradually acquired the mathematical techniques with which to express the theory. In late 1912, he wrote to a friend,

> I occupy myself exclusively with the problem of gravitation and now believe that I will overcome all difficulties with the help of a friendly mathematician here. But one thing is certain: that in all my life I have never before labored at all as hard, and that I have become imbued with a great respect for mathematics, the subtle parts of which, in my innocence, I had till now regarded as pure luxury. Compared with this problem, the original theory of relativity is child's play.

The 'here' was his alma mater, Zurich Polytechnic Institute, where he had just returned as professor of physics, and the 'friendly mathematician' was Marcel Gossmann, an old school friend. The necessary mathematical technique was tensor calculus, Grossmann's specialty.

In addition to the principle of equivalence and the general principle of relativity, Einstein put an additional constraint on the theory, one dictated by aesthetic values which he held to be of paramount importance in physics. Out of the literally thousands of formalisms provided by the tensor calculus consistent with the principle of equivalence and the general principle of relativity, Einstein insisted that only the mathematically simplest formalism would provide a correct description of nature.

In 1914, Einstein left Zurich for Berlin. There, late in 1915, after years of almost constant effort, Einstein arrived at a formalism that seemed to satisfy all his requirements. As a test, Einstein used the theory to calculate the orbit of Mercury. The classical Newtonian theory of gravity was very, very slightly, though undeniably, inconsistent with the observed orbit. The difference was an incredibly small 43 seconds of arc per century in the precession rate of the orbit.

Precession of Mercury's Orbit

Observed precession (in 100 yr)	= 574.2 ± 0.4″
Precession predicted by Newtonian theory	= 531.1 ± 0.2″
Observed discrepancy	= 43.1 ± 0.5″
Pridicted relativistic correction	= 43.03″

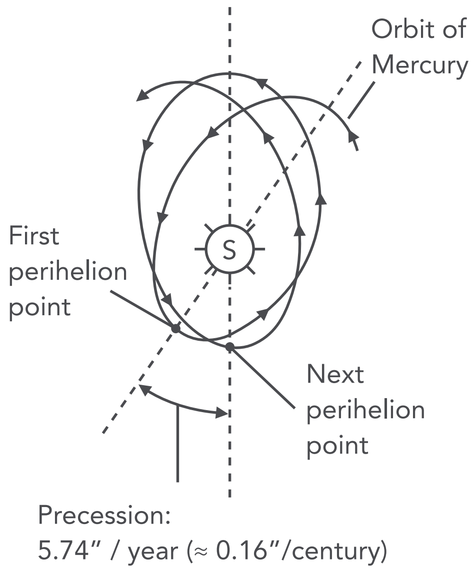


Figure 7.2 Precession of Mercury's Orbit

The observed and the calculated orbit matched perfectly; the additional 43 seconds of arc per century came naturally and of necessity from the theory. It was immediately clear that Einstein's theory is a more accurate description of gravity than Newton's, a theory that for centuries had been assumed to be absolutely correct.

The General Theory of Relativity is, as Einstein had intended from the beginning, a theory of gravity. However, it did not simply produce a new, more general force law. It changed in a very fundamental way our concept of gravity. According to the theory, the effects of gravity are not the result of a force being exerted on an object, but rather are the result of the natural, inertial motion of the object through space-time, the properties of which are determined by the presence of other massive objects.

In the special theory, space and time are interwoven, and separate models of each cannot be constructed. In the general theory, space-time and matter lose their independent meaning. They are different aspects of a single unity; each is meaningless in the absence of the other.

In addition to providing a more accurate description of the effects of gravity, the general theory also eliminated one of the most disturbing aspects of the classical theory – action-at-a-distance. The sun does not pull on the earth. Rather it changes the properties of the space-time in its vicinity. In effect, it curves space-time in a way that inertial motion through it is no longer a straight line.

This is much like an object moving on a curved two-dimensional surface. Imagine a horizontal, flat rubber sheet. A bowling ball placed on it will depress the sheet creating a curved surface surrounding the ball. If a marble is rolled across the rubber surface, it will not travel in a straight line. It will be deflected by the curved surface with the amount of the deflection being greatest where the curvature is greatest. In fact, the speed of the marble can be such that the marble will orbit the bowling ball, much as the earth orbits the sun. The bowling ball is not exerting a force on the marble to produce the orbit. Rather it is the curvature of the surface that is producing the orbit.

The amount of the curvature of the rubber sheet depends on the distance from the ball. In this analogy, the amount of the curvature of the rubber sheet models the strength of the gravitational field. Just as is the case with the strength of the gravitational field, the amount of curvature of the rubber sheet surrounding the bowling ball decreases with distance.

This analogy is not exact, as it explains the motion in terms of a curved two-dimensional surface whereas gravity is a consequence of curved four-dimensional space-time. However, it does provide a visualizable way to represent the nature of gravity in the General Theory of Relativity.

Curvature of Space - Time

In the general theory, gravity is not a force acting at a distance. Rather its effect on motion is the result of a massive object, such as the earth in this example, curving the space-time in its vicinity. The moon orbits the earth as a result of its inertial motion through the curved space-time, not because the earth is pulling on it.

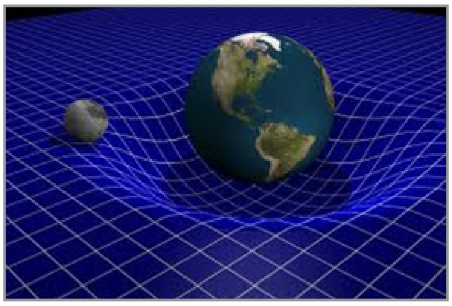


Figure 7.3 Curvature of Space-Time

7.3 EVIDENCE FOR THE GENERAL THEORY OF RELATIVITY

The calculation of the correct motion for the planet Mercury was a tremendous success for the general theory. However, physical theories are judged primarily on predictions of new, unsuspected physical phenomenon that are subject to experimental or observational verification. The Mercury result, though completely unforced, was not a prediction but an explanation of a previously known fact.

Deflection of Light by a Massive Body

The deflection causes the star to appear in a different point in the sky.
The general theory predicted the shift in position that would result.

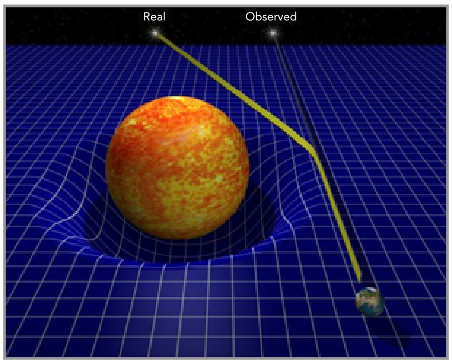


Figure 7.4 Deflection of Light

The first crucial test of Einstein's theory was made during the total eclipse of the sun in May 1919. In 1911 Einstein had proposed an experiment to measure the amount of bending of light as it passed near the sun. During a total eclipse of the sun, the sky is dark as night, and the stars can be seen. A photograph of the stars taken during a total eclipse can be compared to an earlier photograph of the same region of sky taken at night. Light rays passing near the eclipsed sun are bent. For these stars, their relative positions with respect to the other stars will be slightly different on the two photographs.

For the star in the figure, the photograph taken during the eclipse would show a different location in the sky compared to a photograph taken at some other time. The difference between the apparent and the true positions of the star is a direct measure of the amount of bending produced by the gravitational effect of the sun.

Einstein predicted the bending of light in a gravitational field in 1911, well before his theory was complete. He made this prediction based on the principle of equivalence and Newton's theory of gravity. The calculation indicated that a light beam passing very near the surface of the sun would be deflected through an angle of 0.87 seconds of arc or 0.00024 degrees.

Although this angle is extremely small, there was a possibility that it could be measured. In 1914, the German astronomer Erwin Finlay-Freundlich, set off for Russia to observe the total eclipse of the sun and try to verify Einstein's prediction. However, he was prevented from doing so by the outbreak of the First World War. Einstein wrote the following in a letter to a friend:

> Europe, in her insanity, has started something unbelievable. In such times one realizes to what a sad species of animal one belongs. I quietly pursue my peaceful studies and contemplations and feel only pity and disgust. My dear astronomer Freundlich will become a prisoner of war in Russia instead of being able there to observe the eclipse of the sun. I am worried about him.

In fact, it was fortunate that Freundlich was not able to make his measurements. By the end of 1915, it was clear that his 1911 prediction was in error. In 1911 he had assumed that Newton's gravitational theory was appropriate for the calculation, but his general theory indicated that this was not the case. A new calculation based on the general theory produced a value of 1.75 seconds of arc, about twice the earlier prediction.

There was no chance to test the eclipse prediction until the war ended. However, as early as 1917, Sir Arthur Eddington (1882 – 1944), a British physicist, began preparing for two 1919 expeditions. Eddington was one of the first to appreciate the significance of the general theory, but his efforts to organize the test expeditions were motivated by more than just scientific curiosity. Eddington was a Quaker, and like Einstein, was profoundly disturbed by the war. He felt that if a British expedition verified the work of a German theoretical physicist, this would help heal the wounds of war, and, in particular, would reestablish scientific relations between the warring nations.

The results confirmed Einstein's prediction. 1919 was the last year of Einstein's private life. The announcement of the verification of his theory to a war-weary and heartsick people made him a world-wide hero. He became the personification of intelligence, an identification that has survived in spite of the fact that Einstein has been dead for more than a half-century. This fame, which he neither sought nor enjoyed, would later cause him to refer to his years of obscurity in the Bern Patent Office as the happiest of his life.

In early 2016, perhaps the most spectacular experiment confirming a prediction of the general theory --: the first-ever direct detection of gravitational waves, ripples in the fabric of space-time. These waves are produced when two massive accelerating objects collide with each other. In this case, the detected waves were produced by a pair of stellar-mass black holes in a death-spiral into one another.

7.4 THE GENERAL THEORY OF RELATIVITY TODAY

In the decades following its publication, the number of observable effects that distinguished Einstein's General Theory of Relativity from the much simpler Newtonian theory was small, and the magnitude of the differences between the predictions of the two theories was almost negligible. For these and other reasons, interest in the general theory soon all but disappeared – almost, but not quite. During the twenties and thirties, a handful of theoretical physicists and cosmologists were applying the general theory and arriving at results which staggered the imagination, results so inconceivable and incomprehensible that they were generally ignored.

This attitude toward the General Theory of Relativity changed rather dramatically in the 1960s due primarily to developments in the field of astronomy. The discoveries of such exotic objects as quasars and pulsars suddenly made the bizarre predictions of Einstein's theory seem less unreasonable. Today the General Theory of Relativity is again at the forefront of physics. The two most important applications of the theory are the physics of gravitationally collapsed objects, black holes being the most extreme example, and cosmology, the science of the universe as a whole.

A black hole is a region of space-time where gravity is so strong that anything, even light, that enters the region will be trapped there. The possibility of the existence of a black hole was recognized as a direct prediction of the general theory almost immediately after the formulation of the theory. However, it was not until 1939 that Robert Oppenheimer and one of his students, Hartland Snyder, suggested a possible mechanism whereby one might actually form.

When all possible fuels of a star are exhausted, it will begin to contract gravitationally under its own weight. Using the general theory, Oppenheimer and Snyder were able to show that if the mass of the collapsing star exceeded a certain value, today believed to be about three times the mass of our sun, no known force could prevent complete collapse. The star will collapse to smaller and smaller volumes, collapse to the point where electrons, protons, and neutrons are crushed out of existence, crushed to an object of infinite density surrounded by a volume of space where gravity will prevent the emission of light. It is little wonder that at first physicists were reluctant to accept this. Arthur Eddington characterized the idea as absurd.

Today the existence of black holes is a virtual certainty. Black holes from 3 to 10 or so solar masses have been identified in orbit around ordinary stars. A supermassive black hole containing 2.2 million solar masses has been found in the center of our galaxy. Other supermassive black holes have been detected in the centers of other galaxies.

Modern cosmology, one of the greatest intellectual adventures ever undertaken, owes its existence to Einstein's General Theory of Relativity. In 1916, soon after the calculation of Mercury's orbit, Einstein applied his theory to the universe as a whole and got an unwelcome result. The theory indicated that the universe could not be static. The theory clearly predicted that the amount of space in the universe must be either increasing or decreasing. Einstein did not believe that this was true, and the other physicists and astronomers he consulted assured him that it was not. Everyone was sure that the universe was static and unchanging on the large scale.

For once in his life, Einstein lost faith in the basic simplicity of nature, and added an ad hoc term, which he called the cosmological constant, to his theory. Though it marred the beauty and simplicity of the theory, Einstein believed it would eliminate the offending prediction of expanding or contracting space. The effect of the cosmological constant is to provide a universal repulsive force that Einstein thought could balance the attractive force of gravity and allow for a static universe.

In the 1920s two theoretical physicists independently showed that Einstein's 'fix' of his theory did not work. Not only did the original theory require that space be expanding or contracting, but the new version with the cosmological constant made exactly the same prediction. Hardly anyone at the time studied the general theory and no one paid much attention to these papers, including Einstein himself. One of the involved physicists, the Belgium priest, George Lemaitre, was so discouraged by the lack of interest in his paper that he switched his research to another field.

In 1929, the American astronomer, Edwin Hubble (1889 – 1953), published a law stating that the radiation received from distant galaxies is stretched out, that is, has longer wavelengths than radiation received from local sources. Furthermore, the amount of the stretching, called redshift by astronomers is directly proportional to the distance to the galaxy. Hubble knew that this was a very significant result but had no idea what caused it.

When Arthur Eddington heard of this result, he immediately knew the explanation. Eddington was among the few physicists who actually understood the details of the general theory. He was also one of the few who were aware of the papers showing that, regardless of the version of the theory used, the general theory required either an expanding or a contracting universe. Either was consistent with the general theory. As to which described our actual universe, it was an empirical question. Eddington immediately recognized Hubble's Law as empirical evidence that space is expanding. When Einstein heard of this, he called his inclusion of the cosmological constant the "greatest blunder of my scientific career." We will later see that this may not be the case.

Eddington concluded space was expanding because expanding space would stretch out the wavelengths of light from distant galaxies just has Hubble had observed. Further, the light from the more distant galaxies would spend more time traveling through expanding space and thus be stretched more, making the amount of redshift directly proportional to the distance. Had space been contracting, the wavelengths would be compressed, or blueshifted rather than redshifted.

When George Lemaitre learned of these developments, he returned his interest to cosmology. He was the first to think scientifically about what the universe must have been like in the past if space is expanding. He concluded that if time could be run backwards, the universe would become increasingly dense. Using the general theory, it was clear that this increasing density would become infinite at some finite time in the past, that is; that the universe must have had a beginning in time.

The prevailing belief among scientists at the time was that the universe was infinitely old. Thus the prediction by a Catholic priest that the universe had a finite age, (he had probably believed this from childhood) was not taken very seriously. Lemaitre's theory had an additional strike against it. Almost nothing was known about nuclear physics at the time, so Lemaitre was unable to use his theory to make testable predictions regarding the present universe.

The idea was taken up again in the late 1940s and early 1950s by the Russian-American physicist, George Gamow. By that time considerable progress had been made in the area of nuclear physics and Gamow was able to make two significant, testable predictions. The first of these was that the very early universe could not have produced any appreciable amount of chemical elements more massive than helium. Thus, the early universe must have nearly pure hydrogen and helium. The second was that the entire universe would be filled with thermal radiation at a temperature of a few degrees above absolute zero; the cosmic background radiation. Gamow's theory soon became known as the Big Bang theory.

By the late 1950s, it had become clear that the elements heavier than helium are actually created in the interiors of massive stars, and that the explosive deaths of these stars distributed the heavier elements throughout the galaxy, making them available to later generations of stars, including our sun. The evidence was also accumulating that in the past, the galaxy was more nearly pure hydrogen and helium.

The clinching piece of evidence in favor of the Big Bang was the discovery in 1964 of the predicted cosmic background radiation. Arno Penzias and Robert Wilson, two radio astronomers working for Bell Laboratories, detected the radiation using a radio telescope built for trans-Atlantic telephone conversations. The pendulum of scientific opinion immediately swung from the Steady State theory, with an infinitely old universe, to the Big Bang and a finite age for the universe. The latest studies of this cosmic background radiation indicate that the universe began 13.7 billion years ago.

It is now universally accepted that space is expanding. It was also universally accepted that gravity, as understood by the General Theory of Relativity, required that the rate of expansion must be decreasing. In the mid-1990s, two independent research teams set out to measure the rate of slow down. By 1998, their results were in. Much to their (and everyone else's) amazement, both determined that the rate of expansion was actually increasing. The only possible explanation for this is that, in addition to the attraction of gravity, the universe contains a global repulsive force, and, at the present time, the effect of this repulsive force is greater than that of gravity. The nature of this force is unknown and it is just referred to as dark energy, but Einstein's cosmological constant of 1916 seems to be in accord with what has been observed so far of the changing expansion rate of the universe. Perhaps its inclusion was not a 'blunder' after all.

Albert Einstein: Person of the Century

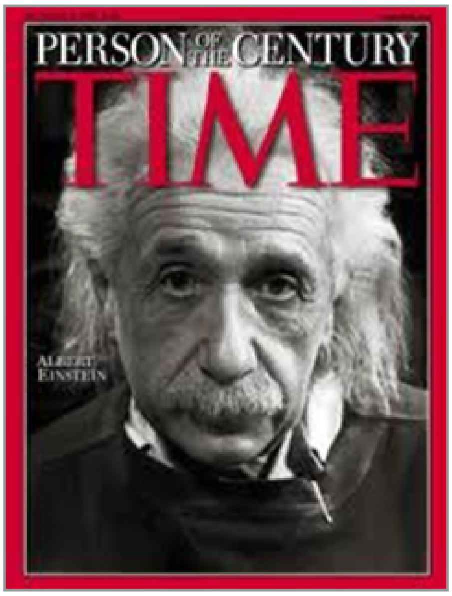


Figure 7.5 Person of the Century

Summary

In late 1915, Albert Einstein formulated the general theory of relativity. This theory is based on the following two postulates: 1. The laws of physics are the same in any frame of reference (the general principle of relativity); and 2. It is impossible to distinguish by any experiment between the effects of acceleration and the effects of gravity (the principle of equivalence). These postulates lead to the logical conclusion that any object obeys simple laws of motion through space-time, but that space-time itself is altered by the presence of masses; hence gravitation is viewed as a property of space-time This theory represents a tremendous advance in our understanding of the concept of space-time (an understanding that began with the special theory of relativity).

The general theory of relativity is best described as a relativistic theory of gravitation. It extends our understanding of gravity into regions where the strength of gravity is too great for the Newtonian theory to be valid. As required by the principle of correspondence, the general theory of relativity yields the same predictions as the Newtonian theory for those situations in which the Newtonian theory had already been shown to be valid.

Like all good theories, the general theory of relativity led to several predictions of previously unknown natural phenomena. The earliest experimental confirmations of predictions from the theory were the observations of the bending of light rays when passing near a very massive object. All observations to date concerning the nature of space-time and gravity are consistent with the general theory of relativity including the latest observation of gravity waves. The most interesting and exciting applications of the general theory have been in the areas of astrophysics and cosmology.

Important concepts

General theory of relativity; general principle of relativity; inertial mass; gravitational mass; principle of equivalence; gravity as a property of space-time; curved space-time; black hole; cosmology; big-bang model

Questions

1. Explain how the general theory of relativity differs from the special theory of relativity. Describe the similarities between the two theories.
2. Distinguish between inertial mass and gravitational mass.
3. State the principle of equivalence and explain its meaning.
4. What role did the orbit of Mercury play in the development, testing, and acceptance of the general theory of relativity?
5. A mass is placed on the end of a suspended spring inside a closed compartment. A physicist inside this compartment observes that the spring stretches as the mass is hung on it. State two different explanations that the physicist might give for this observation in terms of the relationship between the compartment and the rest of the universe.

6. What esthetic constraint did Einstein adopt in developing the general theory of relativity?
7. Explain how a photograph taken during a total eclipse of the sun can be used to provide evidence about the possible bending of light rays when passing near a very large mass.
8. What is a black hole? How can a black hole be detected?
9. What does it mean to say that the universe is expanding? What evidence suggests that this is the case?
10. In what way might the statement "we are stardust" be quite literally valid?

The following questions are of a more general nature. They have no single correct answer and are just something for you to think about. When possible, they are best answered in conversation with others.

11. A koan is a riddle of paradoxical content used by Zen Buddhists to achieve a deeper understanding of reality. Meditating on the apparent logical contradictions involved in answering the riddle is said to force the mind beyond a logical barrier and into a new realm of understanding that transcends the limitations inherent in the form of the question and the usual approaches to its solution. Discuss the role of paradoxical puzzles in the development of the special and general theories of relativity.

12. Einstein's philosophical and physical intuition led him to hold fast to the assumption that all physical laws should have identical and relatively simple forms in all frames of reference. Until the time of Galileo, the philosophical and physical intuition of most natural philosophers led them to hold fast to the assumptions that the earth is at the center of the universe and that a heavy object falls faster than a lighter one. Discuss the role of philosophical and physical intuition in physics. Under what conditions does it advance or hinder the development of physics?
13. Einstein regarded the success of the general theory of relativity in explaining the motions of Mercury as a thrilling vindication of his new theory; other physicists were not terribly impressed. Other physicists (and the general public) were greatly excited by the results of the eclipse expeditions that verified the bending of light rays when passing near the sun; Einstein apparently attached little importance to this experimental evidence. Discuss the role of experimental evidence (especially when it involves a previously unknown phenomenon) verses the power to explain something that is already known, in confirming the validity of a theory.
14. Contrast gravity as it is understood in terms of the general theory of relativity with gravity as it is understood in terms of the classical Newtonian theory.
15. In 1961 the historian of science T. S. Kuhn wrote: "Unlike the special theory, general relativity is today very little studied by students of physics. Within fifty years we may conceivably have lost sight of this aspect of Einstein's contribution." Discuss the reasons this this statement was wrong only twenty years later.
16. The principle of equivalence applies only to small regions of space. Suppose that the closed compartment containing the observer dropping weights is hundreds of miles wide. Describe an experiment by which the observer could determine whether the rectangular compartment is near the earth's surface or is accelerating through empty space.

8 ELECTROMAGNETIC RADIATION: THE CLASSICAL THEORY

In Chapter 2 we described three categories of energy: kinetic energy, rest-mass energy, and radiant energy. We are now ready to consider radiant energy in more detail. For our purposes, we have defined radiant energy as the energy associated with electromagnetic radiation, which includes such phenomena as radio waves, infrared radiation, visible light, ultraviolet radiation, X-rays, and gamma rays. These phenomena are quite distinct from matter (which is made up of atoms or molecules) and can be regarded as energy that is in the process of being transferred between material objects. The nature of the process by which such energy is transferred is a topic that has played a fundamental role in physics with a long and interesting history.

When we think of mechanisms for transferring energy from one point in space to another, two possible models come to mind. Perhaps the most obvious one is the possibility that a particle carries the energy. Energy transferred to the particle becomes kinetic energy of the particle; the particle then travels through space and interacts with some object, transferring the energy to that object. For example, a gun transfers energy to a bullet, which can travel to some distant object and transfer energy to the target (perhaps knocking it over, or altering its shape in some way).

The second (less obvious) model is that of a wave. A wave is an oscillating disturbance in some medium. Although the individual points in the medium simply move back and forth over a limited distance, the wave pattern travels through the medium and carries energy with it. An example is a sound wave. Energy is used to set up a vibration in your vocal cords; this energy is transferred to the medium (air in this case) setting up a wave pattern that travels away from your mouth. When the sound waves reach someone's ear, they transfer energy to the eardrum and set up a vibration there, which is perceived as a sound. Note that the wave mechanism involves only the transfer of energy, whereas the particle mechanism involves the transfer of both energy and matter. There is no net transfer of air from your mouth to the listener's ear. Each vibrating air particle simply sets an adjoining air particle into vibration; energy is transferred across the room, but no individual air particle actually moves any significant distance across the room. The sound wave moves through the air with a particular speed that is determined by the properties of the air (about 343 m/s). A sound wave in water travels at a different speed (about 1500 m/s). It is characteristic of a wave that the speed of propagation is determined by the properties of the medium through which the wave travels.

Another example of a wave is the one that can form on the surface of a body of water. Suppose that you drop a rock into a lake. Energy is transferred from the rock to the water, causing water particles to oscillate up and down near the spot of impact. This disturbance sets up a similar oscillation of neighboring water particles, so that we see a wave pattern move outward as expanding circular waves centered on the point of impact. However, no individual water particle has significant sideways motion; only the wave pattern and associated energy movers along the surface of the lake. (You can verify this by putting a floating object on the surface and observing that it merely bobs up and down rather than being carried sideways by moving water.)

In this chapter and the next, we examine the following question: Is electromagnetic radiation best described as a wave propagating through some medium, or is it best described as a stream of particles? Throughout most of the period that this question has been debated visible light was the only known form of electromagnetic radiation. Therefore, we phrase the question here as it was phrased historically: Is visible light best described as a traveling wave pattern or as a stream of particles?

8.1 THE NATURE OF LIGHT

The problem of explaining the nature of light is an ancient one, but we will take up the story at the end of the seventeenth century. By this time, two distinct models had been proposed, and each had its group of supporters. On the one hand, Robert Hooke and Christian Huygens were the leading proponents of the wave model. They argued that light is an oscillating disturbance propagated through the ether (a material substance that permeates all space) in much the same way that sound is propagated through air. On the other hand, Isaac Newton was the most prominent supporter of the particle model. He believed that light consists of streams of tiny particles moving at high speeds. Neither of these models was developed to any great extent as a detailed quantitative explanation. Each model was able to provide qualitative explanations for some of the properties of light, but neither model could provide a comprehensive explanation of all of the known properties of light. Both Huygens and Newton were familiar with one important property of light: the fact that the path of a light beam is bent when it passes from air into a more dense substance such as water or glass. This property is called refraction (see Figure 8.1).

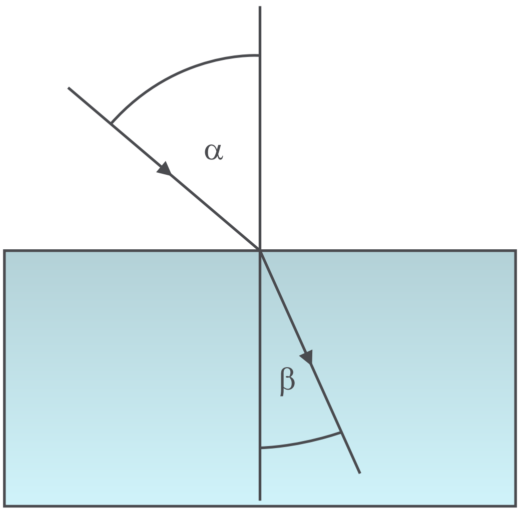


Figure 8.1 Refraction of light

Newton explained refraction in terms of the particle model by assuming that the light particles are attracted toward the more dense material as they approach it. (Do you see how this would produce the bending indicated in Figure 8.1?) Because this attraction would also accelerate the particles, Newton's explanation leads to the prediction that the speed of light is greater in water (or any medium denser than air) than it is in air.

A similar phenomenon of refraction can be observed with waves. For example, surface waves on water are refracted as they pass into more shallow water (where the speed of the waves is slower). In order to explain the observed refraction in terms of the wave model, Huygens was forced to conclude that the speed of light must be greater in the air than it is in a denser medium such as water. We now have a clear-cut, testable distinction between the two models. It should be easy to determine which model is consistent with the actual speeds of light in air and water. Unfortunately, light travels at a tremendous speed. It is no easy matter to measure the speed at which light travels across a room or through a body of water. All experimental techniques available at the time led to the conclusion that light travels almost instantly over any relatively small distance, either through air or through water. The experimentalists could not tell whether it went faster in one medium or in the other.

Through the seventeenth and eighteenth centuries, it was largely a matter of personal taste that determined which model a given physicist would support. Because of the overwhelming success of Newton's theories of motion and gravity, his particle model of light was by far the most widely accepted. (Although Newton preferred the particle to the wave model, he did have his doubts about the model. To explain certain phenomena, he supposed that the particles produced waves in the ether. However, most of Newton's followers simply took his preference as dogma and ignored his doubts.) By the end of the eighteenth century, the wave model was all but forgotten.

The situation began to change, however, in the first few years of the nineteenth century. In 1802, the English physicist Thomas Young (1773 – 1829) showed that light exhibits a wave phenomenon called interference. When two waves (such as two water waves or two sound waves) combine with each other, the waves reinforce each other in certain regions and cancel each other in other regions. Young showed that a similar phenomenon occurs when a light beam is split into two parts and then recombined. If the recombined beam shines on a screen, some parts of the screen are light and others are dark. From the dimensions of the pattern produced, Young calculated that the wavelength of light is on the order of 5×10^{-7} m. This very small value for the wavelength explains why the phenomenon of light interference had not been detected earlier. For such small wavelengths, interference can be observed only under very special conditions. (In fact, the absence of any observed interference effect for light had earlier been argued as evidence against a wave model for light. It was one of the reasons that led Newton to prefer the particle model.)

Over the following years, additional evidence for wave behavior by light was accumulated, primarily by the French experimental physicist Augustin Fresnel, and the pendulum began to swing toward the wave model. The apparent *coup de grace* to the particle model of light was applied in 1850 by another French physicist, Jean Foucault, when he measured the speed of light in water. Recall that the two competing models lead to contradictory predictions on this matter: the particle model predicts a greater speed for light in water than in air, whereas the wave model predicts a lesser speed in water. Foucault found that the speed of light in water is only about three-fourths of that in air, seeming to settle the controversy in favor of the wave model for light.

A problem remained for the wave model, however. If light is propagated as waves, what is the nature of the medium through which the light waves propagate? What is it that is "waving"? Since the seventeenth century, a basic part of scientific thinking had been the existence of a material substance called the ether (often spelled aether). This substance was believed to permeate the entire universe. It was natural to assume that light waves are disturbances propagating through the ether. However, waves travel with the greatest speeds in the most rigid solids. To explain the very great speed of light, it was necessary to assume that the ether is far more rigid than iron or any other known solid. Yet at the same time, it is obvious that the planets, comets, and other bodies move through the ether without any apparent effects of friction or drag. How can a substance be extremely rigid and yet offer no resistance to material objects moving through it? It was difficult to form a clear picture of the nature of the ether, and hence it was difficult to form any clear picture of the nature of light waves as disturbances traveling in the ether.

In the second half of the nineteenth century, the situation was somewhat clarified in a rather unexpected fashion -- through developments in the study of the nature of the electric and magnetic forces.

8.2 ELECTROMAGNETISM

While Hans Christian Oersted (1777 – 1851) was delivering a lecture demonstration in 1820, he accidentally discovered that a magnetic compass needle was deflected whenever an electric current was passed through a nearby wire. This was the first evidence of any connection whatsoever between electricity and magnetism. 0ersted's discovery touched off an avalanche of experimental effort to clarify this connection, culminating in 1831 with the discovery that a magnet can be used to produce an electric current -- a discovery made independently by Michael Faraday (1791 – 1867) in England and Joseph Henry in America.

Faraday was to play an important role in the development of theories about electromagnetism. He introduced into physics the concept of a force field. Until this time, such forces as electricity, magnetism, and gravity were thought to act at a distance without the involvement of any intermediate interactions. To Faraday this was nonsense. To get around the problem of action at a distance, Faraday pictured a magnet as being surrounded by a magnetic field that is a physically real thing occupying the space around the magnet. This magnetic field then exerts the magnetic force directly on any magnetic object that is located within the field. Similarly, Faraday imagined a charged object as surrounded by an electric field that exerts an electric force on other charged particles located within the field, and he visualized a mass as surrounded by a gravitational field that exerts gravitational force upon other masses within the field. Faraday considered the possibility that his fields were related in some way to the ether, but apparently, he was never thoroughly convinced one way or the other about this idea. In terms of electric and magnetic fields, Oersted's discovery can be described as the fact that an electric current produces a magnetic field, which then exerts a force directly on the magnetic compass. Similarly, the discovery of Faraday and Henry can be described as the fact that a changing magnetic field produces an electric field, which then causes an electric current to flow in a wire.

Most of Faraday's contemporaries did not take his field concept very seriously. They regarded it as a crutch for those who could not handle the abstract mathematical formulas that most compactly described the "reality" of electric and magnetic interactions. However, there was one physicist who did recognize the value of Faraday's new idea. James Clerk Maxwell (1831 – 1879), a brilliant theoretical physicist and mathematician, used Faraday's field concept as the basis for his tremendously successful theory of electromagnetism.

In 1865, Maxwell was able to show that all the experimental results concerning electromagnetic phenomena could be represented by a set of four equations involving electric and magnetic fields. However, Maxwell soon realized that these four equations were inconsistent with the law of conservation of charge. That is, if the equations were correct as they had been written, then charge could be created or destroyed. For what can only be called philosophical reasons, Maxwell did not for a minute believe that this could be true. He modified the equations to make them consistent with the law of conservation of charge. In order to do this, he had to assume not only will an electric current give rise to a magnetic field (0ersted's discovery), but also that changes in an electric field will give rise to a magnetic field. At the time, there was no experimental evidence to justify this assumption, but it proved to be a valid one. The four equations written in the form that is consistent with the law of conservation of charge are now known as Maxwell's equations. They constitute one of the most beautifully compact and powerful theoretical structures in physics.

James Clerk Maxwell
(1831 – 1879 * England)

1860 – derived a formula for the distribution of molecular speeds in a gas.

1865 – demonstrated that all electromagnetic phenomena known at the time could be described by a set of four differential equations and that the equations predicted the existence of electromegnetic waves. From the theory, the speed of these waves was calculated to be the same as the recently measured speed of light. He concluded that light was an electromagnetic wave and predicted the existence of other types of electromagnetic waves of longer and shorter wavelengths.

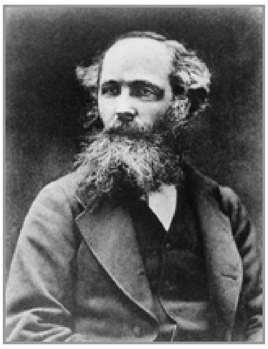


Figure 8.2 Maxwell

Maxwell's assumption that a change in an electric field will give rise to a magnetic field has profound physical consequences. If a changing electric field produces a magnetic field, and a changing magnetic field produces an electric field (Faraday's and Henry's discovery), then there is the possibility that a self-perpetuating disturbance of electric and magnetic fields can be produced. Maxwell was able to show that such a disturbance would propagate as a wave. From his equations and known data about electric and magnetic phenomena, From his theory he was able to calculate the speed at which such an electromagnetic wave would propagate. This speed turned out to be the speed of light, which had been measured some fifteen years earlier. This immediately suggested to Maxwell that light waves must be electromagnetic waves: “We can scarcely avoid the inference that light consists in the transverse undulations of the same medium which is the cause of electric and magnetic phenomena.” This “medium” was of course the ether, with the electric and magnetic fields regarded simply as manifestations of stresses and strains in the ether. Maxwell thus visualized light as an oscillating disturbance propagated through a very rigid, low-density ether. (Later the Michelson-Morley experiment and the theoretical work of Einstein forced the abandonment of the ether concept as an unnecessary and meaningless redundancy.) Maxwell's correct identification of light as an electromagnetic phenomenon constitutes one of the greatest triumphs of theoretical physics. In 1888, Heinrich Hertz was able to generate and detect electromagnetic waves of a type now employed in radio broadcasting. With these experiments, Hertz provided the first experimental verification of Maxwell's theoretical predictions.

8.3 PROPERTIES OF ELECTROMAGNETIC WAVES

Maxwell regarded electromagnetic waves as disturbances propagating through the mechanical medium of the ether. The ether was believed to permeate the entire universe, so it is not surprising that electromagnetic waves will propagate through a vacuum. However, this model of electromagnetic waves leads to the prediction that the measured speed of light will vary with the speed of the observer's motion through the ether. This prediction was not verified by experimental data, and Einstein's special theory of relativity was based upon the postulate that the speed of light is the same for any inertial observer. This theory in turn led to the abandonment of the concept of the ether, leaving physicists in some confusion about the exact nature of electromagnetic waves. There is undeniable experimental evidence that electromagnetic radiation has many wave properties, but it is also clear that it is not a mechanical wave propagating as an oscillating disturbance in some material medium. Perhaps it is best if we avoid saying firmly that electromagnetic radiation is a wave; instead we can simply describe it as having wavelike properties. If we are careful not to push the physical model too far, it remains useful today to describe electromagnetic radiation in terms of its wave properties.

Using Maxwell's theory of electromagnetism, it is possible to describe an electromagnetic wave as oscillating electric and magnetic fields such that the electric field is everywhere perpendicular to the magnetic field and both are perpendicular to the direction in which the wave propagates. This type of wave, in which the direction of the oscillation is perpendicular to the direction of propagation is called a transverse wave.

A wave on the surface of a body of water is another example of a transverse wave; the water molecules oscillate vertically while the wave propagates horizontally. The up and down motion of the water molecules can be shown by a cork floating on the surface of the water. A special terminology has been developed to describe the properties of waves. Some of these terms are useful in discussing the wavelike properties of electromagnetic radiation

Wavelength

Waves are repeating patterns. Consider a wave pattern moving along the surface of a body of water. The straight-line distance between two similar points on adjacent repetitions of the pattern is called the wavelength. The symbol used to represent wavelength is the Greek letter lambda (λ). The unit for wavelength is either meters (m) or angstroms (1Å = 10^{-10} m), depending upon the scale of the waves being described. In his experiments early in the nineteenth century, Thomas Young showed that the property of color is related to the wavelength of the light. Red light has the longest wavelengths (on the order of 7000 Å), whereas violet light has the shortest wavelengths (on the order of 4000 Å). When white light is passed through a prism, it separates on the basis of wavelength to form the familiar rainbow spectrum White light is simply a mixture of many different colors (wavelengths). The "color" black, of course, represents an absence of light of any of the visible wavelengths.

Period

The time required for one complete repetition of the pattern to pass a given position is called the period of the wave and is represented by the symbol T. There is another equivalent way to define the period – it is the time required for one complete oscillation of the disturbed medium.

Frequency

The frequency of a wave is defined as the number of oscillations (or the number of wave patterns passing a given position) per second. This quantity is related to the period but is more useful for our purposes. The symbol for frequency is the Greek letter "nu" (ν), and it is measured in units of cycles per second. In honor of Heinrich Hertz, the cycle per second has been given the name of the hertz (abbreviated Hz): 1 cycle/s = 1 Hz. The frequency of a wave is simply the reciprocal of its period: $\nu = 1/T$. Thus, if one complete oscillation of a wave takes place in $\frac{1}{3}$ s the frequency of the wave is 1/T or 3 cycles per second or 3 Hz.

Note the cycle is not a true physical unit in the sense that the second or meter is. The 'pseudo-unit' cycle is retained where needed for clarity, but it can be dropped at will. For example: (1 m) x (3 cycles/s) = 3 m/s.

Speed

For a mechanical wave, the speed with which the disturbance propagates through the medium is completely determined by the physical properties of the medium. The speed of the wave is the distance traveled by some point in the wave pattern (say, a crest) per second. The wave will travel a distance equal to one wavelength during the time required for one complete oscillation (that is, during one period). Therefore, we can write the following expression for the speed of a wave: $v = \lambda/T$. However, $1/T = \nu$, so we can substitute to obtain the more useful expression:

$$v = \lambda v$$

The speed of a wave is equal to the length of one complete disturbance multiplied by the number of disturbances passing a given point per second.

Although an electromagnetic wave is not a mechanical wave, the relationship is also valid for an electromagnetic wave. From the special theory of relativity, we know that the speed of an electromagnetic wave in a vacuum is the absolute physical constant c, so we obtain the important relationship

$$c = \lambda\nu$$

where λ is the wavelength and ν is the frequency of a particular electromagnetic wave. Because c is an absolute physical constant, the wavelength and frequency are not independent properties of an electromagnetic wave. If we know the value of either quantity for a given electromagnetic wave, we can compute the other.

Example 8.1

The yellow light emitted by a sodium lamp has a wavelength in air of 5890 Å. Determine the frequency of this light.

Solution

From the relationship $c = \lambda\nu$ or $\nu = c/\lambda$

$$\nu = c/\lambda = 3 \times 10^8 \text{ m/s} / 5890 \times 10^{-10} \text{ m} = 5.09 \times 10^{14} \text{ s}^{-1}$$

$$\nu = 5.09 \times 10^{14} \text{ cycles/s} = 5.09 \times 10^{14} \text{ Hz}$$

Intensity

The intensity of a wave is a measure of the total energy being transported by the wave. For a sound wave, the intensity is related to the loudness of the sound. For a light wave, the intensity is related to the brightness of the light. The symbol for intensity is I and its unit is the joule per square meter per second, or J/m^2-s. This property of an electromagnetic wave plays a very important part in our discussion of the photoelectric effect in Chapter 9.

8.4 THE ELECTROMAGNETIC SPECTRUM

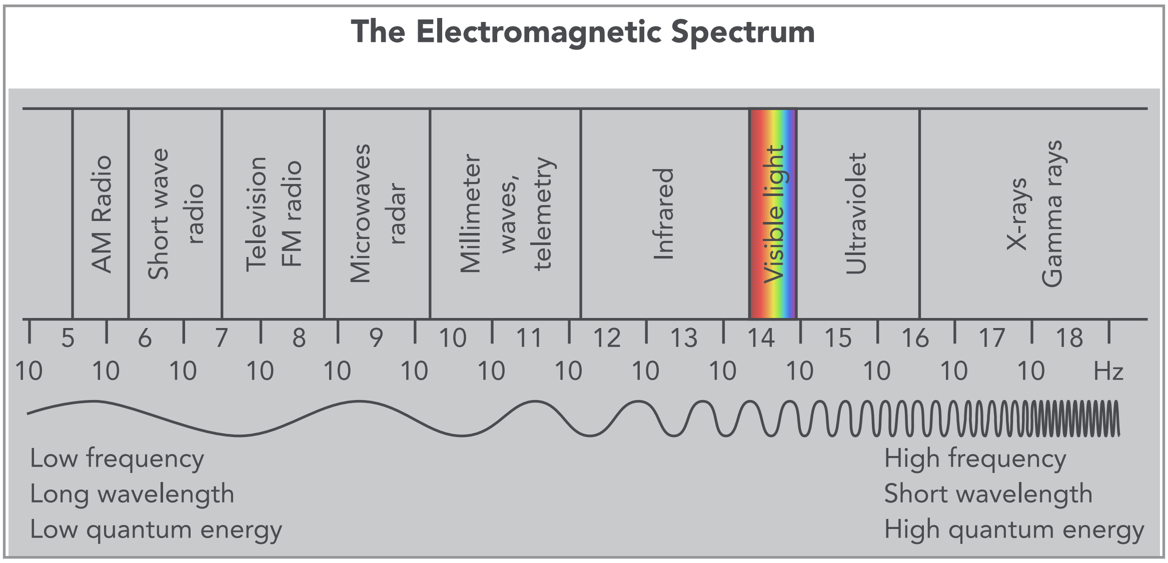


Figure 8.3 Electromagnetic Spectrum

After Maxwell had identified light as an electromagnetic wave, he realized electromagnetic waves other than light (that is, with wavelengths either longer or shorter than those of visible light) also must exist, but doubted they could ever be detected. However, Hertz in 1888 produced and detected electromagnetic waves of a wavelength much longer than light, waves that we would now describe as radio waves. By the first decades of the twentieth century, physicists were familiar with electromagnetic radiation ranging over an entire spectrum of wavelengths (see Figure 8.3). At one end of the spectrum are radio waves of incredibly great wavelength (millions of meters). At the other end are gamma rays of unimaginably small wavelength (less than 10^{-14} m).

All electromagnetic waves travel at the speed c in a vacuum. Any wavelength is theoretically possible. For a given electromagnetic wave, the wavelength and frequency are related by the equation $c = \lambda\nu$, so the spectrum can be described in terms of either wavelength or frequency. Electromagnetic waves can be produced in a number of different ways; each technique for production of waves typically produces waves of a restricted range of wavelengths (or frequencies). The names usually given to various regions of the electromagnetic spectrum are for the most part indicative of the ways in which the waves are produced. It is important to remember that, although electromagnetic waves of greatly differing wavelengths interact with matter differently and are produced by different mechanisms, they are nonetheless of exactly the same nature. Any electromagnetic wave, no matter what its wavelength, is represented by oscillating electric and magnetic fields and is propagated through space with the speed c.

The following list summarizes the major regions of the electromagnetic spectrum, in order of decreasing wavelength:

Radio waves are produced by oscillating electric currents. The waves used for AM radio broadcasts have wavelengths of hundreds of meters and frequencies of around 10^6 Hz (also called l megahertz, or 1 MHz). The waves used for FM and TV broadcasts have slightly higher frequencies and correspondingly smaller wavelengths. The microwaves used for radar, satellite communications, and cooking also are classified as radio waves; their wavelengths are on the order of centimeters (l centimeter = 10^{-2} m).

Infrared, visible, and ultraviolet light are produced by changes in the motions of the outer electrons of atoms. The usual sources of visible light are hot objects (the sun, heating elements, light-bulb filaments), atomic discharge tubes (neon signs, mercury-vapor lights), and materials that emit visible light after absorbing ultraviolet light (fluorescent bulbs). Visible light is defined as that narrow band of wavelengths that excite the receptor cells of the human retina and extends from about 7000 Å (4.3×10^{14}Hz) to about 4000 Å ($7{,}5 \times 10^{14}$ Hz). We see as a result of the reflection or emission of this radiation by the material objects around us. The wavelengths of infrared radiation range from a fraction of a centimeter to about 7000 Å (or 7×10^{-5} cm). We sense infrared radiation as heat, but some animals have special organs similar to eyes that presumably "see" infrared radiation much as we see visible light. The wavelengths of ultraviolet radiation range from about 4000 Å to tens of angstroms. Humans do not sense this radiation directly, but it is responsible for sunburn and for some types of skin cancer.

X-rays are produced by orbit changes in the inner electrons of heavy atoms and also when very high-speed electrons slow down. The penetrating ability of X-rays makes them useful in medical applications, but excessive exposure to X-rays can cause permanent damage to cells. The wavelengths of X-rays range from tens of angstroms to fractions of an angstrom.

Gama rays are emitted by the nuclei of atoms. They are even more penetrating than X-rays and also can be hazardous to our health. Gamma rays of wavelengths as short as 10^{-7} Å have been detected.

> *Ever since the time of Young and Fresnel, we know that light is a wave motion. We know the speed of the waves, we know their length, we know that the waves are transverse; in a word, we know completely the geometric relationships of this motion. These things no longer permit of any doubt, and a refutation of this view is unthinkable to the physicist. Insofar as human beings can know truth, the wave theory is certainty.*
>
> – Heinrrich Hertz, 1889

(Remember this when we discuss the photoelectric effect.)

Summary

Energy can be transferred from one point in space to another either by a wave or by a moving particle. By the time of Newton, the debate about the nature of light had reduced to the question of whether light is a wave or a stream of particles. Newton's support for the particle model carried the day through the eighteenth century. Early in the nineteenth century, experimental evidence of the wavelike properties of light began to accumulate, and by 1850 the evidence seemed firmly in favor of the wave model. In 1865, James Clerk Maxwell used his theory of electromagnetism to predict the existence of electromagnetic waves. Because he calculated the speed of such waves to be equal to the speed of light, Maxwell identified light as an electromagnetic wave. Maxwell also realized that electromagnetic radiation of wavelengths both longer and shorter than those of visible light must exist.

The classical model of electromagnetic radiation pictures oscillating electric and magnetic fields propagating through the ether at the speed of light. These fields were believed to be regions of stress and strain in the ether. Although the physical properties of the ether had to be very peculiar, most physicists at the beginning of the twentieth century accepted this general picture for the nature of electromagnetic radiation.

Einstein's special theory of relativity forced the abandonment of the concept of the ether, but it is clear that light and other types of electromagnetic radiation do have wavelike properties. Among these properties that are useful in describing electromagnetic radiation are wavelength, period, frequency, speed, and intensity. The speed of any wave is equal to the product of its wavelength and its frequency. For an electromagnetic wave traveling through a vacuum, the speed is the absolute physical constant c; thus, the frequency and wavelength of an electromagnetic wave are not independent properties. The major divisions of the electromagnetic spectrum are radio waves, infrared radiation, visible light, ultraviolet radiation, X-rays, and gamma rays. Radio waves have the longest wavelengths and lowest frequencies; gamma rays have the shortest wavelengths and highest frequencies.

Important concepts

Wave model of light; particle model of light; refraction; interference; field; Maxwell's theory of electromagnetism; transverse wave; wavelength; period; frequency; speed of a wave; intensity; electromagnetic spectrum.

Questions

1. The distance from the earth to the moon is determined most accurately by measuring the time required for a light pulse to travel from the earth to the moon and to be reflected back to the earth. In such an experiment, the time interval for the round trip is found to be about 2.6 s Use this fact to determine the distance between earth and moon.
2. A woman living on Mars is talking with her son on the earth by radio. If the son begins to speak as soon as he hears his mother pause, how long does she have to wait for his voice to reach her if Mars is 37 million miles from earth at the time?
3. List the following colors of visible light in order of increasing wavelength: yellow, red, blue, violet, green.
4. A radio wave has a wavelength of $\lambda = 1$ m. What is the frequency ν of this radio wave?
5. In what ways are gamma rays and visible light similar? In what ways are they different?
6. A beam of X-rays has a frequency of $\nu = 2 \times 10^{17}$ Hz. What is the wavelength?
7. List the principle divisions of the electromagnetic spectrum in order of decreasing frequency.
8. You are standing in the surf with waves passing you at a steady rate. You determine the distance between two successive wave crests to be 4 m and the time interval between their arrivals to be 6 s. What is the wavelength of the water waves? What is the frequency of the water waves? What is the speed of the water waves?
9. List several phenomena that demonstrate that electromagnetic waves transport energy through space from one point to another.
10. Show that $hc = 1.24 \times 10^4$ eV-Å.

    The following questions are of a more general nature. They have no single correct answer and are just something for you to think about. When possible, they are best answered in conversation with others.
11. Maxwell changed his equations to make them consistent with the law of conservation of charge, even though the change led to the prediction of physical phenomena for which there was no experimental evidence. Do you think this was a legitimate thing for Maxwell to do? What do you think would have happened to Maxwell's theory if subsequent experiments had failed to confirm his predictions?

12. The concept of light as a wave leads naturally to the question "What is waving?" Maxwell believed that it is the ether that is waving, but Einstein discarded this concept as unnecessary. Does it seem possible for a wave to propagate without a medium? Do you think it is possible to imagine a wave that does not exist in any physical medium? Why or why not?
13. In the general theory of relativity, Einstein showed that a gravitational field can be regarded as a "curvature" of space-time. He spent the rest of his life trying to develop a theory that would explain all force fields in terms of similar properties of space-time, but he was never able to find a satisfactory set of postulates that would account for all the known characteristics of electromagnetic and nuclear force fields. Discuss the reason why such a 'unified field theory' is so appealing for theoretical physicists.

9 THE PHOTOELECTRIC EFFECT

In 1888, while performing his famous experiments that verified Maxwell's theory of electromagnetism, Heinrich Hertz noticed that an electric spark between two electrodes occurs more readily when ultraviolet light is shining on one of the electrodes. Other physicists explored this phenomenon and soon discovered that ultraviolet light facilitates the electric discharge by causing electrons to be emitted from the electrode surface. This phenomenon is known as the photoelectric effect.

Ultraviolet light (and in some cases visible light) can free electrons from the surface of a metal.

It is ironic that Hertz's experiments, which so beautifully confirmed Maxwell's electromagnetic theory, simultaneously resulted in a discovery that would later be used by Albert Einstein to illustrate a fundamental flaw in Maxwell's theory.

9.1 THE PHOTOELECTRIC EXPERIMENT

The fact that electromagnetic radiation can "knock" electrons out of a metal surface is not in itself too surprising. Electromagnetic waves transmit energy, and it is entirely consistent with classical theory that this energy could be transferred to the electrons of the metal, giving them energy to leave the metal surface. However, the detailed results of quantitative experiments are not at all consistent with predictions based on the wave model of electromagnetic radiation. Figure 9.1 shows a typical apparatus used to study the photoelectric effect. A target plate and a collecting plate are enclosed in a vacuum tube, a glass tube from which almost all air has been removed.

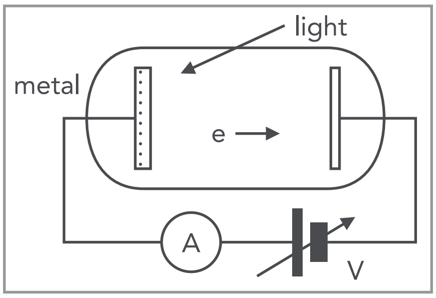


Figure 9.1 The photoelectric effect apparatus.

Electromagnetic radiation is directed onto the target plate. If the photoelectric effect occurs, some of the emitted electrons will travel across the tube and be absorbed by the collector plate. These electrons will then flow through the wire connection back to the target plate. Such a flow of electrons through a metal is called an electric current. Current is measured in the unit of coulombs per second, which is given the special name of amperes abbreviated A. That is, 1 C/s = 1 A. A device called an ammeter provides a continuous measurement of the current flowing through the wire. So long as the photoelectric effect is occurring, this fact will be indicated by some reading greater than zero on the ammeter.

An electric battery is a device that uses chemical reactions to pump electrons from one terminal of the battery to the other. In the apparatus of Figure 9.1, a battery is connected in such a way that it pumps electrons away from the collector plate back toward the target plate. This is done be building up a negative charge on the collector plate, and this charge repels the electrons coming from the target plate. The extent to which the battery will build up such a charge is measured by a property of the battery called its voltage. The greater the voltage of the battery, the larger the negative charge it will produce on the collector plate. The success of the electrons in reaching the collector despite the repelling force is related to the kinetic energy of the electrons leaving the target plate. The greater the initial kinetic energy of the electrons, the greater the voltage needed to prevent them from reaching the collector plate.

This apparatus permits us to measure two things about the photoelectric effect. When the applied voltage is zero, the ammeter reading gives a measure of the rate at which electrons are being emitted from the target plate. If we then increase the applied voltage until the ammeter reads zero (indicating that the applied voltage is great enough to repel the electrons from reaching the collector plate), then we can use this applied voltage to calculate the kinetic energy of the electrons are leaving the target plate.

The photoelectric experiment varies the intensity and the frequency (or wavelength) of the electromagnetic radiation directed onto the target plate and uses the apparatus of Figure 9.1 to measure the rate at which electrons are emitted and the kinetic energy with which they are emitted. The intensity of the radiation is a measure of the rate at which energy is transferred to the target plate.

The first, experiment keeps the frequency (wavelength) of the electromagnetic radiation constant while varying the intensity. As the intensity is increased, the rate at which energy is being transferred to the electrons of the target plate increases. According to the classical theories of physics, we would expect to observe greater numbers of electrons being emitted with greater kinetic energies as we increase the intensity of the radiation. In fact, the actual results of this experiment are somewhat surprising. If the photoelectric effect occurs, the kinetic energy of the electrons remains constant, but the rate at which electrons are emitted increases in direct proportion to the increase in intensity That is, doubling the intensity of the radiation doubles the rate at which electrons are emitted from the target plate but it has no effect on the kinetic energy with which each electron leaves the target plate.

Classical physics cannot explain this result. When more radiant energy is supplied to the metal, more electrons are knocked loose. Yet each electron somehow obtains exactly the same amount of kinetic energy from the interaction. If we make the intensity of the radiation very small, only an occasional electron is emitted, but this electron still emerges with the same kinetic energy. No similar effect can be observed in any experiment where we use water waves or sound waves to transfer kinetic energy to particles. It is impossible to imagine any explanation that will account for this result in terms of the classical wave picture of electromagnetic radiation.

Now let's look at another experiment. This time we keep the intensity of the electromagnetic radiation is held constant and its frequency (wavelength) is varied. Because the energy content of the radiation is held constant, we would expect the variations in frequency to have no effect upon the rate at which electrons are emitted or the energy with which they emerge from the metal. However, the actual experiment again produces results quite different from our prediction. If we start with low frequency radiation, we find that no photoelectric effect occurs at all. When the frequency ν is increased to some value ν_o (called the threshold frequency), the photoelectric effect begins to occur. As we increase the frequency beyond the threshold frequency, the kinetic energy of the emitted electrons increases linearly with increasing frequency, whereas the rate at which electrons are emitted decreases linearly with increasing frequency (see Figure 9.2).

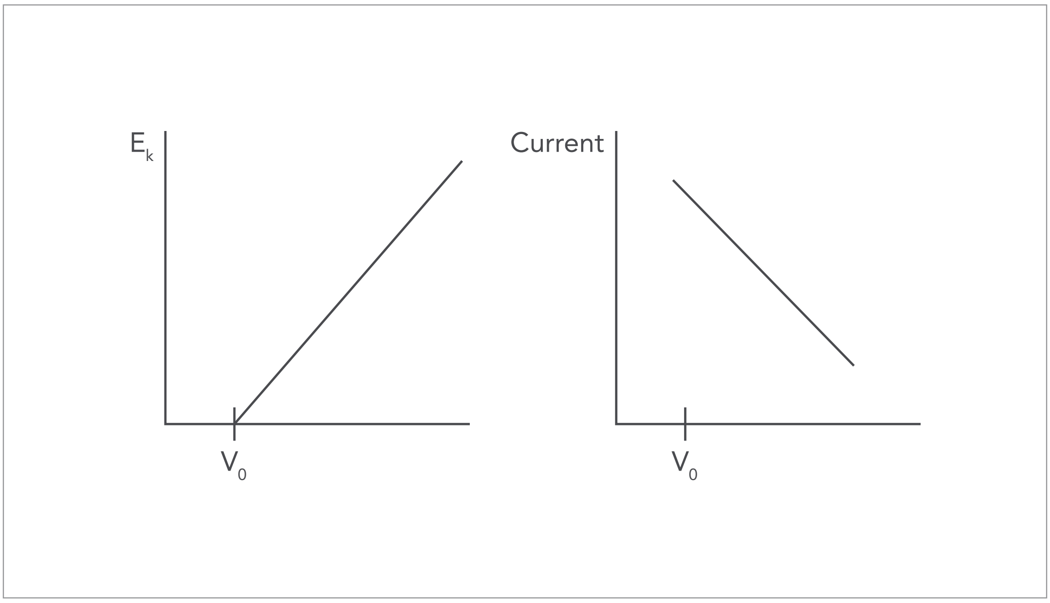


Figure 9.2

The results of this second experiment are quite unexpected. A high-frequency beam of electromagnetic radiation liberates electrons with a certain kinetic energy. If we decrease the frequency, we observe an increase in the rate at which electrons are emitted, and each electron emerges with a smaller kinetic energy. In some way, the frequency of the radiation determines how the energy falling on the target plate is transferred to the emitted electrons. In this experiment, the rate at which energy falls on the plate is kept constant. Yet we find that a higher frequency causes the emission of fewer electrons each having greater kinetic energy, whereas a lower frequency causes the emission of more electrons each having smaller kinetic energy.. Furthermore, when we lower the frequency of the radiation below the threshold frequency, we find that no electrons at all are emitted even though the same amount of energy is being delivered to the target plate. Once again, it is quite impossible to find any explanation based on the classical wave picture of electromagnetic radiation that will account for these results.

Consider some specific experimental results. If we use a target plate made of sodium and shine violet a light of frequency 6.7×10^{14} Hz on the target plate, the photoelectric effect will occur. If we keep the intensity of the light constant but decrease its frequency (increase the wavelength), blue and green light also produce the photoelectric effect, but the electrons are emitted in greater numbers with: correspondingly smaller kinetic energies. Suddenly, a frequency of 5.0×10^{14} Hz, just as we are about to pass from green light to yellow light, the photoelectric effect ceases. If we use yellow, orange, or red light to illuminate the sodium target plate, we observe no photoelectric effect. In fact, a low-intensity violet light will free electrons from sodium, whereas a high-intensity red light will not.

Further experiments show that the threshold frequency depends on the metal that is used for the target plate. Each metal has a different threshold frequency, but in each case the kinetic energy increases linearly with increasing frequency after the threshold frequency has been passed (see Figure 9.2).

These experimental results are completely incompatible with the accepted classical model of electromagnetic radiation where the energy is transferred to the surface of the metal in the continuous manner of a wave.

9.2 EINSTEIN'S QUANTUM HYPOTHESIS

The photoelectric experiments were not the first indication that something unexpected happens when electromagnetic radiation interacts with matter. The classical theories had also failed to provide an explanation for the frequencies of radiation emitted by heated objects. In 1900, Max Planck (1858 – 1947) was able to provide a successful explanation of this thermal radiation. Planck's equations (derived to fit the experimental evidence), suggested that the radiant energy is emitted from heated objects in discrete packets of energy, called quanta, with the energy of each quanta equal to hν where ν is the frequency of the radiation and h is the constant 4.14 x 10^{-15} eV-s. This constant, now called Planck's constant, has subsequently proved to be one of the fundamental constants in nature. Like the constant c it appears in many important physical laws, but we have no explanation in our theories of why it should have this particular value. The units of Planck's constant are energy times time. In SI units this would be joule-second but for our purposes the smaller energy unit, the electron-volt (eV) is more appropriate. 1 eV = 1.602 x 10^{-19} J. In SI units, h = 6.63 x 10^{-34} J-s.

Max Planck
(1858 – 1947 *Germany)

1900 – developed the concet of quantized energy states to explain the details of the thermal radiation spectrum. This was the beginning of the quantum revolution in physics.

1914 – established a prestigious professorship to lure Einstein back to Germany.

1918 – Nobel Prize in Physics.

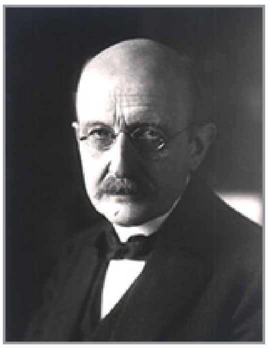


Figure 9.3 Max Planck

Planck did not really believe that energy existed in such particle-like packets, but he offered his theory with the hope that someone else would find a more reasonable explanation for the success of his mathematical equations. In any case, the assumption that energy is quantized was needed only for the moment of emission. Like all other physicists, Planck never doubted that the radiant energy exists in the form of waves after it is emitted.

Because of the theoretical similarities between electromagnetic radiation and an ideal gas, Einstein began to believe that electromagnetic radiation may possess a particle-like nature. In one of the papers Einstein published in 1905, he proposed such a model and suggested it as an explanation for the results of the photoelectric experiment. What Einstein proposed was that radiant energy is not only emitted as discrete quanta but also remains quantized throughout its existence. That is, he extended the idea inherent in Planck's explanation of thermal radiation to the hypothesis that electromagnetic radiation travels through space and interacts with matter as discrete, localized packets of energy. .Einstein's Quantum Hypothesis was based on the following postulates.

1. **Electromagnetic radiation is propagated in the form of discrete packets (or quanta) of energy. These quanta were later given the name photons.**
2. **The energy of a photon is hν.**
3. **Photons travel at the speed of light.**

Using the basic relationship $\nu = c/\lambda$, we can write $E = hc/\lambda$ where $hc = 1.24 \times 10^4$ eV .

Albert Einstein
(1879 – 1955 * Germany)

1905 – developed special relativity and the photon model of electromagnetic radiation.

1916 – developed the general theory of relativity.

1935 – with two colleagues, developed the EPR thought experiment which he believed showed quantum mechanics to be an incomplete description of physical reality.

1921 – Nobel Prize in Physics.

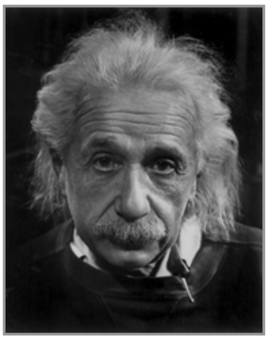


Figure 9.4 Albert Einstein

It is difficult today to appreciate how bold Einstein's 1905 quantum hypothesis was. This young and unknown physicist was suggesting a return to the long-rejected particle model of light at a time when the leading physicists were sure beyond any doubt that light must be a wave. (See the quote at the end of Chapter eight.) Furthermore, Einstein's formulation treats light as particles, but retains the properties of frequency and wavelength that have meaning only when defined in terms of waves. Like Planck, Einstein regarded his quantum hypothesis as an unsatisfying model that would eventually be replaced by a more complete model giving a comprehensible picture of electromagnetic radiation.

Example 9.1

Radio waves are traveling through space with a frequency of 2.0×10^{7} Hz. What is the energy associated with a single photon of this radiation?

Solution

Using Einstein's relationship $E = h\nu$, we have

$$E = (4.14 \times 10^{-15} \text{ eV s}) \times (2.0 \times 10^{7} \text{ cycles/s}) = 8.28 \times 10^{-8} \text{ eV}.$$

(Recall that the cycle is not a true unit and is dropped when possible.)

Example 9.2

What is the energy associated with a beam of gamma rays of frequency Hz?

Solution

$$E = h\nu = (4.14 \times 10^{-15} \text{ eV s}) \times (1.5 \times 10^{21} \text{ cycles/s}) = 6.21 \times 10^{6} \text{ eV} = 6.21 \text{ MeV}$$

where an MeV = 1 million eV. Note that a gamma ray photon has approximately 10^{14} times more energy than a radio photon.

Example 9.3

A beam of X-rays has a wavelength of 12.4 Å. What is the energy of an individual photon? If the beam transports 600 J of energy per seconds, how many photons must the beam deliver per second?

Solution

$$E = hc/\lambda = (1.5 \times 10^{21} \text{ eV})/12.4 = 1000 \text{ eV}$$

In SI units, the energy of a single photon is

$$E = 1000 \text{ eV} \times (1.6 \times 10^{-19} \text{ J/eV}) = 1.6 \times 10^{-16} \text{ J}$$

If 600 J are delivered by the beam each second, then the number of photons delivered per second is

$$(600 \text{ J})/(1.6 \times 10^{-16} \text{ J/photon}) = 3.75 \times 10^{18} \text{ photons.}$$

This example is typical in that a beam of electromagnetic radiation producing macroscopically observable effects involved a tremendously large number of photons.

Einstein explained the photoelectric effect in terms of his quantum hypothesis as follows. An electron in the target plate can be knocked out of the plate only if it happens to be struck by a single photon in the beam of electromagnetic radiation. When this happens, all of the photon's energy is transferred to the electron. (Remember a photon is essentially just a packet of energy; it no longer exists after its energy is transferred to the electron.) Part of this newly acquired energy becomes rest-mass energy that frees the electron from the metal, and any energy that is left over becomes kinetic energy of the free electron. Thus Einstein proposed his photoelectric equation:

$$E_k = h\nu - W$$

where E_k is the kinetic energy of the emitted electron, $h\nu$ is the energy of the photon absorbed by the electron, and W is the energy required to free an electron from the particular metal used for the target plate. W is called the work function of the metal. It is the energy needed to overcome the attractive electric force between the negatively charged electron and the positive charge left behind if the electron leaves the surface of the metal.. The work function is also the increase in rest-mass energy when changing from a state where the electron is within the metal to a state where the electron is free from the metal.

Example 9.4

A target plate made of magnesium is irradiated with ultraviolet light of frequency 1.5×10^{15} Hz. The work function of magnesium is 3.7 eV. What will be the kinetic energy of electrons emitted from the target plate?

Solution

The energy of the photons in the beam is

$$E = h\nu = (4.14 \times 10^{-15} \text{ eV s}) \times (1.5 \times 10^{15} \text{ cycles/s}) = 6.2 \text{ eV}.$$

When one of these photons is absorbed by an electron in the magnesium, the electron acquires 6.2 eV of energy. In freeing the electron from the magnesium, 3.7 eV is used up to free the electron from the magnesium surface.

$$E_k = h\nu - W = (6.2 \text{ eV}) - (3,7 \text{ eV}) = 2.5 \text{ eV}.$$

As an analogy for Einstein's explanation of the photoelectric effect, consider a baseball suspended by a string. Energy is transferred to the baseball when it is struck by a bat. Part of the energy given to the baseball is used to break the string. If the bat does not provide enough energy, the baseball remains attached to the string. If the bat does provide enough energy to break the string, then the ball moves off with kinetic energy equal to the amount of energy provided by the bat minus the amount that was used to break the string. The amount of energy needed to break the string is the mechanical analog of the work function in the photoelectric effect. When the energy given to the baseball is less than the amount needed to break the string, the baseball simply bounces around on the end of the string until the energy it received from the bat is dissipated. Similarly, if the photon does not provide an amount of energy as large as the work function the electron remains a bound particle and the absorbed energy is dissipated as heat.

Einstein's explanation of the photoelectric effect is completely consistent with the experimental data described early in this chapter. First, consider the experiment where the electromagnetic radiation is maintained at a constant frequency while the intensity is varied. In Einstein's model, intensity (a measure of the rate at which energy is transferred by the beam) is determined by the energy per photon and the number of photons per second delivered by the beam. Increasing the intensity of the beam while holding the frequency constant means increasing the number of photons striking the target plate each second. (The energy of each photon is constant because the frequency of the radiation is held constant.) Such a change should lead to an increase in the number of electrons emitted from the plate per second, but no change in the kinetic energy of each electron emitted. (Recall that $E_k = h\nu - W$, so E_k should not change in this experiment.) This is consistent with the observed experimental results.

Now consider the other experiment, in which the intensity of the electromagnetic radiation is held constant while its frequency is increased. This would correspond to a constant rate at which energy is delivered to the plate. In Einstein's model, increasing the frequency would increase the energy per photon, so in order to hold the intensity constant, the number of photons striking the plate per second decreases. So long as $h\nu$ is smaller than W for the target plate, none of the photon-electron collisions will free any electrons. When the frequency ν is increased past a threshold value, $h\nu_o = W$, some electrons will begin to emerge from the metal with very small kinetic energies. As the frequency is increased, the kinetic energy of each emitted electron also increases. However, because the number of photons in the beam is decreasing, the number of collisions and hence the number of emitted electrons also decreases. (Be sure you see why a constant intensity with increasing frequency implies a decreasing rate of photons striking the target plate.) Again, the predictions of Einstein's model are consistent with the observed experimental results. We should point out that only sketchy experimental results were available when Einstein proposed his quantum hypothesis in 1905. It was almost another decade before the fully detailed results we have described were available. Once again, we see that Einstein relied more upon his own physical intuition than upon empirical data in deriving his theoretical formulations.

9.3 PARTICLES OR WAVES?

As indicated by the quotation from Heinrich Hertz at the end of Chapter 8, physicists at the beginning of the twentieth century were convinced that electromagnetic radiation (including light) is a wave phenomenon. The experimental evidence seemed to have established the validity of the wave model beyond any possible doubt. As a result, Einstein's quantum hypothesis was greeted with almost total skepticism in the physics community.

In 1913, in a letter to the Prussian Academy of Science recommending the establishment of a special position for Einstein, Planck wrote:

> There is hardly one among the great problems, in which modern physics is so rich, to which Einstein has not made an important contribution. That he may sometimes have missed the target in his speculations, as, for example, in his hypothesis of light quanta, cannot really be held too much against him, for it is not possible to introduce fundamentally new ideas without occasionally taking a risk.

As late as 1917, the great American experimental physicist Robert Millikan (1868 –1953) wrote:

> Despite then the apparently complete success of the Einstein equation [describing the photoelectric effect], the physical theory, of which it was designed to be the symbolic expression, is … untenable

Millikan made this statement despite the fact that the confirmation of the success of Einstein's equation came from experiments carried out by Millikan himself. Millikan went on to say, "we are in the position of having built a very perfect structure [the photoelectric equation] and then knocked out entirely the underpinning [the photon model] without causing the building to fall." The feelings of Planck and Millikan were typical of the attitude of the general community of physicists.

Not only was the wave nature of light well established through such phenomena as interference and refraction, but Einstein's photoelectric equation itself ascribed the wave properties of wavelength and frequency to the particle-like photons. The idea that something could simultaneously be both a wave and a particle seemed nonsensical. Attempts to form a single model combining both earlier models proved unsuccessful. For example, such models as particles moving along wavy paths or waves clustered into particle-like bundles failed to produce predictions consistent with the experimental evidence.

It was not until 1923 that certain experimental evidence (to be discussed in the next chapter) finally forced physicists to begin taking seriously Einstein's photon model of electromagnetic radiation. By this time there could no longer be any doubt that light exhibits both wavelike and particle-like properties, but no physical model could successfully explain the experimental observations. This situation was most unsettling for physicists, including Einstein, who wrote in 1924: "We now have two theories of light, both indispensable, but, it must be admitted, without any logical connection between them, despite twenty years of colossal effort by theoretical physicists."

Let us summarize the quandary posed by the question of the nature of light. Certain physical phenomena (such as the photoelectric effect) can only be explained if electromagnetic radiation is regarded as a stream of particles called photons. The most important property of a photon is its energy, which is proportional to the frequency of the radiation (a property measured by using the assumption that the radiation is a wave phenomenon). The photon energy depends upon the process by which the radiation was produced, and in turn the photon energy determines the way in which the radiation will interact with matter. For example, the photons of visible light have just the correct range of energies to excite the receptor cells in the retina of the eye. Different photon energies within this range are perceived as different colors. Photons of infrared light do not have sufficient energy to excite the receptor cells, so the human eye does not perceive infrared light. On the other hand, a photon of ultraviolet light has too much energy to interact with a receptor cell in the way that produces visual effects, so this light also is invisible to humans. The photons of radio "waves" have the lowest energies, and they pass through the human body with very little interaction. The photons of gamma rays have the highest energies, and these photons can cause significant damage to the body if they collide with electrons in the atoms of the organism. Figure 9.6 illustrates the entire electromagnetic spectrum as a function of photon energy, and Figure 9.7 shows an expansion of this diagram for the visible portion of the spectrum On the other hand, the photon model fails entirely to account for such phenomena as light interference. There can be no doubt that light sometimes behaves in a fashion that can be explained only by a wave model, and its wave properties of frequency and wavelength remain important even in the photon model. These difficulties were resolved with the development in the 1920s of an entirely new physical theory called quantum mechanics. Its implications for the nature of the universe are even stranger and more disturbing than those of the theory of relativity. The quantum hypothesis put forward with such skepticism by Planck and Einstein led to a revolution in our ideas about reality that these pioneers and many other physicists were never able to accept. Nonetheless, the theory of quantum mechanics has proven valid in terms of its experimental predictions, and physicists today have few doubts about it. We shall explore the implications of the quantum hypothesis throughout most of the rest of this book.

> *All these fifty years of conscious brooding have brought me no nearer to the answer to the question "What are light quanta?" Nowadays every Tom, Dick, and Harry thinks he knows it, but he is mistaken.*
>
> – Albert Einstein, 1951

Summary

In the photoelectric effect, electrons are emitted from a metal surface that is illuminated by ultraviolet (or, in some cases, visible) light. The results of quantitative experimental studies of this phenomenon are inconsistent with the predictions of the classical wave model of electromagnetic radiation. For a given metal, the photoelectric effect does not occur when the incident radiation has a low frequency. There is a certain threshold frequency for the incident radiation that can cause the photoelectric effect. When the photoelectric effect does occur, the kinetic energy of the emitted electrons increases with increasing frequency of the incident radiation. For a given frequency, a higher intensity of radiation leads to a greater number of emitted electrons but does not affect their kinetic energies. These results cannot be explained if the radiation is assumed to be a wave.

In 1905, Albert Einstein suggested an explanation for the effect based on the assumption that the energy of the electromagnetic radiation is not continuously distributed but is localized in quanta (later called photons). The energy of each photon is directly proportional to the frequency of the radiation. That is, $E = h\nu = hc/\lambda$, where h is Planck's constant. Einstein based his work primarily upon theoretical considerations, but detailed experimental results became available a decade later, and these results were in complete accord with the predictions of Einstein's quantum hypothesis.

According to Einstein's model, the photoelectric effect occurs when the entire energy of a single photon is transferred to an electron on the surface of a metal. Part of this energy is used to free the electron from the metal; this work function is a characteristic of the metal. Any remaining energy becomes the kinetic energy of the freed electron. If the energy of the photon is less than the work function, then the photoelectric effect will not occur. It is the energy of the photon (proportional to the frequency of the radiation) that determines whether or not the photoelectric effect will occur and if it does occur, what the kinetic energy of the emitted electrons will be. The intensity of the radiation is determined by the energy of each photon and the number of photons striking the metal per second. The rate at which electrons are emitted from the metal is determined by the number of photons striking the metal per second. The success of Einstein's quantum hypothesis in predicting the experimental results of studies of the photoelectric effect led to a disturbing paradox. Electromagnetic radiation clearly acts in many ways as a wave phenomenon, and yet in the photoelectric effect and in the case of thermal radiation it exhibits particle properties. Neither a wave model alone nor a particle model alone can account for all the properties of electromagnetic radiation. Yet there seems to be no way to combine these two models into a single satisfactory physical model for the nature of electromagnetic radiation. Light seems to behave under some circumstances as a stream of particles and under other circumstances as a wave. What is its true physical nature?

Important concepts

Photoelectric effect; Einstein's photoelectric equation; threshold frequency; work function of a metal; Planck's constant; quantum hypothesis; photon; wave-particle duality.

Questions

1. Describe the photoelectric effect, outlining the main features of the experimental observations of the effect. In particular, describe how the properties of the ejected electrons depend upon the intensity and the frequency of the incident radiation.
2. Describe Einstein's explanation of the mechanism of the photoelectric effect. State and explain Einstein's photoelectric equation.
3. What does the work function represent? According to Einstein, how is it related to the threshold frequency?
4. The energy output of a 100-watt bulb is 100 joules per second (1 watt equals l J/s). Suppose that all this energy is emitted in the form of radiation of frequency 6×10^{14} Hz. What is the energy of a single photon in this radiation? Approximately how many photons are emitted per second by the bulb?
5. Silver bromide is a light-sensitive substance used in some types of photographic film. When a photon having sufficient energy strikes a molecule of this

substance, the molecule dissociates, leaving a dark spot of silver metal. Explain why such film may be handled in a darkroom under dim red illumination without exposing the film. Would it make any difference if the intensity of the red light is increased greatly? Could the film be handled without exposure under dim blue light?

6. A particular metal has a threshold frequency of 9.2×10^{14} Hz. What is the work function of this metal? In a particular photoelectric experiment, electrons are released from the surface of this metal with 3.0 eV of kinetic energy. What is the wavelength of the incident radiation in this experiment?
7. A metallic surface has a work function of 2.5 eV. What is the minimum frequency of incident light that will cause the photoelectric effect? If electromagnetic radiation of wavelength 3100 Å strikes this metal, what will be the kinetic energy of the emitted electrons?
8. In a photoelectric experiment, the intensity of the incident beam of ultraviolet light is held constant, but the frequency of the light is doubled. What will be the effect on the rate at which electrons are emitted? Why? What will be the effect on the kinetic energy of the emitted electrons? Why?
9. In the experiment of question 8, explain the effects expected as a result of the following experimental procedures. Both the frequency and the intensity of the incident radiation are doubled. Both the wavelength and the intensity are halved. The wavelength is increased above the threshold wavelength while the intensity is held constant. The intensity is doubled while the frequency is held constant.
10. Radiation of wavelength 775 Å strikes a target made of the metal nickel, whose work function is 5.0 eV. Calculate the speed of the emitted electrons.
11. Ultraviolet light with a wavelength of 2700 Å is observed to cause the emission of electrons with kinetic energies of 1.5 eV from a metal target. Will blue light of wavelength 4400 Å cause the emission of electrons from this target? Explain the reasoning behind your answer.
12. Green light causes the photoelectric effect from a certain metal, but yellow light does not. Will red light cause the photoelectric effect from this same metal? Will blue light? Explain your answers.
13. Using the value $c = 3 \times 10^8$ m/s, show that $hc = 1.24 \times 10^4$ eV Å.
14. Use the information on sodium given in this chapter to calculate the work function of sodium.
15. An incandescent light bulb produces radiant energy in the form of visible light at a rate of about 5 J/s. Assuming that the average photon energy visible light is 2.5 eV, how many photons of visible light does the bulb emit per second?

16. The human eye can barely detect a yellow light that delivers energy at a rate of 1.7×10^{-18} J/s. If the wavelength of the light is 5890 Å, how many photons are reaching the retina each second?

    The following question is of a more general nature. It has no single correct answer and is just something for you to think about. When possible, questions like this are best answered in conversation with others.

17. Using some appropriate physical analogy, explain why the results of the photoelectric experiments are inconsistent with the assumption that the incident electromagnetic radiation is a wave.
18. Could the photoelectric effect be caused by a very intense beam of radio waves focused on a metal plate? Be careful. Explain your answer.
19. Einstein once commented that the most incomprehensible thing about the universe is the fact that it is comprehensible. Comment on this statement and discuss its application to the study of the photoelectric effect.
20. Comment on the quotation from Hertz given at the end of Chapter 8. Considering the quotations at the end of this chapter, contrast the attitudes of Hertz and Einstein. Does this difference in attitudes have anything to do with the differences between classical physics and twentieth-century physics?

10 MORE EVIDENCE OF PHOTONS

Classical physics had produced thoroughly convincing evidence that electromagnetic radiation is a wave phenomenon. However, the wave model predicts that radiant energy should be emitted or absorbed in a continuous fashion, whereas the phenomena of thermal radiation and the photoelectric effect can be explained only with the assumption that radiant energy exists in discrete quanta. Even the creators of the quantum hypothesis were unable to believe in its ultimate validity. Although its predictions were consistent with experimental data, Planck and Einstein remained convinced that someone would develop a more acceptable explanation. Especially disturbing was the lack of any physical model to explain how electromagnetic radiation can have both wave and particle properties.

In the 1920s, however, other experimental evidence emerged for the quantum nature of electromagnetic radiation. Although physicists remained unable to find any satisfying picture of the nature of light, they were forced to the reluctant conclusion that the wave model alone is not satisfactory. Under certain conditions, there could no longer be any doubt that radiant energy is in fact quantized.

10.1 THE COMPTON EFFECT

Compton Effect

A photon of wavelength λ comes in from the left, collides with a target at rest, and a new photon of wavelength λ' emerges at an angel θ . $\lambda'$is greater than λ .

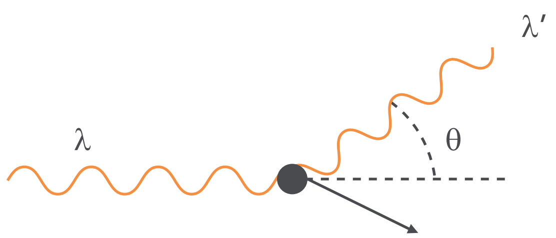


Figure 10.1 The Compton Effect

In the early 1920s, Arthur Compton (1892 – 1962) noticed that a scattered beam of X-rays contains radiation of a slightly greater wavelength (smaller frequency) than that of the incident radiation (see Figure 10.1). The exact wavelength of the scattered X-rays varies in a precise and rather simple way, depending on the angle at which the X-rays are scattered.

Despite his best efforts, Compton was unable to explain this phenomenon in terms of the classical wave model. By late 1922, Compton was forced to try seeking an explanation in terms of Einstein's quantum hypothesis.

When Einstein formulated the special theory of relativity, he chose to assume that the law of conservation of momentum is valid in relativistic physics. This led to the necessity for a new definition of momentum. In the second of his two 1905 papers on relativity, he introduced the relationship

$$E^2 = p^2c^2 + E_o^{\,2}$$

where E is the total energy of a particle (rest-mass energy plus kinetic energy), p is its relativistic momentum, and E_o is its rest-mass energy. Photons have zero rest mass, so the equation applied to a photon reduces to $E = pc$ or $hc/\lambda = pc$ or finally

$$p = h/\lambda = h\nu/c$$

By treating the photon and the electron as colliding particles and applying the laws of conservation of momentum and energy, Compton was able to calculate the expected energy of the scattered photons as a function of the scattering angle. The wavelength or frequency of the scattered X-rays could then be calculated from the energy of the scattered photons. The results of Compton's calculations were in complete agreement with his experimental data. Compton's experiment provided particularly direct and convincing evidence for the existence of photons. For his studies of what came to be called the Compton Effect, Compton shared the 1927 Nobel prize in physics.

In the Compton Effect, a photon of energy $h\nu$ and momentum $h\nu/c$ collides with an electron whose initial kinetic energy is very small compared to the photon energy. In the collision, energy and momentum are transferred to the electron, causing it to be ejected from the material, and a new photon is emitted with energy and momentum smaller than those of the incident photon. Thus, the photon is not simply scattered by bouncing off the electron. Instead, the incident photon ceases to exist, and a new photon of smaller energy moving in a different direction is produced (see Figure 10.1). The transfer of energy and momentum from incoming photon to emitted electron and photon occurs in such a way that total energy and momentum are conserved. If the energy of the emitted photon is $h\nu'$ (where ν' is the frequency of the scattered radiation) and the kinetic energy of the ejected electron is E_k, then the law of conservation of energy indicates that

$$h\nu = h\nu' + E_k$$

The work function of the metal has been omitted from this equation because it is so much smaller than the other energies that it can be neglected. (Remember that the photon energies for X-rays are far greater than those for the visible or ultraviolet light involved in the photoelectric effect.) The law of conservation of momentum leads to a more complicated vector expression because direction must be considered when dealing with momentum. When this expression is combined with the energy expression, it is possible to calculate the angle at which the photon is emitted (the scattering angle) given the angle at which the electron is ejected. Because this calculation involves vector algebra, we will not attempt to follow the details of the full quantitative explanation.

Example 10.1

Radiation of wavelength 10 Å is incident upon a thin metal foil. When a photon strikes an electron in the metal, the electron is ejected from the foil with a kinetic energy of 500 eV. What is the wavelength of the scattered radiation?

Solution

The energy of the incident photon is

$$E = hc/\lambda = (1.24 \text{ x } 10^4 \text{ eV Å })/10 \text{ Å} - 1.24 \text{ x } 10^3 \text{ eV} = 1240 \text{ eV}$$

Of this energy, 500 eV is transferred to the electron. The balance must be the energy of the scattered photon.

$$E' = 1024 \text{ eV} - 500 \text{ eV} = 740 \text{ eV}$$

From $E' = hc/\lambda'$ we obtain $\lambda' = hc/ E'$

$$\lambda' = hc/ E' = (1.24 \text{ x } 10^4 \text{ eV Å})/740 \text{ eV} = 16.8 \text{ Å}.$$

Note that the transfer of energy to the ejected electron results in an increase in the wavelength of the scattered radiation. We have neglected the work function of the metal because it is negligible compared to the other energies involved.

Although Compton had made every effort to explain his results in terms of the wave model, he was finally convinced that only the quantum hypothesis could account for the quantitative details. When he announced his results, he argued strongly that they confirm the validity of the quantum hypothesis. Other physicists reluctantly admitted there seemed to be no alternative but to accept the fact that electromagnetic radiation sometimes behaves as a stream of photons. As one prominent physicist joked, it seems like God runs electromagnetics by the wave model on Monday, Wednesday, and Friday, and the devil runs it by the quantum hypothesis on Tuesday, Thursday, and Saturday.

10.2 ANTIMATTER

Further evidence for the existence of photons came later from the study of certain types of particles called antiparticles. Before examining that evidence, let's first discuss the general concept of antimatter.

In 1929, P. A. M. Dirac (1902 – 1984) published a paper investigating the consequences of making quantum mechanics consistent with the special theory of relativity. His work led to many important understandings about the nature of matter, and for it he shared the Nobel Prize in physics in 1933. One of the results of Dirac's treatment was the prediction there should exist a particle having the same mass as an electron but having a positive rather than a negative charge. The existence of such a particle was confirmed in 1932 by Carl Anderson. Anderson was working at Caltech, carrying on the studies of cosmic rays begun by Millikan. Among the tracks created by cosmic rays in his cloud chamber, Anderson found tracks whose properties could be explained only by assuming that they were created by particles like those Dirac had predicted. Such particles are known as antielectrons, or positrons. Anderson shared the 1936 Nobel Prize in physics for his discovery of the positron. Further extension of Dirac's work led to the prediction that antiprotons and antineutrons should also exist. Because of their much greater masses, these particles are far more difficult to produce. However, antiprotons were finally produced and identified in 1955, and antineutrons were discovered in the following year. Theoretically, it should be possible for such antiparticles to combine to form anti-atoms, which in turn could form anti-molecules. Thus, it should be possible to have antimatter that is made up entirely of antiparticles. For example, an anti-helium atom would consist of two antielectrons (positively charged) surrounding a nucleus of two antiprotons (negatively charged) and two antineutrons (with zero charge). It is very difficult to construct such anti-atoms here on earth, because an antiparticle can combine with a normal particle in such a way that both wi1l disappear (be annihilated) with all of their rest-mass energies being converted to radiant energy. However, despite the difficulties, simple anti-atoms have been produced and identified experimentally.

By definition, the matter that makes up the earth is ordinary matter. We know that the other bodies of our solar system also are made of ordinary matter, because the sun ejects streams of particles that do not produce violent annihilation reactions when they contact the other objects in the solar system. Most physicists believe that the entire universe is constructed of ordinary matter, with the appearance of antimatter being a rare and short-lived event everywhere just as it is here on earth. The most widely accepted theory about the origin of the universe (the modern big-bang theory) assumes that matter has predominated over antimatter from very early in the history of the universe.

10.3 PAIR PRODUCTION

Like the photoelectric effect and the Compton Effect, the process that results in the creation of a positron is a phenomenon that can be explained only through the quantum hypothesis. This process is called pain production.

> **Under certain circumstances, a photon with sufficiently high energy can cease to exist and be replaced by a pair of particles – an ordinary particle and its corresponding antiparticle.**

Because the electron is the least massive particle, electron-positron pairs are the easiest to produce. The process of pair production is consistent with the law of conservation of charge. The photon has no charge and the charge on the electron and positon are of equal magnitude but opposite sign. Therefore, the net charge is zero before and after the interaction. However, it turns out that a process of pair production in isolation cannot be consistent with the law of conservation of momentum. The photon possesses momentum in the direction of its propagation. It turns out to be impossible for the net momentum of the electron-positron pair to be as large as the photon's momentum. Thus the process of pair production can occur only under certain circumstances, namely in the vicinity of a heavy nucleus.

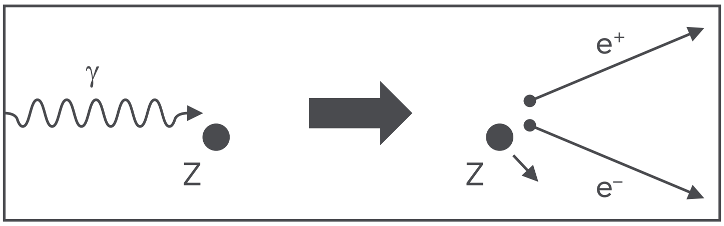


Figure 10.2 Pair Production

If we sum the momenta of the electron, the positron, and the heavy nucleus after the pair production, we can obtain a net momentum equal to the momentum of the photon before the pair. Because the heavy nucleus does not change charge during the pair production, the law of conservation of charge is still satisfied. The presence of the heavy nucleus is necessary because momentum can be conserved only through a transfer of part of the photon momentum to the recoil of the heavy nucleus.

The process of pair production is also consistent with conservation energy. Before the interaction, we have the radiant energy, hν, of the photon and the rest-mass energy of the heavy nucleus. After the interaction, we have the rest-mass energies and kinetic energies of the heavy nucleus, the positron, and the electron. The speed of the heavy nucleus after the interaction is always so small that the kinetic energy of the nucleus is extremely tiny compared to the kinetic energies of the positron and the electron. Although the heavy nucleus is necessary in order to conserve momentum, it is negligible in terms of conservation of energy. Essentially all of the energy of the photon goes to rest-mass energy and kinetic energy of the electron-positron pair.

$$h\nu = 2\ m_{oe}c^2 + E_k^+ + E_k^-$$

The value of m_{oe} has been accurately measured as m = 0.00055 u, so we have

$$2\ m_{oe}c^2 = 2 \times (0.00055\ u) \times (931\ MeV/u) = 1.02\ MeV$$

(Review chapter 6, Units and Calculations, if you don't see where we got these values.) We can now write the energy-conservation equation for pair production as

$$h\nu = 1.02\ MeV + E_k^+ + E_k^-$$

From this equation we see that pair production can only occur if the photon has an energy of at least 1.02 MeV (the energy required to create the rest masses of the electron-positron pair). Any energy above this his minimum value will appear after the interaction as the kinetic energy of the pair. Only gamma rays have photon energies great enough to cause pair production.

All of the predictions of the photon model for pair production have been confirmed experimentally. Pair production does occur only in the vicinity of a heavy nucleus, and the energy-conservation equation and the predicted minimum photon energy are both observed as expected. There is no way to explain these experimental data in terms of a wave model for the gamma rays. Here is another case where the behavior of electromagnetic radiation can be explained quantitatively only through acceptance of the quantum hypothesis.

It is important to recognize one characteristic typical of quantum interactions. Pair production is a discontinuous process in space-time. The photon does not split apart into the pair of particles. The photon does not gradually change into an electron-positron pair. Instead, pair production is an abrupt, discontinuous process. The photon ceases to exist, and the electron and positron appear. There is no intermediate stage. At one instant the photon exists and at the next an electron-positron pair exists. This is a very radical and completely non-classical concept, but it is one that forms a normal and basic part of the quantum picture of reality. We have already noted a similar discontinuity in the Compton Effect; the incident photon ceases to exist, and a new scattered photon appears as the electron acquires kinetic energy. The quantum hypothesis leads to a similar discontinuous model for the photoelectric effect. There will be much more to say about this discontinuity of quantum interactions in following chapters.

Example 10.2

A certain type of nucleus emits gamma rays with a frequency of 7.5×10^{20} Hz. What is the energy of a single photon in this radiation? If such a photon undergoes pair production, what will be the total kinetic energy of the electron-positron pair?

Solution

The energy of a photon is

$$E = h\nu = (4.14 \times 10^{-15} \text{ eV s}) \times (7.5 \times 10^{20} \text{ cycles/s}) = 31 \times 105 \text{ eV} = 3.1 \text{ MeV}.$$

If the photon undergoes pair production, 1.02 MeV of energy will be used to create the rest masses of the electron-positron pair. The remaining energy will become the kinetic energy of the pair:

$$E_k^+ + E_k^- = (3.1 \text{ MeV}) - (1.02 \text{ MeV}) = 2.08 \text{ MeV}$$

The electron and positron need not split this energy evenly. In some cases the electron gets more of it, and in other cases the positron gets more. The only requirement is that, however the energy is divided, both momentum and energy must be conserved in the process.

The production of proton-antiproton and neutron-antineutron pairs requires far greater energy to supply the rest-mass energies of the much more massive particles. Gamma rays cannot provide such large amounts of energy. These pair-production processes occur when some of the kinetic energy of an extremely high-speed particle is converted to rest-mass energy. This can be accomplished by having one high-speed particle collide with another that is at rest, or better yet with one that is moving in the opposite direction at high speed. When two high-speed particles collide head-on, most of their kinetic energy is converted to rest-mass energy, producing particle-antiparticle pairs. Many different types of particle-antiparticle pairs have been created in experiments where beams of high-speed particles are sent through each other in opposite directions.

10.4 PAIR ANNIHILATION

An antiparticle is annihilated soon after it is created, along with the corresponding ordinary particle. This process is called pair annihilation. As a positron moves through the ordinary matter in which it was created, it loses its kinetic energy as a result of interactions with the oppositely charged electrons in the matter. After it has lost almost all of its kinetic energy, it combines with an electron to form what is called a positronium atom. The positronium atom is held together by the electric attraction between the positron and the electron, just as a hydrogen atom is held together by the attraction between an electron and a proton.

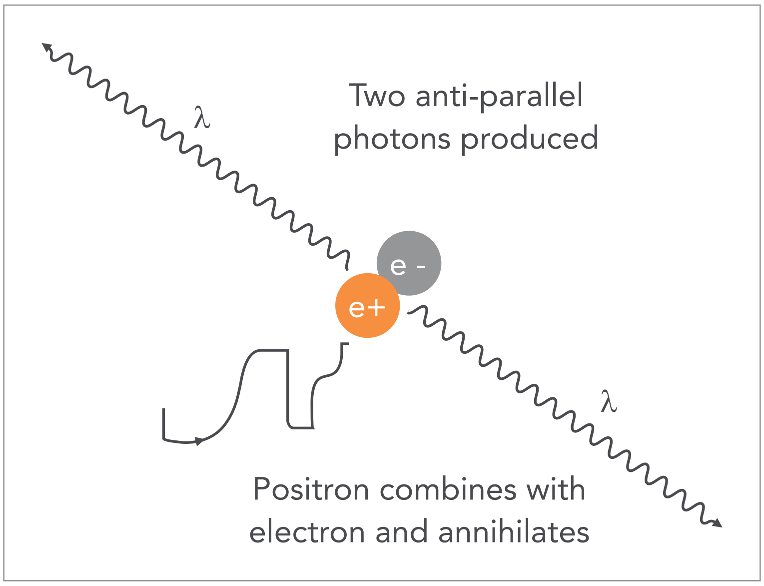


Figure 10.3 Pair Annihilation

After a brief lifetime of about 10^{-10} s, the positronium atom (that is, the electron-positron pair) annihilates. -- the pair of particles ceases to exist, and radiant energy is produced. This interaction obeys the laws of conservation of charge, momentum, and energy. The kinetic energies of the positron and electron before the annihilation are very small compared to their rest-mass energies, so the total energy of the positronium atom is approximately $2\ m_{oe}c^2$ = 1.02 MeV. Thus 1.02 MeV of radiant energy must be emitted in the pair annihilation. The net momentum of the positronium atom is approximately zero, so the net momentum after the annihilation also must approximately zero.

In a typical pair annihilation, two photons are emitted. Conservation of momentum requires that they have momenta of equal magnitudes but opposite directions. Because the momentum of a photon is proportional to its energy ($p = E/c$) the two photons also must have equal energies. Conservation of energy requires that the total energy of the two photons be 1.02 MeV. Thus a typical process of pair annihilation produces two photons, each of energy 0.51 MeV, moving in opposite directions (see Figure 10.3).

Occasionally, a process of pair annihilation is observed to create either a single photon or three photons. An annihilation producing only a single photon must occur in the vicinity of a heavy nucleus that can recoil to conserve momentum. This process is essentially the reverse of the process of pair production discussed earlier. As in pair production, the kinetic energy of the recoiling nucleus is negligibly small in comparison to the other energies involved, so the energy of the single photon emitted is approximately 1.02 MeV. The rather rare process producing three photons involves far more complicated relationships for the conservation of momentum and energy, so we will not consider it in detail here.

Although the process of pair annihilation involving formation of a positronium atom is by far the most probable process for such an annihilation, there is a very slight probability that the annihilation can occur as the result of a direct collision between a positron and an electron. A very important research tool in the study of elementary particles utilizes this fact by directing a beam of high-speed positrons through a beam of high-speed electrons moving in the opposite direction. In such a colliding-beam experiment, Pair annihilation is a very energetic process because the large kinetic energies are available in addition to the rest-mass energies of the electron-positron pair. This energy in turn produces particle-antiparticle pairs of relatively large rest-mass energies.

Example 10.3

What is the minimum energy released in the annihilation of a proton-antiproton pair?

Solution

The minimum energy will be released when the pair has negligible kinetic energy before the annihilation. In that case, only the rest-mass energy of the pair will be released in the annihilation. The rest mass of a proton (and therefore of an antiproton also) is almost exactly 1 u. Thus the total rese-mass of the pair before annihilation is about 2 u, so that the total rest-mass energy of the pair is

$$E_o = 2m_pc^2 = (2\ \text{u}) \times (931\ \text{MeV/u}) = 1862\ \text{MeV}$$

This is the minimum energy that can be released in the annihilation of a proton-antiproton pair.

10.5 WAVE-PARTICLE DUALITY

The wave-like nature of light was firmly established in the first half of the nineteenth century when interference experiments and measurements of the speed of light in various media (such as water and air) seemed to disprove the competing particle model for the nature of light. However, Einstein's 1905 explanation of the photoelectric effect and Compton's 1922 explanation of X-ray scattering by using a particle model for light established that light sometimes acts like a stream of particles. Does this mean that light (electromagnetic radiation) is both a wave and a stream of particles? No, that is nonsense.

The resolution of this dilemma lies partly in the realization that our descriptive abilities are limited. In describing electromagnetic radiation, we have used models based on our experiences with the macroscopic world (surface waves on water and bullets). We have forced a macroscopic description on a submicroscopic phenomenon, and we should not be terribly surprised if it does not fit. In the familiar macroscopic world, there is a clear distinction between waves and particles. Something is either a wave or it is a particle, and there are no ambiguities. On the atomic scale, we have been forced by experimental data to recognize that this distinction does not seem to exist. Just as we were forced to abandon our commonsense models when dealing with objects moving at speeds near the speed of light, we are apparently forced to abandon commonsense models when dealing with the nature of electromagnetic radiation in its interactions with particles on the atomic scale.

Electromagnetic radiation is wave-like in certain situations and it is particle-like in other situations. There is no macroscopic analogy for such a wave-particle duality, so we cannot visualize the true nature of electromagnetic radiation in terms of an analogy to our everyday experiences. Generally, the common approach is to use either the wave model or the particle model, whichever is more appropriate to the situation being considered. In discussing interference effects, it is useful to visualize electromagnetic radiation as if it were a wave. In discussing the photoelectric effect or Compton Effect, it is useful to visualize electromagnetic radiation as a stream of particles. Both models are needed for a complete description of electromagnetic radiation.

Wave-particle duality has actually been explicit in our treatment since introducing the quantum nature of electromagnetic radiation. The photon (particle description) has an energy that is proportional to the frequency (wave description) of the radiation. We will learn in a later chapter that wave-particle duality is a fundamental aspect of nature. It applies not just to electromagnetic radiation but to the entire micro-world. Electron, protons, atoms, etc., cannot be completely described in terms of particle properties. Rather, in some of their interactions, they exhibit wave properties. Einstein made a remark in a different context that is applicable here: "It is only with reluctance that one's desire for knowledge endures a dualism of this kind."

Summary

In the Compton Effect, a beam of X-rays is scattered by electrons. The scattered radiation has a longer wavelength than the incident radiation. After trying unsuccessfully to explain this using the wave model, in 1922 he applied the particle model, completely explaining the phenomenon.

In 1929, Dirac's relativistic wave equation predicted the existence of the positron (an anti-electron). Carl Anderson discovered the positron in 1932. Particle-antiparticle pairs are created when energy, either in the form of electromagnetic radiation or kinetic energy, is converted into rest-mass energy. Conversely, when an ordinary particle comes together with its corresponding antiparticle a process called pair annihilation occurs where the rest mass of the two is converted into electromagnetic radiation.

Reluctantly, physicists have concluded that neither the wave model alone nor the particle model alone can account for all the phenomena associated with electromagnetic radiation. That is, there is no visualisible model for electromagnetic radiation. Because models help us to understand the world around us, we use the wave model when radiation behaves like a wave and the particle model when it behaves like a stream of particles. This approach is called wave-particle dualism.

Important concepts

Compton Effect; antimatter; positron; pair production; pair annihilation; wave-particle duality.

Questions

1. What is the Compton Effect?
2. What does it mean to say that the Compton Effect is discontinuous in space-time?
3. What is the wavelength of a photon with a momentum of 5.22×10^{-25} kg m/s?
4. In a Compton scattering event, the incident radiation has an energy of 0.055 MeV and the ejected electron has a kinetic energy of 340 eV. What is the wavelength of the scattered photon?
5. What is meant by pair production? By pair annihilation?
6. Why is the presence of a heavy nucleus necessary for pair production?
7. A photon of energy 2.22 MeV results in the production of an electron-positron pair. What is the sum of the kinetic energies of the particles?
8. What is the maximum wavelength of a photon that can produce an electron-positron pair?
9. List several phenomena that can only be explained by the wave model for electromagnetic radiation. List several phenomena that can only be explained by the particle model for electromagnetic radiation.

   The following question is of a more general nature. It has no single correct answer and is just something for you to think about. When possible, it is best answered in conversation with others.
10. What do you think Einstein meant when he said of wave-particle dualism "It is only with reluctance that one's desire for knowledge endures a dualism of this kind?"

11 CLASSICAL MODELS OF THE ATOM

The first expression of the concept of the atom is generally credited to the Greek philosopher Democritus around 400 BCE. Although matter appears to the senses as continuous, Democritus argued on philosophical grounds that it must be actually constructed of discrete units that are too small to be apparent. In Democritus' philosophy, the most important property of atoms was their elementary nature. An atom was an indivisible and indestructible unit. It had no internal structure. The word atom comes from the Greek word meaning "indivisible."

Democritus taught that "nothing exists except atoms and the void; all else is mere J opinion." Democritus' atomic model never gained wide acceptance in Greek thought and was rejected by Aristotle. During the Middle Ages, the Catholic church adopted the Aristotelian viewpoint and regarded atomic models of matter as contrary to church doctrine. It was not until the seventeenth century that the atomic model of matter gained any semblance of respectability. Galileo, Newton, and most of their contemporaries were atomists, although more for philosophical reasons than for scientific ones. There certainly was no experimental evidence available at the time to confirm or even suggest the existence of atoms.

11.1 DALTON'S INDIVISIBLE ATOMS

Newton postulated the existence of indivisible atoms as the ultimate building blocks of matter, but he offered no experimental evidence to support this assumption. Atomic theory made its first steps toward scientific respectability early in the nineteenth century when a quantitative atomic theory was proposed by John Dalton, an English chemist and schoolmaster. Dalton's ideas provided a theoretical basis for the science of chemistry, and they led to predictions concerning chemical reactions that later were verified. However, there was no direct physical evidence of the existence of atoms. No one was able to propose an experiment whose results would unequivocally confirm or deny the atomic model. Although most of the experimental data obtained by chemists were most easily explained by the atomic model, many other data were not well explained by it.

The atomic model was useful in forming a coherent and rational picture of many natural phenomena, so that by the late nineteenth century it was easier to believe in the existence of atoms than to deny their existence. Belief in the reality of the atom, however, was far from universal. Because of the lack of direct experimental evidence, some influential physicists and chemists remained skeptical, even at the beginning of the twentieth century.

In 1905, Einstein provided a quantitative explanation of the observed motions of suspended particles in a stationary fluid (a phenomenon called Brownian motion). Einstein's explanation was based on the assumption of the existence of real atoms of a definite (nonzero) size and mass. The quantitative predictions of Einstein's model were confirmed in experiments reported by the French physicist Jean Perrin. Since the publication of Perrin's results in 1908, physicist or chemist have accepted the validity of the atomic model of matter. For his work, Perrin received the Nobel Prize in physics in 1926.

Dalton's atomic model could be called the marble model. Dalton (much like Newton) viewed an atom as a hard sphere with no internal structure, indestructible and indivisible as the name implied. However, toward the end of the nineteenth century, it was becoming increasingly clear that this marble model cannot account for all the experimental data.

Many electrical experiments performed in the nineteenth century explored the effects that result when a very strong battery is connected across the two ends of a glass tube from which most of the air has been removed. These experiments showed that invisible rays are emitted from the negatively charged electrode, or cathode (see Figure 11.1). It was shown in 1869 that these cathode rays travel in straight lines. The next year they were shown to have both energy and momentum. In 1895, Jean Perrin demonstrated that the cathode rays carry negative charge. In 1897, J. J. Thomson (1856 – 1940) was able to show that this radiation consists of a stream of negatively charged particles, and that all of these particles are identical, no matter what the nature of the material from which they are emitted. He found that the mass of one of these particles is very much less than the mass of even the lightest atom. These particles came to be called electrons, and J. J. Thomson received the 1906 Nobel Prize in physics for his role in their discovery.

Cathode Rays

Cathode rays (electrons) are emitted by the negatively charged cathode and attracted toward the positively charged anode. Electrons missing the anode cross will fluoresce a screen at the end of the tube, leaving a shadow because the electrons could not pass through the cross.

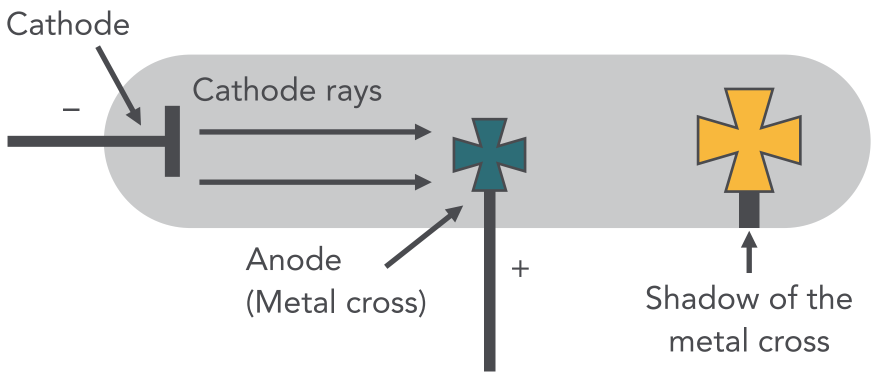


Figure 11.1 Cathode rays

By 1906 it seemed clear that every atom contains electrons as a part of the atomic structure. Dalton's marble model of the atom was no longer tenable, and the race was on to develop an atomic model that provided a logical place for the electron in some kind of atomic structure.

11.2 THE NUCLEAR MODEL OF THE ATOM

In 1909, the most widely accepted model of the atom was the one proposed by J. J. Thomson, He visualized the positive charge of the atom as being spread uniformly throughout a sphere about 10^{-10} m in diameter, with the electrons as smaller particles scattered through the atom. In early versions of this model, the electrons were assumed to be scattered randomly through the atom, like plums or raisins in a pudding. It is often called the plum-pudding model of the atom.

J.J Thomson
(1856 – 1940 * England)

1897 – descover that cathode rays, eventually called electrons , had masses much less that the lightest and surmised that they were internal constituents of atoms.

1904 – developed the plum-pudding model of the atom in which negatively charged electrons were imbedded in a sphere of positive charge which balanced the negative charge of the electrons.

1906 – Nobel Prize in Physics.

Figure 11.2 J. J. Thomson

However, calculations based on classical theories of electricity showed that the electrons in such an atom would arrange themselves in regular patterns or shells because of the repulsive electrical forces between electrons and the attractive electrical forces between the electrons and the "pudding". In a neutral atom, the total positive charge of the "pudding" would be balanced exactly by the negative charge of the electrons.

If a scientific model is to be useful, it must explain a substantial body of experimental data that already exists, and it also should generate predictions that can be tested experimentally. Thomson's plum-pudding model was consistent with the existence of electrons as discrete particles present in all atoms, but that was about all the model could accomplish. It did not predict any new effect that could be experimentally tested. The model did not lead to any better understanding of the nature of the atom.

In 1911, an alternative model of the atom was proposed by Ernest Rutherford (1871 – 1937). This nuclear model was more solidly based in experimental physics, and it quickly led to experimental tests that generated a great deal of information about atomic structure. Before discussing the nuclear model in detail, we will digress slightly to discuss the events that led Rutherford to develop the nuclear model of the atom.

In 1896, while the great French physicist Henri Becquerel was searching for X-ray emissions from a compound of uranium, he accidentally discovered strange radiations coming from the compound. This phenomenon later became known as radioactivity. Rutherford soon showed that this radiation is complex and includes at least two distinct types of radiation. One type, which he called alpha (α) rays, can be stopped by several layers of metal foil. A second type beta (β) rays, had much greater penetrating ability. Later, a third type called gamma (γ) rays was discovered, with the far greatest penetrating ability of the three types. (Alpha, beta, and gamma are simply the first three letters of the Greek alphabet. These names are the equivalent of calling the three types A, B, and C.) In 1903, Rutherford was able to determine that alpha rays consist of a stream of positively charged particles and was able to measure the charge-to-mass ratio for these alpha particles. Later he determined the charge of the alpha particle and concluded that an alpha particle is a doubly ionized helium atom; that is, a helium atom with two electrons removed.

After Rutherford became professor of physics at Manchester in 1907, he realized alpha particles could be used as probes to explore the structure of the atom. A stream of alpha particles will pass through a thin sheet of metal foil. If the alpha particles interact with the atoms in the foil, they should be deflected from their straight-line paths. Using the classical theories of mechanics and electromagnetism, it is possible to calculate the amount of deflection expected from a particular model of the metal atom in the foil. From rough calculations based on Thomson's plum-pudding model, Rutherford expected the alpha particles to be deflected only a few degrees from their straight-line paths when passing through a foil that was a few atoms in thickness

In 1908, Rutherford and his assistant Hans Geiger (1882 – 1945) who later invented the well-known Geiger counter, developed a method for counting individual alpha particles. Geiger then used this scintillation method of counting to examine the scattering of alpha particles passing through a thin metal foil. Geiger found that most of the alpha particles passed straight through the foil without measurable deflection. Some were scattered through very slight angles, but the number of scattered particles decreased very quickly as the angle of scattering increased. The largest scattering angles observed in these experiments were only a few degrees. These results seemed to be in good general agreement with the predictions based on the Thomson model of the atom.

Ernest Rutherford
(1871 – 1937 * New Zealand)

1900 to 1910 – made significant contributions to the study of radioactivity. Coined the terms alpha and beta radiation.

1911 – proposed the nuclear model of the atom.

1917 – the first to deliberately transmute one chemical element into another through induced nuclear reactions.

1908 – Nobel Prize in Chemistry.

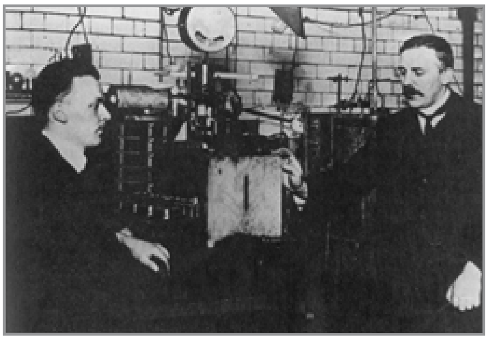
Figure 11.3 Ernest Rutherford

Like most other physicists of the day, Rutherford felt that Thomson's model was essentially correct, and he was not surprised when the few trial runs seemed to confirm predictions based on the model. Although the quantitative agreement with the predictions was not as perfect as might be desired, further refinement of the predictions and the experimental procedure seemed to be a rather tedious task that would simply result in a routine confirmation of expectations. Thus, Rutherford was not very interested in pursuing the matter. It seemed unlikely to yield any important new insights about the nature of the atom. Fortunately, however, Rutherford did regard this research as an appropriate assignment for his young research assistants. In 1909, an undergraduate named Ernest Marsden was brought in on the research. Rutherford later told the story this way.

> One day Geiger came to me and said, "Don't you think that young Marsden, whom I am training in radioactive methods, ought to begin a small research?" Now I had thought that too, so I said, "Why not let him see if any alpha particles can be scattered through a large angle?" I may tell you in confidence that I did not believe that there would be, since we knew that the alpha particle was a very fast massive particle, with a great deal of energy, and you could show that if the scattering was due to the accumulated effect of a number of small scatterings the chance of an alpha particle being scattered backwards was very small. Then I remember two or three days later Geiger coming to me in great excitement and saying, "We have been able to get some of the alpha particles coming backwards. It was quite the most incredible event that has ever happened to me in my life. 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.

Discovery of the Nuclear Atom

Using the nuclear model of the atom, Rutherford was able to determine from theory the alpha counting rate as a function of the scattering angel. Geiger's experimental results exactly matched the theoretical predictions. Their paper was published in 1913.



Rutherford and Hans Geiger in their Manchester Laboratory

Figure 11.4 Discovery of the Nuclear Atom

Rutherford immediately realized that the results of the Geiger-Marsden experiment could not be explained by the Thomson model of the atom, and he set his mind to the task of explaining these strange results. It took a long time, but he found an explanation. Geiger tells the story.

> One day [in l911] Rutherford, obviously in the best spirits, came into my room and told me that he now knew what the atom looked like and how to explain the large deflections of alpha particles. On the very same day I began an experiment to test the relations expected by Rutherford between the number of scattered particles and the angle of scattering.

Alpha Scattering

Most alphas pass through with little or no deflection, showing that the atom is mostly empty space. A few interact with the much more massive gold nucleus and are scattered through larger angles. A near head-on collision will result in back scattering.

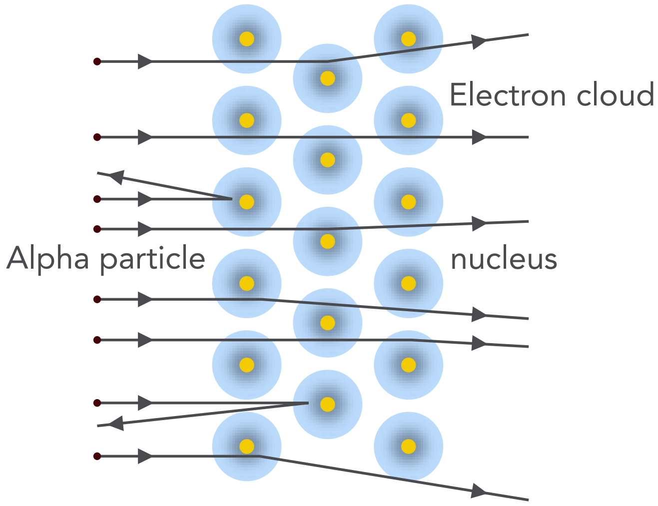


Figure 11.5 Alpha scattering

What Rutherford proposed was a nuclear model of the atom, with the positive charge and almost all of the mass of the atom concentrated in a tiny nucleus at the center of the atom, The diameter of the nucleus is only about 10^{-14} m or about one ten-thousandth of the diameter of the atom. (Rutherf0rd used the analogy of a fly in a cathedral") The electrons are located outside the nucleus, forming a sphere about $l0^{-10}$ m in diameter as the outer surface of the atom.

Rutherford's nuclear model provides a qualitative explanation for the Geiger-Marsden results. The atom is mostly empty space, so most of the alpha particles pass through the thin metal foil with little or no deflection. If an alpha particle does interact with one of the electrons in the metal atoms, it will experience a deflection so slight as to be very difficult to measure. The alpha particle is 8000 times more massive than an electron, so the effect of a collision with an electron is about like a marble colliding with a sand grain. However, an extremely small fraction of the alpha particles will happen to come close to the tiny nucleus in the center of a metal atom. This nucleus, in the case of the gold atoms being used in the foil of the Geiger experiments, is about 50 times more massive than an alpha particle. Therefore, it is quite reasonable to expect a small fraction of the alpha particles will be scattered through large angles, and even that a very few particles could undergo head-on collisions with nuclei and be scattered back in the direction from which they came. The effect of such a collision is rather like that of a high-speed marble running into a billiard ball.

Because of the structure of the nuclear atomic model, it was possible to calculate quantitative predictions for the exact results to be expected in alpha-scattering experiments. Rutherford predicted from the model the exact rate at which alpha particles should be deflected through each of the possible scattering angles. Geiger immediately began performing experiments to test these predictions. Because of the precision required, the final results were not published until 1913. These results are in excellent quantitative agreement with the predictions based on the nuclear model of the atom.

11.3 THE CLASSICAL PLANETARY MODEL OF THE ATOM

It is often the case in physics that the solution to one problem leads to the discovery of another problem of even more fundamental significance. Such as the case with Rutherford's nuclear model of the atom. The Geiger-Marsden experiments confirmed beyond any reasonable doubt the validity of the nuclear model. However, if all of the positive charge and almost all of the mass of the atom are concentrated in the tiny nucleus at the center, then where are the electrons located? It is clear that they cannot be at rest in the space outside the nucleus. The negatively charged electrons would be attracted toward the positively charged nucleus by the electrical force between the unlike charges. The electrons would swiftly fall into the nucleus, and the atom would collapse down to nuclear size. However, the experimental results indicate the volume of the atom is about 10^{12} times greater than the volume of the nucleus.

We know that attractive forces do not always lead to the collapse of a system of bodies, however. The earth and all the other planets are attracted toward the sun by the force of gravity, but the solar system does not collapse. It is the motion of the planets that prevents collapse. The gravitational force pulls at right angles to the motion of the planet, so the planet moves in an almost circular orbit around the sun. We might then imagine that each electron is in a stable orbit around the nucleus, in much the same way that a planet moves in a stable orbit around the sun. In this case, however, it is the attractive force between unlike charges rather than gravity that holds the atom together. This planetary model of the atom is particularly attractive because the equations representing the gravitational force and the electrical force between charged particles are of identical form. In both cases, the strength of the force between interacting objects is inversely proportional to the square of the distance between the objects. That is, F is proportional to l/r^2, where r is the distance between the centers of the masses in the gravitational force law or the distance between the centers of the charges in the electrical force law.

$F_g = G\ m_1 m_2/r^2$ where $G = 6.67 \times 10^{-11}\ N\ m^2/kg^2$. (Newton's law of gravity.)

$F_e = kq_1q_2/\ r^2$ where $k = 9 \times 10^9\ N\ m^2/C^2$ (Coulomb's law for the force between charges.)

Thus the motions of the electrons orbiting the nucleus should obey the same mathematical laws as the motions of the planets orbiting the sun. The analogy of the atom as a miniature solar system should be an almost exact one. Consider the planetary model of the simplest atom that exists in nature, the hydrogen atom. This atom consists of a single proton of charge +e and a single electron of charge –e. (The symbol e represents the fundamental quantity of charge, 1.6×10^{-19} C.) The proton is much more massive than the electron ($m_p = 1840\ m_e$), so the attractive force between them will change the motion of the electron much more than it changes the motion of the proton. We can think of the electron as being in orbit around a stationary proton (which is the nucleus in this case), just as the earth is in orbit around the much more massive sun.

If we use classical calculations to derive the energy of gravitational system of a planet and the sun, we get

$$Eg = -G\ m_p m_s/2r$$

A similar calculation of the energy of hydrogen atom

$$E = -\ k\ e^2/2r$$

The energy is negative because we have to add energy to the systems to separate the planet from the sun and also to separate the electron and proton in the atom. Since energy must be added to produce a state of separation, the energy of the bound systems is negative. As we discussed in an earlier chapter, the added energy increases the rest-mass energy of the system. In both cases the change in rest mass is immeasurably small relative to the rest mass of the system.

Experiments show that 13.6 eV of energy must be provided to separate the electron and proton of the hydrogen atom. If a sample of hydrogen s irradiated with electromagnetic radiation, no free electrons are produced until the photon energy exceeds 13.6 eV. Thus the total energy of the hydrogen atom is -13.6 eV. If we substitute this value into the classical equation for the energy of the hydrogen atom and solve for the radius, we obtain a value of $r = 5.3 \times 10^{-11}$ m = 0.53 Å, which is in excellent agreement with the experimentally determined value of the radius. At last we seem to be getting somewhere in our efforts to produce a quantitative model of the atom.

Because the form of the electrical force law (Coulomb's law) is mathematically the same as the form of the law of gravitation, classical Newtonian mechanics predicts that the electron should be able to move in a stable orbit around the nucleus. However, a difficulty arises when we apply Maxwell's theory of electromagnetism to this model. According to Maxwell's theory, any charged particle moving in a closed path must continuously emit electromagnetic radiation. Electromagnetic radiation is a form of energy, so Maxwell's theory indicates a planetary atom will continuously lose energy, causing the electron to spiral into the nucleus. Calculations show that a planetary hydrogen atom should completely collapse in about 10^{-6} s. Thus, according to classical theories of physics, the planetary model of the atom cannot explain the stability that is observed in nature.

A hydrogen atom does not normally emit electromagnetic radiation and it certainly exists for longer than 10^{-6} s. Something must be wrong with the model or with Maxwell's theory (or with both). In the next chapter, we will assume that the model is correct and modify the physical theories in such a way that we can retain the planetary model of the atom. When we treat quantum mechanics in a later chapter, we will find that it is ultimately necessary to modify both the model and the classical theories of physics.

The stability of the atom was not the only difficulty posed by the planetary model of the atom. Over the last few decades of the nineteenth century, a tremendous body of data was collected on the phenomenon of atomic spectra. Under certain circumstances, the individual atoms of a gas do emit electromagnetic radiation. However, atoms of a given element emit only certain frequencies (or wavelengths) of radiation. The pattern of frequencies emitted by a given kind of atom forms a distinctive "fingerprint" useful in identifying the element. There is a subtle mathematical regularity in the pattern of such an atomic spectrum, and this pattern must somehow be determined by the internal structure of the atom. Any satisfactory atomic model must provide both qualitative and quantitative explanations for atomic spectra. The planetary model of the atom does predict that an atom will emit electromagnetic radiation, but it predicts that the radiation will change its frequency continuously over the very brief lifetime of the collapsing atom. Instead, the atom is found to emit radiation only under certain circumstances, and then only at certain frequencies. The planetary model provides no explanation (either qualitative or quantitative) of atomic spectra.

In the next chapter, we will turn to a more detailed study of this problem and look at the first attempts of twentieth-century physicists to solve it. As you might already suspect, the best hope for a solution seemed to lie in the quantum hypothesis.

Summary

The concept that the material world is constructed from elementary building blocks is an ancient one. The Greek notion of small, indivisible, and indestructible atoms survived with very little modification to the end of the nineteenth century. In 1897, J. J. Thomson discovered the electron and concluded that it is part of the internal structure of the atoms of all elements. By 1913, experimental evidence clearly indicated that an atom is for the most part empty space; almost all of the mass of an atom is concentrated in the nucleus, an incredibly small volume at the center of the atom. The negatively charged electrons are bound to the positively charged nucleus by attractive electrical forces. The distribution of the electrons in space determines the size of the atom. This nuclear model of the atom was proposed by Ernest Rutherford and was firmly established by the results of the Geiger-Marsden alpha-scattering experiments. It is clear that the electrical force within the atom does not pull the electrons into the nucleus. This was explained by assuming that the electrons move in orbits around the nucleus, much as the planets move in orbits around the sun under the attractive force of gravity. This classical planetary model of the atom is made plausible by the equivalence of the mathematical forms of the force laws for the electrical and gravitational forces. Using the classical planetary model, we can calculate the energy for the hydrogen atom. We find that the energy of the hydrogen atom depends only on the radius of the electron orbit around the nucleus. This result is extremely encouraging because, when the measured energy of the hydrogen atom is inserted in the equation, the computed radius is equal to the experimentally determined radius of the hydrogen atom.

However, a problem arises when Maxwell's theory of electromagnetism is applied to the classical planetary model of the atom. According to Maxwell's theory, the electron should continuously emit electromagnetic radiation as it orbits the nucleus. This loss of energy from the atom should cause the electron to spiral down in a path of continuously decreasing radius, plunging it into the nucleus after a lifetime of only about 10^{-6} s. Thus application of all theories of classical physics to the planetary model of the atom indicates that such an atom is unstable, which is clearly inconsistent with experimental observation. We are forced to conclude that either the planetary model of the atom or the classical theories of physics (or both) is wrong.

Important concepts

Atom; electron; radioactivity; alpha rays; beta rays; gamma rays; nuclear model of the atom; nucleus; planetary model of the atom; total energy of the hydrogen atom.

Questions

1. Explain the origin of the word atom and its original meaning.
2. What are cathode rays?
3. What important discovery led to the downfall of the Dalton model of the atom?

4. Compare and contrast the models of the atom proposed by Dalton and by Thomson.
5. Determine the fraction of the volume of the atom that is occupied by the nucleus. Roughly, what percentage of the mass of the atom is contained in this volume?
6. Compare and contrast the models of the atom proposed by Thomson and by Rutherford.
7. Why are alpha particles scattered by atoms?
8. Explain the details of the Geiger-Marsden experiment. What was Rutherford so surprised about the in the initial results?
9. In the equation for the total energy of the hydrogen atom, the value of the constant k is 9×10^9 N m^2/C^2. The radius of the hydrogen atom is $r = 5.3 \times 10^{-11}$ m. What is the total energy of the hydrogen atom in joules? What is the total energy in electron-volts?
10. In the universal law of gravitation, the constant $G = 6.67 \times 10^{-11}$ N m^2/kg^2. The mass of the electron is 9.11×10^{-31} kg, and the mass of the proton is 1.67×10^{-27} kg. Find the gravitational force that the electron and proton exert on each other in the hydrogen atom. Using data from question 9, find the electrical force that the electron and proton exert on each other in the hydrogen atom. What is the ratio of the electrical force to the gravitational force?
11. A hydrogen atom absorbs a photon with 17.5 eV of energy. What will be the energy of the separated electron and proton?
12. Outline the problems posed for the planetary model of the hydrogen atom by the experimental evidence.
13. Although the sun attracts the earth by a gravitational force, the earth does not spiral in toward the sun. Explain why this is so, using classical theories of physics. If the classical theories lead to an explanation of the stable orbit of the earth, why do the same theories not explain the stable orbit of the electron in the planetary model of the hydrogen atom?

    The following question is of a more general nature. It has no single correct answer and is just something for you to think about. When possible, it is best answered in conversation with others.
14. The models of an internal structure for the atom were quite pleasing to many physicists, even though the idea of indivisible atoms had endured for so long. Instead of dozens of different kinds of atoms, the nature of matter could now be explained in terms of a few basic kinds of particles such as electrons and protons, combined in various structures. Discuss the reasons why such a model should seem more attractive. How do our preconceptions about simplicity and beauty affect our scientific theories about the nature of the universe?

Ernest Rutherford (1871 – 1937)

While earning his B.A. and M.A. at the University of New Zealand, Ernest Rutherford distinguished himself in both mathematics and physical science. He did experimental work on electromagnetic waves, and he was able to transmit and receive radio waves over a distance of two miles. This work was done well before the more famous inventions of Guglielmo Marconi in Italy, but Rutherford had no interest in the practical applications of the effect. The research did win him a scholarship to Cambridge University, however, although only after the first winner chosen refused the honor for family reasons. In 1895, Rutherford arrived at the Cavendish Laboratory to work as a research student under J. J. Thomson. Rutherford soon became well known as a bright young researcher, expanding his interests into work with X-rays and radioactivity. In 1898, he accepted the offer of a professorship at McGill University in Montreal, Canada, where he established a center for research in radioactivity. His reputation was already great enough to attract many bright young students to work as assistants there.

During the nine years he spent in Montreal, Rutherford performed many important experiments demonstrating the properties of the radiation emitted from radioactive substances. He worked with Frederick Soddy to find an explanation for the events involved in radioactivity, showing that an atom of uranium gradually changes through a series of other elements as it gives off radioactive emissions. He began a series of experiments to study the way that alpha particles are scattered by a thin sheet of metal foil.

In 1907, Rutherford returned to England to become professor of physics at the University of Manchester. There he and his students Hans Geiger and Ernest Marsden carried out the experiments that led to the development of the nuclear model of the atom. For all this important work, Rutherford received the 1908 Nobel Prize in chemistry, and he was knighted in 1914. Meanwhile, he was carrying out a series of experiments in which he devised various ways to make quantitative measurements of radioactive phenomena. In 1919, he succeeded J. J. Thomson as head of the Cavendish Laboratory. There he performed experiments in which a beam of alpha particles was used to knock protons out of the nuclei of atoms, thus for the first time in history creating an artificial nuclear reaction in which one element is converted into another. However, only about l in 300,000 of the incident alpha particles would undergo such a collision with a nucleus, so the process had no apparent practical application. Rutherford was skeptical that such artificial nuclear reactions could ever be put to any significant use, and he was to die just a few years before the age of nuclear bombs and reactors began.

Although Rutherford was one of the greatest experimental physicists who has ever lived, he by no means devoted all of his energies to physics. He was a voracious reader of all types of books. He enjoyed bridge, golf, and just about any type of open-air activity. In 1910, he bought an automobile and spent many hours driving with his family around the countryside.

He received many honors for his scientific work. In 1930, he was made Lord Rutherford, Baron of Nelson. He was noted at the Cavendish for his boisterous manner and booming voice. Other researchers complained that his talking interfered with the delicate apparatus, and they posted a large sign reminding him to “talk softly please.” He was never known for false modesty. When an envious colleague remarked of his honors that he was “lucky to be riding the crest of a wave,” Rutherford replied (with some justification), “Lucky, nothing! I made the wave.

12 THE BOHR MODEL OF THE ATOM

The planetary model of the atom comes tantalizingly close to providing an explanation for known properties of the hydrogen atom, but it suffers from one fatal flaw. According to Maxwell's theory of electromagnetism, a hydrogen atom should collapse within about a millionth of a second, emitting a burst of radiation of constantly changing frequency. In reality, of course, the hydrogen atom is a stable system that can exist indefinitely. However, atoms do emit electromagnetic radiation under certain circumstances, but the nature of this radiation is not consistent with expectations from the planetary model. We look now at the radiation emitted by atoms, and then turn to a surprising modification of the planetary model that was proposed to account for these observations. Once again, the quantum hypothesis proves to be useful in trying to reconcile classical theories with experimental results.

12.1 ATOMIC SPECTRA

An electric discharge tube contains a small quantity of some particular element in gaseous form. If an electric current is passed through the tube, the atoms of the gas emit electromagnetic radiation that is characteristic of the element. The emitted radiation contains only certain frequencies (or wavelengths) scattered through the infrared, visible, and ultraviolet regions of the electromagnetic spectrum. The pattern of particular wavelengths is different for each element.

A common example of an electrical discharge tube is a neon sign; its characteristic red-orange light is due to emission of light at particular wavelengths in the visible part of the spectrum. Another example is the mercury-vapor light whose bluish-green light now commonly lights parking areas and highways.

Everyone is familiar with the ability of a glass prism to separate a beam of white light into a rainbow of colors. The white light contains a range of wavelengths across the entire visible spectrum. As the light passes through the prism, light of the longest wavelength (red light) is bent through the smallest angle, whereas light of the shortest wavelength (violet light) is bent through the greatest angle (see Figure 12.1). The emerging light forms a continuous sequence of wavelengths from red to violet. Such a spectrum is called a continuous spectrum. The rainbow that often follows a rainstorm is created when sunlight passes through water drops in the air; the many individual drops act much like a prism to separate the wavelengths.

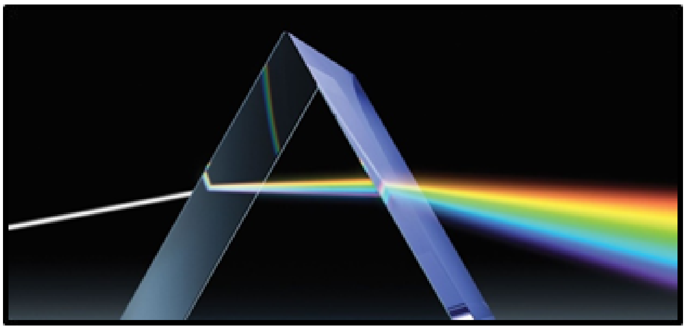


Figure 12.1 The effect of a prism on light.

If the light from an electric-discharge tube is passed through a prism, this light is also spread out on the basis of wavelength. The resulting spectrum, however, is markedly different from the familiar rainbow. Rather than a continuous spectrum, the light forms a discrete spectrum, or line spectrum. Only certain wavelengths are present in the light, so they form separate and distinct lines at certain points in the spectrum. Most of the wavelengths of the rainbow are absent from the line spectrum. This type of spectrum is called a bright-line spectrum, or emission spectrum (see Figures 12.2). Emission spectra are often called characteristic spectra, because the positions and intensities of the lines are characteristic of the gas in the discharge tube.

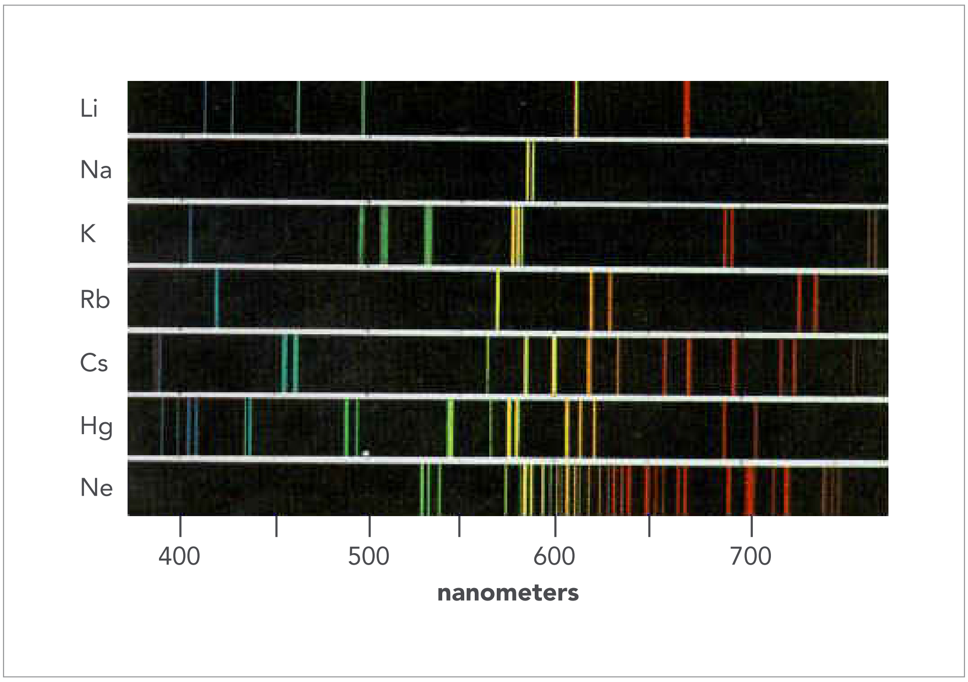


Figure 12.2 Bright-line spectra

Another useful way to represent a spectrum is by a graph that plots intensity of the radiation as a function of wavelength (or frequency). Figure 12.3 shows such a plot of the bright-line spectra of several elements.

Atomic Emission Spectra

As opposed to continuous thermal emission spectra, emission from individual atoms is discrete. Only certain wavelengths are emitted and the wavelengths and intensities are characteristic of the chemical element. The figure show emission spectra for hydrogen, gold (AU), aluminum (Al), and silver (Ag).

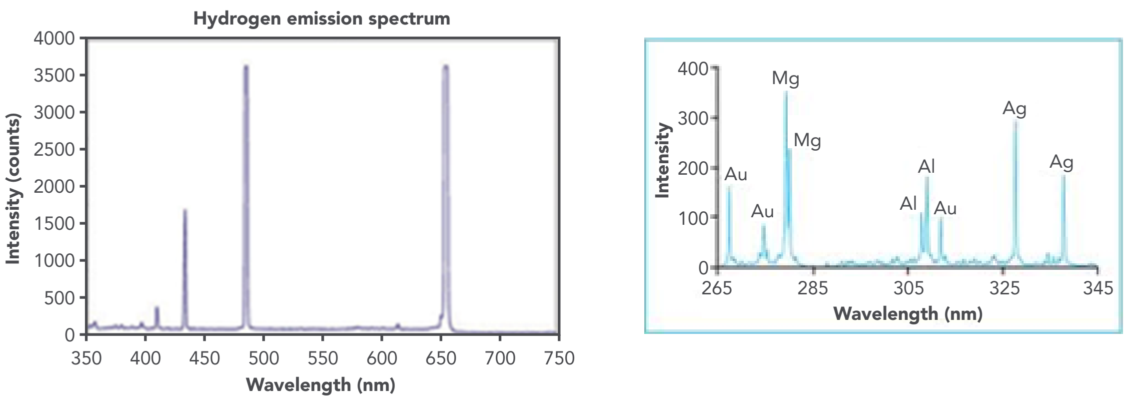


Figure 12.3 Bright-line spectra

A related phenomenon occurs when white light is passed through a tube that contains a particular element in gaseous form. When the emerging light is then passed through a prism to form a spectrum, dark lines are found in the continuous spectrum of the white light at positions corresponding to the positions of the bright lines in the emission spectrum of the same gas. Apparently, the gas absorbs light at the same wavelengths as it emits light in the electric-discharge tube. Figure 12.4 shows a dark-line spectrum, or an absorption spectrum.

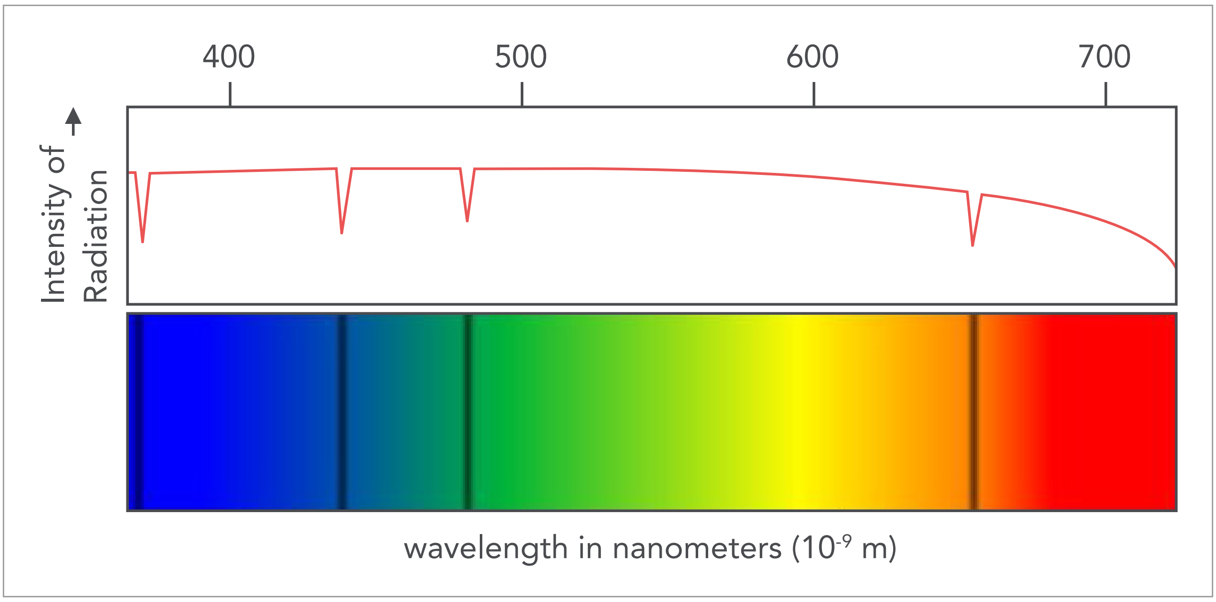


Figure 12.4 A dark-line spectrum.

The wavelengths corresponding to the lines in emission and absorption spectra can be measured very precisely. Handbooks of physical data give the wavelengths of these lines to six or more significant figures. The spectra of various gases had been quite well measured by the late nineteenth century, but no one had discovered any regular pattern among these data. In 1885, a sixty-year-old mathematics teacher at a Swiss girl's school, Johann Balmer, discovered a simple mathematical relationship among the wavelengths of the lines in the visible portion of the hydrogen emission spectrum. The Swedish physicist Johannes Rydberg later expressed this relationship in a more general form:

$$1/\lambda = R\ [(1/2^2) - (1/n^2)]$$

where λ is the wavelength of the visible light emitted, n is a small integer greater than 2, and the Rydberg constant R equals 1.097×10^{-3} Å^{-1} . (The unit Å^{-1}, reads inverse angstroms and is equal to l/ Å. With R expressed in Å^{-1}, the equation yields a value for λ in angstroms.)

If we put the values n = 3, 4, 5, or 6 into this equation, we obtain wavelengths that correspond exactly to those measured experimentally for the bright lines in the visible4spectrum of hydrogen. The four wavelengths obtained correspond to the four lines shown in the hydrogen spectrum in Figure12.3. Balmer's relationship was obtained simply by seeking some mathematical regularity among the experimentally determined wavelengths. Neither Balmer nor Rydberg could offer any explanation or model to explain why this equation should be valid. There seems to be no connection between this equation and any theory or physical model for the emission process. Thus the equation was not of much immediate use to physicists. For some time, it remained simply an interesting curiosity. However, any relationship that fits experimental data as accurately as this one does is very likely to contain some hidden physical significance. Later physicists did discover a model of the atom that provides a physical explanation for the validity of the Balmer-Rydberg equation.

Example 12.1

Calculate the wavelength of the bright line in the visible spectrum of hydrogen that is associated with the value n = 3 in the Balmer-Rydberg equation. What color is this particular radiation?

Solution

From the Balmer-Rydberg equation,

$$1/\lambda = R\,[(1/2^2) - (1/n^2)] = 1.097 \times 10^{-3}\ \text{Å}^{-1}\,[(1/2^2) - (1/3^2)]$$

$$1/\lambda = 1.097 \times 10^{-3}\ \text{Å}^{-1}\,[(1/4) - (1/9)] = 1.52 \times 10^{-4}\ \text{Å}^{-1}$$

$$\text{or } \lambda = 6563\ \text{Å}.$$

The wavelength lies in the red region of the visible spectrum (see Figure 8.4).

Matter in the solid form produces a continuous spectrum of electromagnetic radiation. The continuous spectrum emitted by an incandescent lamp is formed by emission from the hot metal filament of the lamp. The intensities of various parts of the continuous spectrum depend only on the temperature of the emitting solid and not on the elements present in the solid. Liquids and high-density gases also emit a continuous spectrum. Emission spectra are produced by matter in the form of a low-density gas and absorption spectra are produced when a continuous spectrum passes through a low density gas. Because the emission and absorption spectra are characteristic of the particular elements of the gas involved in their formation, they are often called atomic spectra. It is the study of atomic spectra in the light from astronomical bodies that enabled physicists to identify the elements present in the sun and stars and even in the atmospheres of planets. These spectra played an important role in the search for a satisfactory model of the atom.

12.2 QUANTIZING THE PLANETARY MODEL OF THE ATOM

In the year 1911, the year that Rutherford proposed the nuclear model of the atom, Niels Bohr (1885 – 1962), a young Danish physicist obtained his Ph.D. and began a one-year fellowship under J. J. Thomson. Bohr's training had been strongly theoretical, and he brought to England with him a thorough knowledge of the quantum hypotheses of Planck and Einstein. In England, quantum theory was less popular than it was on the continent. This was particularly true in Thomson's laboratory, where much time was being invested in an effort to explain atomic spectra by applying the physics of Newton and Maxwell to Thomson's "plum-pudding" model of the atom.

In early 1912, Rutherford came to Cambridge to visit Thomson. Rutherford took an immediate liking to young Bohr and invited him to visit Manchester. Bohr was delighted to accept because he had been frustrated in his efforts to interest Thomson in his ideas. They must have been a strange pair, the boisterous and booming Rutherford and the shy young Bohr whose voice was almost a whisper, with pauses of many minutes in his conversation as he struggled to find just the right words to express his ideas. But they did communicate and each was impressed by the other's ideas. Bohr returned to Thomson's laboratory convinced that Rutherford was on the right track with his nuclear model of the atom. This was the last straw as far as Thomson was concerned. He suggested that it might be best for everyone if Bohr completed his fellowship at Manchester under Rutherford.

Niels Bohr
(1885 – 1962 * Denmark)

1913 – developed Bohr model of the hydrogen atom, quantizing the energy states and explaining atomic spectra.

1920s – established the Institute for Theoretical Physics in Copenhagen, Denmark, the leading mentor of young physicist developing quantum physics.

1927 – introduced the Copenhagen interpretation of quantum mechanics.

1922 – Nobel Prize in Physics.

Figure 12.5 Niels Bohr

Over the next few months at Manchester, Bohr began working out a new model of the atom, based on Rutherford's nuclear model but with very significant modifications based upon the ideas of quantum theory. Unlike Thomson, Rutherford provided constant support and encouragement for Bohr's efforts to challenge existing models of the atom. Bohr focused his attention on the simplest atom – the hydrogen atom.

Bohr's model of the hydrogen atom is based on four postulates. He accepted the planetary model of the atom, but this raised the difficulty of the atomic collapse predicted by Maxwell's theory. In his first postulate, Bohr boldly assumed (on the basis of the experimental fact that atoms do not collapse) that Maxwell's theory does not apply to electrons orbiting the nucleus of an atom.

Postulate I. An electron orbiting the nucleus of an atom does not emit electromagnetic radiation.

With this postulate, he could now use the planetary model of the atom worked out in Chapter 11. This model gives a value for the total energy of the hydrogen atom of

$$E = -ke^2/2r$$

where r is the radius of the electron's orbit, e is the basic quantity of charge, and k is the constant from Coulomb's law. Because of postulate I, he no longer expected the atom to collapse. However, he still had to explain the existence of discrete emission and absorption spectra. Bohr made two additional assumptions to fit the model to these facts.

Postulate II. The energy of the atom is quantized. That is, only certain values of the total energy E are allowed to exist.

Postulate III. Energy is emitted or absorbed by an atom when and only when it makes a transition from one allowed energy state to another.

Bohr called the allowed energy states of postulate II "stationary states" in order to stress that no radiation is emitted when the atom is in one of these allowed states. Because the planetary model of the atom indicates that the energy depends solely on the radius of the electron's orbit, each allowed stationary state must correspond to an allowed value of the radius of the orbit Thus, postulate II is equivalent to the assumption that only certain values of the radius are allowed, and all other values of the radius are forbidden. Because of the negative sign in the energy equation, the larger orbits correspond to higher atomic energies – that is, energies closer to zero. When the electron is in one of the allowed orbits (that is, the atom is in one of the allowed energy states), then no electromagnetic energy is emitted (postulate I).

If energy is absorbed by the atom, then the atom must jump to a higher energy state -- that is, the radius of the orbit must jump to one of the larger allowed values. Similarly, if energy is emitted by the atom, then the radius of the orbit must jump to a lower allowed value. The amount of energy emitted or absorbed must correspond to the difference between two allowed energy states. This means that only certain amounts of energy can be emitted or absorbed. (If we accept the Planck-Einstein quantum hypothesis, then these packets of energy should correspond to certain wavelengths of radiation, and we have an explanation for the discrete atomic spectra.)

The next task is to seek a quantitative match between the model and the experimental data. From the quantum hypothesis, the energy E that is emitted or absorbed must represent a quantum of radiation of frequency ν, where $E = h\nu$. If E_u is the energy of the greater (upper) energy state and E_l is the energy of the lesser (lower) energy state, then the energy of the photon emitted or absorbed must be $E = h\nu = E_u - E_l$.

If the transition is from the upper state to the lower state, then the photon of energy $h\nu$ will be emitted. If the transition is from the lower to the upper state, then the photon must be absorbed. (Like most physicists at the time, Bohr did not believe that electromagnetic radiation actually exists in the form of discrete photons. However, we will discuss his theory in more modern terms and make use of the handy concept of photons.) It is clear that the assumption of quantized energy states for the atom can lead to a set of discrete frequencies (or wavelengths, $\lambda = c/\nu$) for the radiation of the atomic spectra. Because the allowed energy values depend on the element, the atomic spectra are unique to each element. We can write

$$h\nu = hc/\lambda = E_u - E_l$$

In terms of wavelength λ, then, we have

$$1/\lambda = (1/hc)(E_u - E_l)$$

where E_u and E_l are allowed energy states of the atom, and λ is some wavelength present in the atomic spectrum.

The problem now is to find a rule for the allowed orbits that will lead to just the right allowed energy states to explain the experimentally observed wavelengths in the atomic spectra. Before Bohr had time to carry his work any farther, his fellowship in England ran out and he returned to Copenhagen to marry and assume a teaching position. In January 1913 a fortuitous event occurred. Bohr was explaining his ideas to an old classmate, who asked if this model might not explain the Balmer-Rydberg equation. Bohr was not familiar with this equation (which was simply a curious mathematical regularity, mostly known only to the study of atomic spectra). His friend suggested that he look it up. Bohr later recalled, "As soon as I saw Balmer's formula, the whole thing was immediately clear to me." Compare the Balmer-Rydberg equation to the equation we have just derived for $1/\lambda$ from Bohr's model. From the similarities between the two equations, Bohr was able to work out the necessary rule for the allowed orbits of the electron in the hydrogen atom. We will begin with Bohr's rule and work back to the Balmer-Rydberg equation.

> **Postulate IV. The angular momentum of the electron in its orbit must be an integer multiple of Planck's constant divided by 2π.**

The angular momentum of an object in a circular orbit is the product of its linear momentum mv and the radius r of the orbit. Thus, Bohr's postulate can be written as

$$m_e vr = nh/2\pi \text{ for } n = 1, 2, 3, \ldots\ldots$$

where m_e is the mass of the electron, and n can be any positive integer. Note that this postulate assumes that the angular momentum of the electron is quantized; it can have only certain discrete values. Such a quantization of angular momentum is not predicted by the laws of classical mechanics. In fact, each of Bohr's postulates involves an assumption that contradicts predictions based on classical theories of physics.

Using Newton's classical mechanics together with his fourth postulate, Bohr was able to show that

$$E_n = - (1/n^2)(2p^2 m_e k^2 e^4/h^2)$$

where E_n is the energy of the stationary state represented by the value n. E_1 is the value of the lowest energy state, E_2 the energy of the next highest state and so on. The expression in parentheses includes only constants and the combination has the unit of energy. Evaluating this term and expressing the result in electrons-volts, we get

$$E_n = - (1/n^2)(13.6 \text{ eV})$$

where E_n is the energy of the hydrogen atom when the electron has angular momentum $nh/2\pi$ and n is a positive integer.

The lowest possible energy for the hydrogen atom is E_1 = -13.6 eV. This energy state is called the ground state, and the lowest possible energy for the atom is called the ground-state energy. For larger values of n, the energy E_n is larger (that is, less negative), and the energy approaches zero as n approaches infinity. These higher-energy states are called excited states. An atom generally spends very little time in any excited state (typically less than 10^{-8} s) before it makes a spontaneous transition to a lower state by emitting a photon of the appropriate energy It is these transitions from higher energy states to lower ones that are responsible for the emission spectra. The spontaneous transition can occur to any lower state. Several transitions may occur before the atom returns to the ground state.

Figures 12.6 and 12.7 are diagrams of the allowed energy states for the hydrogen atom.

Bohr Model of the Hydrogen Atom

Energy level diagram for the hydrogen atom. The n=1 level is the lowest allowed energy and the smallest orbital radius. Transitions to this level produce ultraviolet photons. Transitions from higher energies to the n=2 level produce photons in the visible region of the spectrum. Transitions form higher energies to the n=3 level produce infrared photons. Transitions from lower to higher energy levels occur when the hydrogen atom absorbs a photon of the exact energy needed to transition to a higher energy level.

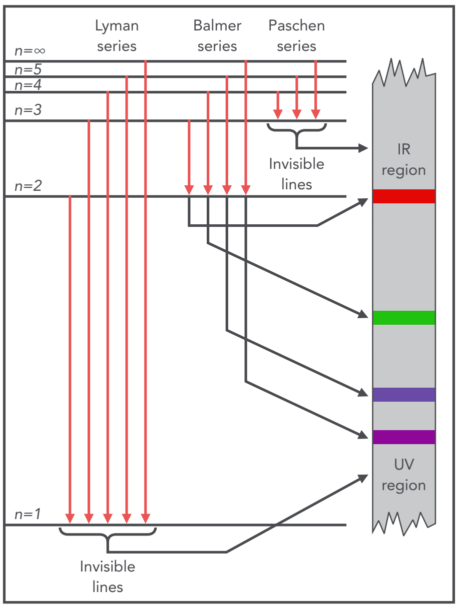


Figure 12.6 Energy-level diagram for hydrogen

By absorbing the appropriate amount of energy, an atom in the ground state can make a transition to an excited state. If this absorbed energy is sufficient (greater than 13.6 eV for the hydrogen atom), the electron can be completely removed from the atom. An atom from which one or more electrons have been removed is said to be ionized. The minimum energy required to ionize the atom is called the ionization energy. The ionization energy for hydrogen is 13.6 eV.

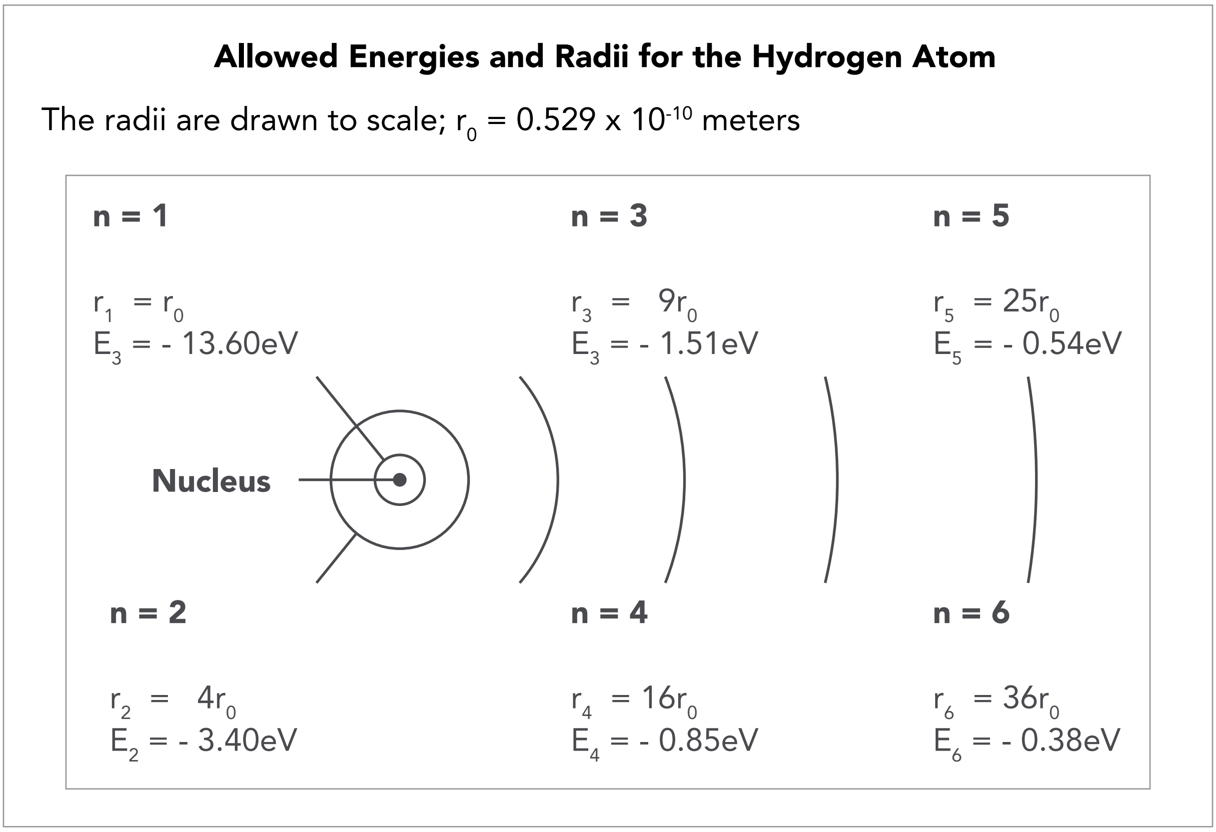


Figure 12.7 Allowed hydrogen atom radii.

In the electric-discharge tube, high-speed electrons collide with the atoms of the gas in the tube, transferring enough energy to individual atoms to either ionize or excite them. Emission then occurs as excited atoms drop in one or more steps back to the ground state. Another way to excite atoms is to heat them to high temperatures. If the atoms have sufficiently large kinetic energy when they collide with each other, ionization or excitation can result. A third method of excitation is to irradiate the atoms with a continuous spectrum of light. Those photons that correspond to allowed transitions will be absorbed; the others will not. This is how a dark-line, absorption spectrum is produced. Because the transitions are the same for absorbing or for emitting, the wavelengths of the bright lines in the emission spectrum are the same as those of the dark lines in the absorption spectrum.

Figure 12.6 shows the transitions from higher excited states to the states with n = l, n = 2, and n = 3. Consider the transitions to the state n = 2. In this case n = 2 is the lower energy state and the upper states correspond to n = 3, 4, 5, …. in Bohr's relationship,

$$1 / \lambda = (1/hc)(E_u - E_l) = (1/hc)(E_u - E_2) \qquad \text{for } n = 3, 4, 5, \ldots.$$

Using the equation for E_n derived earlier, we have

$$1 / \lambda = (1/hc) \times (13.6 \text{ eV}) \times (-1/n^2 + 1/2^2) \text{ so,}$$

$$1 / \lambda = (13.6 \text{ eV}/hc) \times (1/4 - 1/n^2) \qquad \text{for } n = 3, 4, 5,$$

Evaluating the constants, we find that (13.6 eV)/hc = 1.097 x 10-3 $Å^{-1}$, which is exactly the experimentally determined value of the Rydberg constant R in the Balmer-Rydberg equation. Thus the postulates of Bohr's theory lead quite naturally to the derivation of the Balmer-Rydberg equation. The Bohr model of the hydrogen atom, although it requires some drastic assumptions that conflict with classical theories, does provide a quite satisfying match with the experimental data. This was a tremendous success for the Bohr model, providing convincing evidence that he was on the right track in applying quantum hypotheses to the structure of the atom.

The series of transitions from higher states to the state n = 2 is now called the Balmer series. As we have seen, this series corresponds to the bright lines in the visible portion of the emission spectrum of hydrogen. We can easily calculate the wavelengths expected for transitions from higher states to the ground state (with n = 1). When we carry out the calculations, we find that these transitions correspond to wavelengths in the ultraviolet portion of the electromagnetic spectrum (that is, wavelengths less than 3500 Å). This series of transitions corresponds exactly with the observed wavelengths of lines in the ultraviolet portion of the hydrogen emission spectrum; it is called the Lyman series.

Similarly, transitions to the n = 3 state result in emission of infrared light with wavelength greater than 7500 K, and the corresponding series of lines does exist in the hydrogen spectrum (it is called the Paschen series). Transitions to the states with n greater than three also result in emission of infrared radiation. Bohr's model was completely successful in accounting for the general features of the hydrogen emission spectrum and other physicists certainly were impressed with his accomplishment. However, they were very disturbed by some of the implications of Bohr's postulates.

According to Bohr's model of the atom, the transitions between energy states are discontinuous events. The atom ceases to exist in one allowed energy state and begins to exist in another, but the transition from one state (one orbit) to the other cannot be followed continuously in space-time. We cannot represent the transition between energy states as a continuous motion of the electron through space-time from one orbit to another. We must simply describe the initial and final states and forget about any possibility of describing the transition itself. To emphasize the discontinuous nature of the transitions, Bohr called them "quantum jumps." Nearly all physicists at the time (and even some today) were disturbed by this concept. For example, Rutherford initially reacted to Bohr's model by commenting:

> There appears to me one grave difficulty with your hypothesis, which I have no doubt you fully realize, namely, how does an electron decide what frequency it is going to vibrate at when it passes from one stationary state to the other? It seems to me that you would have to assume that the electron knows beforehand where it is going to stop.

In other words, the electron ceases to exist in one orbit and appears in another, as a photon of energy is simultaneously emitted. There is no continuity in space-time between these events, and yet the electron somehow makes the proper jump that corresponds to the energy of the emitted photon. Any attempts to explain these events in terms of classical ideas of cause and effect lead immediately to contradictions. For example, it is necessary to postulate that the electron cannot exist except in allowed orbits. The jump must be discontinuous so that the quantum of energy (the photon) can be emitted as a single entity. But then, how does the electron "know" that it must emit a certain amount of energy in order to arrive at another allowed orbit?

Acceptance of the quantized model of the atom seems to lead to severe problems in trying to picture how things work at the atomic level. All of our classical notions about continuity of events and about cause and effect are knocked askew by the postulates of Bohr's model. The debate among physicists about these matters was to last for decades and it led in the end to the development of a new theory of quantum mechanics that has even more startling implications about the ultimate nature of reality. We will return to this topic in the next chapter.

Example 12.2

What is the wavelength and the frequency of the radiation emitted in the transition from the state n = 5 to the state n = 3 of hydrogen?

Solution

For the hydrogen atom, the energy of state n is given by

$$E_n = - (1/n^2)(13.6\ \text{eV})$$

For a hydrogen atom in the n = 3 state

$$E_3 = - (1/9)(13.6\ \text{eV}) = - 1.51\ \text{eV}.$$

For the n = 5 state,

$$E_5 = - (1/25)(13.6\ \text{eV}) = - 0{,}54.\text{eV}$$

$$\text{So } E_u - E_l = (- 0.54\ \text{eV}) - (-1.51\ \text{eV}) = 0.97\ \text{eV}.$$

The energy of the atom decreased as it moves from the n = 5 state to the n = 3 state. This energy is emitted as electromagnetic radiation of energy $h\nu = hc/\lambda = 0.97$ eV.

$$\lambda = hc/\ 0.97\ \text{eV} = 1.24 \times 10^4\ \text{eV Å} /0.97\ \text{eV} = 12{,}800\ \text{Å}.$$

As expected, this line in the Paschen series lies in the infrared region of the spectrum.

Example 12.3

As it makes the transition from some higher state to the state n = 2, a hydrogen atom emits a photon of wavelength 4336 Å. What was the energy of the hydrogen atom before the transition occurs?

Solution

The energy of the emitted photon is

$$E = hc/\lambda = 1.24 \times 10^4 \text{ eV Å} / 4336 \text{ Å} = 2.86 \text{ eV}$$

$$2.86 \text{ eV} = E_u - E_l = E_n - E_2$$

where E_n is the energy of the initial state.

$$2.86 \text{ eV} = [(-1/n^2)(13.6 \text{ eV}) - (-1/2^2)(13.6 \text{ eV})] = [(-1/n^2)(13.6 \text{ eV}) + 3.4 \text{ eV}]$$

$$(1/n^2)(13.6 \text{ eV}) = 3.4 \text{ eV} - 2.86 \text{ eV} = 0.04 \text{ eV or } n^2 = 13.6 \text{ eV} / 0.04 \text{ eV} = 25$$

So n = 5 for the initial state.

12.3 DIFFICULTIES WITH THE BOHR MODEL

Although it led to numerous successes, the Bohr model also presented some difficulties. For emission spectra, both the wavelength and intensity of the lines are characteristic for a given emitting substance. Bohr's model led to excellent predictions of the wavelengths but could not explain the intensities. Furthermore, Bohr's model could be applied successfully only to one-electron atoms such as the hydrogen atom or other light atoms with all but one electron removed. And even this success was limited. As spectroscopic measurements improved, single lines in the spectrum of hydrogen proved to be actually groups of closely spaced lines. Bohr's model could offer no explanation for this phenomenon. Perhaps the most dissatisfying aspect of the Bohr model was the fact that it was neither a classical model nor a quantum model, but rather a combination of the two. Bohr used the classical physics of Newton and Maxwell to discuss the electron in its allowed orbits and to calculate the energy associated with these states. On the other hand, his assumptions of the existence of stationary states, the mechanism for emission and absorption of light by the atom, and the quantization of angular momentum are strictly non-classical. Even Rutherford, whose case for a nuclear model of the atom was strengthened by Bohr's quantized version, had some concern on this point. In a letter, he wrote to Bohr:

> Your ideas as to the mode of origin of the spectra in hydrogen are very ingenious and seem to work out very well; but the mixture of Planck's ideas and the old mechanics makes it very difficult to form a physical idea of what is the basis of it all.
> Quoted in Helge Kargh, *Niels Bohr and the Quantum Atom.*

It was clear that classical physics could never offer a satisfactory model of the atom. It was beginning to become clear that what was needed was a complete break with classical theories -- the development of a new quantum theory of mechanics that would permit a completely quantum-based model of the atom. In 1925; Erwin Schrodinger and Werner Heisenberg independently proposed just such a completely quantum-mechanical treatment of the hydrogen atom. This treatment provided quantitative explanations for all that the Bohr model could explain, and it could be applied successfully in all of the areas where the Bohr model failed. For instance, quantum mechanics predicted intensities for the spectral lines that matched the experimental data; the new theory can, in principle, be used for more complex atoms and molecules, although the mathematical procedures are too complicated to permit detailed solutions for very large atoms; quantum mechanics explained the splitting of the spectral lines into groups of lines; and the new theory explained the binding together of atoms in molecules, liquids, and crystals.

To date, quantum mechanics has been completely successful in dealing with all aspects of atomic physics. There is really only one disturbing fact about the quantum-mechanical model of the atom: it does not lend itself to picturing the atomic processes in terms of our everyday experience. No physical model of the atom corresponds to the extremely satisfactory mathematical expressions of quantum theory.

The Bohr Family

Niels married Margrethe Norlund in 1912. They had six sons, two of whom died in childhood. The four surviving sons all went on to successful careers. Aage followed in his father's footsteps, becoming a theoretical physicist. He won the Nobel Prize in Physics in 1975 for his work on the atomic nucleus.

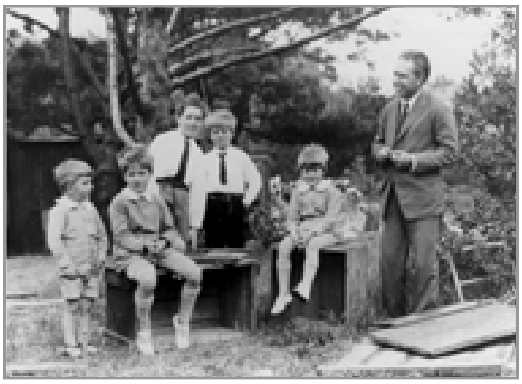


Figure 12.8 The Bohr family

Although the Bohr model was replaced after 1925 by the theory of quantum mechanics, Bohr's model was a very significant step forward in the quest for an understanding of the atom. Even today, the Bohr model is used in many cases because it gives good quantitative results from relatively simple calculations. When we wish to visualize atomic processes, we normally use some version of Bohr's model.

No one recognized the shortcomings of the Bohr model more fully than Niels Bohr himself. In the next chapter, we'll see that his inspiration and encouragement played a major role in the development of quantum mechanics, and his unrelenting efforts to understand the underlying philosophical meaning of the new physical theory led to the interpretation that is most widely accepted today.

Summary

When a low-pressure gas is excited (for instance, by running an electric current through it, or by heating it to a high temperature), the gas will emit electromagnetic radiation. When the emitted radiation is analyzed on the basis of wavelength, it is found to contain only certain wavelengths that are characteristic of the particular chemical element. Each chemical element has its own distinctive emission spectrum. Because it is clear that the radiation is being emitted by the individual atoms of the gas, a good atomic model should have something to say about the nature of this radiation.

By 1913, the Danish physicist Niels Bohr had taken the first step toward providing an atomic model that can account for characteristic emission spectra, at least for that of hydrogen. Bohr assumed that the hydrogen atom can exist only in certain allowed energy states. Contrary to Maxwell's theory of electromagnetism, Bohr assumed that the atom will not give off electromagnetic radiation so long as it remains in one of the allowed energy states. However, the hydrogen atom can absorb or emit energy in the form of electromagnetic radiation by making transitions between allowed energy states. For example, a hydrogen atom can go from a higher energy state to a lower energy state by giving off electromagnetic radiation of the appropriate energy. Because only certain energy states are allowed, only photons corresponding to energy differences between these allowed states will be emitted. Because the wavelength of the radiation is related to the energy of the photon, ($E = hc / \lambda$), this model accounts for the fact that the atom emits only certain wavelengths. His rule for the quantization of the electron's angular momentum allowed Bohr to calculate the allowed energies for the hydrogen atom, and the calculated energy differences corresponded to the wavelengths observed in the emission spectrum of hydrogen.

Bohr's model of the hydrogen atom is a combination of classical physics with some of the newer quantum ideas. The energy calculations for the electron's orbits and the planetary model itself are based on the classical physical theories. However, the quantization of energy and angular momentum, and the suspension of Maxwell's theory, are sharp breaks with classical physics. It was obvious to Bohr and other physicists that this hodgepodge of classical and quantum ideas was unsatisfactory. A new physical theory of mechanics based upon the quantum hypothesis would be required for the building of a fully satisfactory model of the atom.

Important concepts

Continuous spectrum; emission spectrum; absorption spectrum; atomic spectrum; Bohr's postulates; ground state; excited state; energy-level diagram for hydrogen.

Questions

1. Explain the difference between a continuous spectrum and a bright-line emission spectrum. Under what conditions is each produced?
2. List all of the natural or artificial sources of light that you can think of. In each case, state whether you would expect the spectrum of light from this source to be a continuous spectrum or a line spectrum.
3. What is the Balmer-Rydberg equation? How was it obtained? Explain exactly to which experimental data it can be applied.
4. Explain why the principles of classical physics lead to the prediction that Ruther£ord's atomic model is unstable. How did Bohr deal with this difficulty?
5. Explain what an energy level is. According to the Bohr model, how does an atom change its energy?
6. State Bohr's postulate about the angular momentum of an orbiting electron. What is meant by the following statement: "The angular momentum of the electron is quantized"?
7. An atom is often described as being like a miniature solar system. In what ways does the Bohr model of the atom resemble the solar system? In what ways is it different?
8. Show that $2p^2m_ek^2e^4/h^2 = 13.6$ eV.

9. Describe the process of ionization in terms of the Bohr model. What is meant by the ionization energy? What is the ionization energy for hydrogen?
10. Hydrogen gas at room temperature absorbs light of wavelengths corresponding to the lines of the Lyman series, but it does not absorb light of wavelengths corresponding to the lines of the Balmer series. Explain this fact in terms of the Bohr model.
11. Can a hydrogen atom in the ground state absorb a photon of energy 5.2 eV? Explain your answer.
12. The electron in a hydrogen atom has an angular momentum of $2h/2\pi$. What is the energy of the atom in this state?
13. A hydrogen atom emits a photon of frequency 2.46×10^{15} Hz as it makes a transition from some higher energy state to the ground state. What was the energy of the hydrogen atom before the transition?
14. Explain how the Bohr model accounts for the main features of emission spectra. Why must a gas be excited in order to give off an emission spectrum?
15. Using the energy-level diagram for hydrogen, calculate the wavelength of the light emitted and identify the region of the electromagnetic spectrum where the light will be found for the following transitions: from the state $n = 3$ to the state $n = 1$; from the state $n = 4$ to the state $n = 2$, from the state $n = \infty$ to the state $n = 2$, from the state $n = 4$ to the state $n = 3$.
16. Use the Bohr model to find the value of the Rydberg constant R,
17. A hydrogen atom in an excited state makes a transition to the ground state by emitting a photon of frequency 2.92×10^{15} Hz. What is the wavelength of the emitted radiation? What was the initial energy of the hydrogen atom?

    The following questions are of a more general nature. They have no single correct answer and are just something for you to think about. When possible, they are best answered in conversation with others.
18. A helium atom normally contains two electrons. Singly ionized helium has lost one of the electrons, so an ion consists of the helium nucleus and one electron. Using the Bohr postulates, construct an energy-level diagram for singly ionized helium. (Hint: In this case $Z = 2$.)
19. Discuss the nature and role of the postulates in a physical theory. What do you think of Bohr's postulates? Do you think they contribute to our understanding of the hydrogen atom?
20. Obviously, Bohr chose to quantize the angular momentum of the electron rather than some other property of the electron in order to obtain results that agree with experimental data. (He could, for example, have chosen to quantize the electron's kinetic energy or its speed.) Does this mean that Bohr's model is simply built to match the facts that were already known? Why were physicists impressed by the success of the Bohr model in "predicting" results that agree with the existing experimental data?

Niels Bohr (1885 – 1962)

The 1920s were a time of revolution, both political and intellectual. Many Europeans returned from World War I passionately committed to overthrowing the old guard. This was true in science as well as the arts. Just as Paris became the eye of the artistic hurricane, the Copenhagen of Niels Bohr became the focal point for upheaval in theoretical physics.

The Carlsberg Brewery contributed to this intellectual ferment by founding an institute of atomic studies to be headed by Bohr, a 1922 Nobel laureate. A steady stream of young physicists would travel to Copenhagen to debate the implications of each new development. Unlike Einstein, who worked best in isolation, Bohr was a gregarious and active fellow who preferred to think out laud – across the ping-pong table, along a hiking trail or a soccer field. (See the following chapter for more discussion of the famous Copenhagen debates.)

Bohr's influence as a mentor continued through the next three decades. During the 1930s, the gentle and humane Bohr helped many Jewish colleagues find safe haven outside Hitler's Germany. He himself eventually had to stage a harrowing escape to England. He devoted his later years to the development of peaceful uses for atomic energy.

13 QUANTUM MECHANICS

The quantum revolution began with Max Planck's proposal of the quantum hypothesis in 1900 in his explanation of thermal radiation. It continued with Einstein's explanation of the photoelectric effect in 1905 and Bohr's model of the hydrogen atom in 1913. But the revolution had not yet reached a satisfactory conclusion. The old classical theories had been toppled from the position of absolute power, but they had been replaced by a hodgepodge of the old ideas, arbitrary quantum postulates, and computational recipes. Clearly this is a very flimsy foundation upon which to build an understanding of the micro-world. What was needed was a new physical theory having the same sort of generality as that of classical mechanics and electromagnetism. Such a theory cannot be based on a large number of postulates, each chosen to fit a restricted range of experimental results. It must be based on a few simple, elegant, powerful postulates chosen because they seem the best logical basis for a coherent description of physical reality. Just such a theory had its beginnings in the somewhat metaphysical musings of a young French aristocrat, Prince Louis Victor de Broglie (1892 – 1987).

13.1 DE BROGLIE WAVES

Louis de Brog1ie's 1926 doctoral thesis was inspired by Einstein's photon model of light. Einstein ascribed a particle-like nature to electromagnetic radiation, in contrast to the classical view that this radiation is a wave phenomenon. However, such phenomena as diffraction and interference can be explained only in terms of a wave model. Thus, electromagnetic radiation was seen to have a dual nature. Only a particle model can explain such phenomena as the photoelectric effect and the Compton Effect. Only a wave model can explain such phenomena as interference and diffraction. De Broglie reasoned as follows: the universe consists of electromagnetic radiation and material particles (such as electrons, protons, and atoms). If electromagnetic radiation has a dual wave-particle nature, then symmetry suggests that material particles should also possess a dual nature. That is, material "particles" should, under certain circumstances, behave in a wavelike fashion. What an absurd idea. Everyone "knows" that electrons and protons are like tiny marbles and are not in the least wavelike in their behavior.

The faculty at the Sorbonne found itself in a quandary over de Broglie's thesis. Here was a young man of obvious intelligence, from one of the most influential families in France, and the brother of a distinguished physicist. Yet his thesis was truly bizarre. This convert from the humanities had proposed the most outrageous hypothesis, offering no experimental evidence in its favor. On almost metaphysical grounds, de Broglie proposed that each material particle has associated with it what he called a "pilot wave," but he offered no physical interpretation of the nature of this wave. He characterized his theory as "a formal scheme whose physical content is not yet determined." The faculty resolved the problem by sending a copy of the thesis to Einstein for his opinion. Einstein replied, "it may look crazy, but it is completely sound." With no less an authority than the renowned Einstein standing behind de Broglie's work, the faculty was off the hook. They were able to accept the thesis and award the doctorate to de Broglie without qualms. In retrospect, it was a very wise decision. De Broglie received the 1929 Nobel Prize in physics for the ideas set forth in his thesis, giving him the distinction of being the first person to receive this award for a doctoral thesis.

Louis Victor de Broglie
(1892 – 1987 * France)

1924 – making an analogy with electromagnetic radiation, submits a PhD. thesis suggesting that particles such as electrons can exhibit wave properties.

He sided with Einstein and Schrodinger in opposing Bohr, Heisenberg, and Bom's interpretation of the wave function.

1929 – Nobel Prize in Physics.

Figure 13.1 Louis de Broglie

De Broglie's wave model of material particles was based on an analogy with electromagnetic radiation. The wave properties (wavelength and frequency) of electromagnetic radiation are related to the particle properties (energy and momentum) in a precise, well-defined manner:

$$E = h\nu \text{ and } p = h / \lambda$$

where E is the energy and p is the momentum of the photon; and ν is the frequency and λ is the wavelength of the wave. The photon, because it travels at the speed c, must have a rest mass of zero according to the theory of relativity. Material particles such as electrons and protons have a nonzero rest mass and travel at speeds that are smaller than c. However, de Broglie assumed that the relationship $p = h / \lambda$ can also be applied to particles. Therefore, the wavelength of a particle is

$$\lambda = h / p = h / mv.$$

Example 13.1

An electron is traveling at a speed of 0.01c, or 3 x 10^6 m/s. What is its wavelength?

Solution

The speed of the electron is a small fraction of the speed of light, so we can use its rest mass, m_{eo}. = 9.11 x 10^{-31} kg.

$$\lambda = h / mv = 6.63 \times 10^{-34} \text{ J s} / (9.11 \times 10^{-31} \text{ kg}) \times (3 \times 10^6 \text{ m/s}).$$

$$\lambda = 2.42 \times 10^{-10} \text{ m} = 2.42 \text{ Å}$$

For comparison, electromagnetic radiation with this wavelength would be in the X-ray portion of the spectrum. However, although the electron has this wavelength associated with it, it is not electromagnetic radiation and its properties are quite different from those of X-rays.

Example 13.2

A baseball of mass 0.15 kg travels at a speed of 30 m/s. What is its wavelength?

Solution

As before, we write

$$\lambda = h / mv = 6.63 \times 10^{-34} \text{ J s} / (0.15 \text{ kg}) \times (30 \text{ m/s}) = 1.47 \times 10^{-34} \text{ m}.$$

This wavelength is only about one hundred-million-million-millionths of the diameter of the nucleus and is immeasurably small. In other words, the wave properties of a baseball traveling at 30 m/s are undetectable.

Example 13.3

If the wavelength of the baseball in Example 13.2 were, say on the order of one meter, it would seem that we might be able to detect the wave behavior of the baseball and check the relationship proposed by de Broglie. If the wavelength is to be l m, what must be the speed of the baseball?

Solution

$$\lambda = h / mv \text{ or } v = h / m\lambda$$

$$v = 6.63 \times 10^{-34} \text{ J s} / (0.15 \text{ kg}) \times (1 \text{ m}) = 4.42 \times 10^{-33} \text{ m/s}.$$

(Recall 1 J = 1 kg m^2/s^2.) At this speed, it would take the baseball 7.15×10^{24} years to travel one meter. For all practical purposes, the baseball is at rest, so again its wave properties are unmeasurable.

The Correspondence Principle requires that if the de Broglie hypothesis applies to an electron, it must also apply to a baseball. It is well established that such material objects as baseballs are described completely in terms of their behavior as particles. A baseball does not exhibit interference or diffraction or other wavelike properties. However, the de Broglie's hypothesis leads only to immeasurable predictions for the wave properties of a baseball. However, the predicted wave properties of electrons can be tested experimentally.

Even before any experimental evidence was available, there was one other argument that made the de Broglie hypothesis seem interesting to physicists. The Bohr model of the atom involves the postulate that electron orbits are quantized -- only orbits of certain sizes are allowed. If we think of the electron as a small particle circling the nucleus, it is difficult to see why certain orbits should be permitted and others forbidden. However, if we think of the electron as a wave, then it seems quite natural that the "orbits" should be quantized. As an analogy, consider a guitar string. The string vibrates only with certain "allowed" wavelengths; the other "forbidden" wavelengths die out almost immediately after the string is excited (plucked). Because the string is held in a fixed position at its ends, it can vibrate only with wavelengths such that

$$2 L = n \lambda \text{ for } n = 1.\ 2.\ 3.\ \ldots..$$

where L is the length of the string (see Figure 13.2).

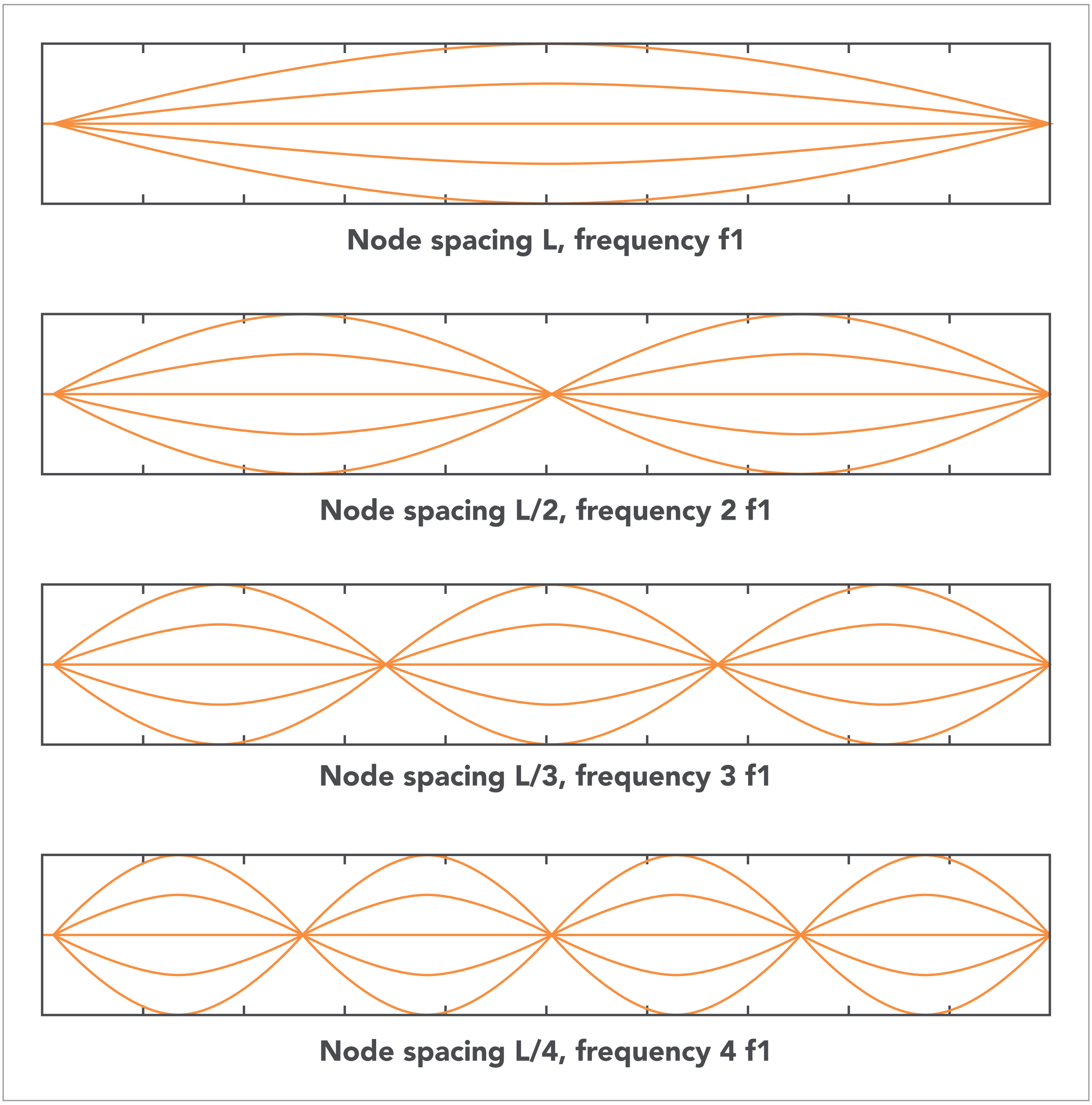


Figure 13.2 guitar string

We can think of the electron in its orbit as being rather like a wave on a circular guitar string. If the wave pattern is to be stable, the motions of the wave must match where the "ends" of the string are wrapped around and joined together. That is, the circumference of the orbit must contain some whole number of wavelengths. The circumference of an orbit of radius r is $2\pi r$, so we can write

$$2\pi r = n\lambda \text{ for } n = 1, 2, 3, \ldots$$

This equation should describe the stable orbits for an electron of wavelength λ. From de Broglie's relationship, λ = h/mv. So we can write

$$2\pi r = nh / mv \text{ or } mvr = nh / 2\pi$$

This is exactly postulate IV of Bohr's model of the atom! The quantum rule that appeared as an arbitrary assumption in Bohr's model now can be seen as a natural consequence of the assumption that the electron has a wavelike nature.

Physicists were impressed by the fact that de Broglie's hypothesis led to such a natural derivation of Bohr's quantum postulate. However, they remained skeptical about the validity of de Broglie's ideas until more direct experimental confirmation was obtained in 1927. For many years, two American physicists, Clinton Davisson and Lester Germer, had been studying the scattering of electrons by a solid. They directed a beam of electrons perpendicular to the face of a crystal and then measured the rate of electron scattering as a function of scattering angle. Their results were not consistent with predictions based on the classical treatment of the electrons as particles. However, they found that their results were quite consistent with predictions based on treating the beam of electrons as a wave of the wavelength predicted by de Brog1ie's relationship and computing the interference effects to be expected. This experiment was quite similar to the one used in 1913 by the German physicist Max von Laue to demonstrate the wave nature of X-rays. When the Davisson-Germer results were announced in 1927, there remained little doubt that electrons do in fact behave as waves just as de Broglie had predicted. Since then, the wave nature of other elementary particles, atoms, and even molecules has been demonstrated.

Ironically enough, the wave nature of electrons was independently demonstrated at the same time by George Thomson, the son of J. J. Thomson, who had first demonstrated that electrons are particles. George Thomson and Clinton Davisson shared the 1937 Nobel Prize in physics for their work.

13.2 SCHRODINGER'S WAVE EQUATION

De Broglie predicted (and experiments confirmed) that each material particle has associated with it wave properties such as a wavelength. For a particle with a well-defined momentum mv, the wavelength is given by the simple relationship $\lambda = h / mv$. However, there is much more that must be known about these mysterious matter waves if they are to be of use in physics. Such macroscopic wave phenomena as sound and water waves can be described by wave equations based on the theory of classical mechanics. Maxwell's theory of electromagnetism provides wave equations to describe electromagnetic waves. What was needed was a general wave equation that would fully describe the properties of the matter wave associated with a particular material particle.

Einstein was among the first to become aware of de Broglie's ideas. Early in 1925, in a paper on gases, Einstein referred to de Broglie's work and stated, "I believe that it involves more than merely an analogy." This remark led Erwin Schrodinger (1887 – 1961), a physicist at the University of Zurich, to begin studying matter waves and possible wave equations for them. However, he failed to come up with any equation that was consistent with experimental data, and he abandoned his efforts.

Several months later, the eminent physical chemist Peter Debye was instrumental in rekindling Schrodinger's interest in matter waves. Debye was in charge of the physics colloquium at the Swiss Federal Institute of Technology in Zurich (Einstein's alma mater). One day Debye suggested to Schrodinger that, since he was not working on anything important at the time, he should prepare a presentation for the colloquium on the recent thesis of de Broglie. After the presentation, Debye casually remarked that it seemed childish to discuss these hypothetical waves if there was no corresponding wave equation from which the properties of the waves could be calculated. Only a few weeks later, Schrodinger opened another presentation to the colloquium by saying, "My colleague Debye suggested that one should have a wave equation; well, I have found one."

By making an analogy between the matter waves and mechanical waves, Schrodinger was able to produce a general wave equation whose solution would correspond at all points in space and time to the properties of the matter wave associated with a particular microscopic particle. Schrodinger's wave equation can also be applied to macroscopic objects. In these cases, as required by the correspondence principle, the solution of the wave equation yields results identical with those predicted on the basis of classical mechanics in which the object has no wave properties. Schrodinger's equation is written,

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \left[\frac{-\hbar^2}{2\mu}\nabla^2 + V(\mathbf{r},t)\right]\Psi(\mathbf{r},t)$$

The symbol ψ (the Greek letter *psi,* pronounced sigh*)* represents the wave function, a mathematical expression describing the matter wave. The expressions ∇^2 and $\frac{\partial}{\partial t}$ represent mathematical operations that are to be performed on the wave function ψ , and V is the potential energy of the system. The symbol 'i' is the imaginary number $\sqrt{-1}$. This is the first time that an imaginary number had appeared in a physical law.

Mathematicians would describe the form of Schrodinger's equation as a partial differential equation. A good deal of training in calculus is needed to solve this type of equation. We will not be able to solve the equation for even the simplest physical situation. The solution to the wave equation represented by the Greek letter psi, is called the wave function. The wave function is a mathematical expression that describes the matter wave at all points in space and time.

Erwin Schrodinger
(1887 – 1961 * Austria)

1926 – developed wave equation for de Broglie waves which became the foundation for wave mechanics.

1926 – demonstrated that wave mechanics and matrix mechanics were two different mathematical formulations of the same theory, now known as quantum mechanics.

He attempted, unsuccessfully, to interpret the wave function as physical real, hoping to eliminate wave-particle duality and quantum jumps.

1933 – Nobel Prize in Physics.

Figure 13.3 Edwin Schrodinger

As a test of his theory, Schrodinger solved the wave equation for the wave function of the hydrogen atom. The result, published early in 1926, was astounding. For an electron bound to the hydrogen nucleus, the Schrodinger theory yields a series of wave functions, each corresponding to an allowed energy state of the atom. These quantized energies proved to be identical to those of the Bohr model, which had been amply confirmed experimentally.

Schrodinger's theory of wave mechanics seemed to provide just the sort of new theory that was needed. It was based on a completely consistent set of postulates rather than an arbitrary mixture of classical and quantum ideas. As required by the correspondence principle, it predicted results consistent with classical theories for macroscopic phenomena. It not only provided a more satisfying explanation for the Bohr model of the atom, but it went on to explain other experimental data that could not be explained by the Bohr model. There was only one shadow over this great success; another German physicist at the same time had produced an equally satisfying theory that seemed to be based on different assumptions,

13.3 HEISENBERG'S THEORY OF MATRIX MECHANICS

After obtaining his doctorate from the University of Munich, Werner Heisenberg (1901 – 1976) went in early 1924 to Copenhagen to study under Niels Bohr. This was the beginning of a collaboration that would last for years and would produce some of the most revolutionary and significant theoretical results in the history of physics. While at Copenhagen, Heisenberg became convinced that the difficulties with the Bohr model of the atom resulted from the use of such unobservable entities as orbits and transitions between orbits. He set himself the task of creating an atomic theory based entirely on such observable quantities as the frequencies and intensities of the radiation emitted by atoms. The task was formidable. It could not be accomplished with the mathematical tools then used by physicists, but Heisenberg found an appropriate mathematical technique called matrix algebra among the branches of abstract mathematics. With the help of two eminent colleagues, Pascal Jordan and Max Born, Heisenberg completed his new theory of matrix mechanics by the end of 1925.

Werner Heisenberg
(1901 – 1976 * Germany)

1924 – developed matrix mechanics, one of two formalisms of quantum mechanics.

1927 – developed the uncertainty principle establishing the inability to make stimultaneous measurements of certain pairs of variables.

1927 – played a significant role in the development of the Copenhagen interpretation of quantum mechanics.

1932 – Nobel Prize in Physics.

Figure 13.4 Werner Heisenberg

At about the same time, Wolfgang Pauli applied the theory to the hydrogen atom, showing Heisenberg's matrix mechanics to be just as successful as Schrodinger's wave mechanics in explaining the Bohr model of the atom and in explaining other experimental data not covered by the Bohr model.

Wolfgang Pauli
(1900 – 1958 * Austria)

1925 – formulated the Pauli exclusion principle stating that no two electrons could occupy the same quantum state, thus explaining how electrons fill the allowed orbitals of the atom.

1925 – applied Heisenberg's matrix mechanics to the hydrogen atom showing that it reproduces the experimental data.

1945 – Nobel Prize in Physics.

Figure 13.5 Wolfgang Pauli

For a decade, physicists had recognized the inadequacy of the half-classical, half-quantum Bohr model. They knew that what was needed was a new physical theory, one that could stand on its own without the arbitrary use of certain classical ideas and the arbitrary rejection of other classical ideas. Now suddenly, in the early months of 1926, they had not one but two such theories! An embarrassment of riches, to be sure.

Much to the relief of the physics community, it was not long before Schrodinger was able to show that the two theories, despite their different approaches, are in fact mathematically equivalent. Although they use different mathematical tools and languages, they are based on the same physical assumptions and they lead to the same physical predictions. Thus the two theories were seen to be actually a single theory of quantum mechanics, which can be expressed in either the language of wave mechanics or the language of matrix mechanics.

At last we have the theory physicists were searching for. Quantum mechanics completely supplants the arbitrary quantum hypotheses of Planck, Einstein, and Bohr. Not only does it duplicate all of their successes, but it extends well beyond them to encompass physical phenomena that the old quantum models could not explain. It is every bit as elegant as the theories of classical mechanics and electromagnetism. It is a more general theory than classical mechanics, including classical mechanics as a special case in the limit as sizes become large enough to approach the macroscopic scale. This is required by the correspondence principle.

Quantum mechanics, as it was developed in 1926, was a nonrelativistic theory. As you might suspect, this led to difficulties, especially when dealing with particles such as electrons that can be accelerated in the laboratory to relativistic speeds. In 1928, the brilliant young English physicist P. A. M. Dirac succeeded in formulating a relativistic wave equation for electrons, thus completing the development of the quantum theory. (You may recall from chapter 10 that in the course of this work Dirac first predicted the existence of antiparticles.)

P. A. M. Dirac
(1902 – 1984 * England)

1925 – Reformulated Heisenberg's matrix mechanics.

1928 – Relativistic wave equation. Predicted positron.

1933 – Nobel Prize in physics.

Figure 13.6 P. A. M. Dirac

The relativistic wave equation was not Dirac's first significant contribution to quantum mechanics. After reading the proofs of Heisenberg's first paper on matrix mechanics, Dirac developed a neater and more convenient mathematical expression for the theory.)

The mathematical techniques of matrix mechanics are even more abstract and difficult than the partial differential equations required for wave mechanics. Therefore, quantum mechanics usually is described in terms of wave mechanics and the wave function. Schrodinger's nonrelativistic formulation is sufficient for most applications, but Dirac's relativistic version of the theory is used where necessary. The theory of quantum mechanics has been completely successful in accounting for experimental data dealing with atomic phenomena.

13.4 WHAT IS A MATTER WAVE?

We have not yet addressed one very fundamental question about Schrodinger's formulation of quantum mechanics. In developing his wave equation, Schrodinger used an analogy with such mechanical waves as sound or water waves. For such a mechanical wave, the wave function that is the solution to the wave equation has a simple physical interpretation. For example, the wave function represents the amplitude of a water wave or the pressure of a sound wave. But what is represented by the wave function in Schrodinger's wave equation?

Thus far, we have treated the wave function strictly as a computational device, an un-interpreted mathematical expression that can be used to calculate certain physically meaningful quantities such as energy or wavelength. Is there more to it than that? Just how does the wave function correspond to "reality" in the traditional sense of the word? Is it a measurable physical quantity? Before we can consider quantum mechanics a satisfactory theory, we must have some satisfactory physical interpretation of the theory, which must include answers to questions such as these. Schrodinger made one attempt at such an interpretation soon after he completed the mathematical formulation of wave mechanics.

Schrodinger suggested that the wave function represents a continuous distribution of charge and mass in space-time. The concept of wave-particle duality cannot be visualized in terms of any physical model, so Schrodinger suggested that the electron be regarded as a matter wave with the mass and charge being smeared out over the region of space where the wave function is not equal to zero. This picture seems to fit quite well for the atom. The electron can be regarded as a wave pattern surrounding the nucleus rather than a particle orbiting around it, so that the model of the atom no longer conflicts with Maxwell's theory of electromagnetism. Recall that Maxwell's theory predicts electromagnetic energy will be radiated by an orbiting charged particle, so that the atom should quickly collapse as it loses energy. However, if the electron is simply a stable matter-wave pattern, then the charge is not circling the atom, so no electromagnetic radiation should be emitted. (Bohr's model had simply postulated that no radiation is emitted when the electron is in an allowed orbit, but this was an arbitrary assumption that really satisfied no one as an explanation.)

Perhaps the most appealing aspect of Schrodinger's interpretation is that it appears to restore continuity to physics. In his model of the hydrogen atom, Bohr had introduced the idea of quantum jumps – processes that cannot be given continuous descriptions in space-time. For obvious reasons, this concept of quantum jumps was profoundly disturbing to most physicists; the mechanistic-deterministic world view was still fundamental to their thinking. According to Schrodinger's interpretation of quantum mechanics, an atomic transition could be viewed as the gradual disappearance of one wave pattern with the simultaneous appearance of another wave pattern. Such a description could be viewed as a continuous process in space-time, one that seemed much more comfortably familiar in terms of everyday macroscopic experience. For these reasons, Schrodinger's interpretation was enthusiastically supported by the more traditional-minded physicists such as Einstein and Planck.

In Schrodinger's interpretation, an electron is regarded as a matter wave spread out over a region of space. But what about the many experiments that seem to indicate a particle nature for the electron? In such experiments, the results indicate that the mass and charge associated with the electron are localized in a very small region of space. In the hydrogen atom, according to Schrodinger's interpretation, the matter wave of an electron is spread out over a volume millions of millions of times greater than the volume of the nucleus. However, a free electron seems to have its mass and charge localized in a volume that is no larger than the volume of the nucleus of a hydrogen atom. Can Schrodinger's wave interpretation explain such observations? Schrodinger initially thought he saw an answer to this problem.

It is a well-known property of waves that two or more waves occupying the same region of space at the same time will interfere with each other. This interference can be destructive (that is, the waves can cancel each other) or it can be constructive (that is, the waves can reinforce each other). If a large number of waves of slightly differing wavelengths interfere with each other, then the resulting wave pattern will be confined to a very small region of space, and the waves will completely cancel each other everywhere else. Such a localized wave is called a wave packet, and it has properties very much like those of a particle. With the appropriate combination of waves, it is possible to construct a wave packet that has a size and speed to match those of free electrons. So far, so good. There is only one problem.

No matter how one adjusts the properties of the matter waves that make up the wave packet, one finds that it stubbornly refuses to stay small. According to Schrodinger's own wave equation, the smallest wave packet will very quickly spread out to occupy a large volume of space. Thus Schrodinger's interpretation leads to the prediction that a free electron will quickly lose its identity as a small material particle; its mass and charge will swiftly become smeared out over space. This prediction is not consistent with the experimental evidence. Thus physicists reluctantly concluded that Schrodinger's interpretation cannot be a valid explanation of the physical meaning of the wave function

13.5 THE PROBABILITY INTERPRETATION OF THE WAVE FUNCTION

At about the same time Schrodinger was formulating his interpretation of the physical meaning of the wave function, another interpretation was being worked out at the University of Gottingen by Max Born (1882 – 1970), one of the principle developers of matrix mechanics. Born was strongly influenced by the interpretation that Einstein had presented for the wave-particle duality applied to electromagnetic radiation. Einstein had suggested that the electromagnetic wave is a kind of "phantom wave" that serves to guide the light quanta (photons). This concept is very similar to the one de Broglie suggested in his original description of "pilot waves" associated with material particles. In fact, the electromagnetic wave can be regarded as the de Broglie wave associated with a photon. The square of the wave amplitude at any point in space is proportional to the probability of finding a photon at that point. Carrying this idea directly over to matter waves, Born suggested that the square of the wave function at any point in space represents the probability of finding the electron at that point. In other words, the electron is most likely to be found where the square of the wave function is large and is less likely to be found where it is small.

Max Born
(1882 – 1970 * Germany)

1925 – recognized that the mathematics of Heisenberg's new quantum theory as matrix algebra, and with Heisenberg and Pascal Jordan, developed matrix mechanics.

1926 – formulated the, not standard, interpretation of the wave function times its complex conjugate as a probability density function.

1954 – Nobel Prize in Physics.

Figure13.7 Max Born

Whereas Schrodinger tried to treat the electron as a matter wave smeared out over space. Born seems to take just the opposite view. He seems to treat the electron as a particle whose probability of being found at any particular point in space and time is given by the wave function, an abstract mathematical function that simply represents a probability. According to the Born interpretation, the wave function does not correspond to any physical quantity Instead. it is a kind of probability wave that indicates the likelihood of finding the particle at particular points in space and time. The Born interpretation is extremely useful in accounting for a wide range of experimental results in atomic physics. However, like the Schrodinger interpretation, it leads to difficulties in certain kinds of experiments.

In the double-slit experiment, a beam of electrons (or photons) is aimed at a phosphorescent screen. When electrons strike the phosphorescent screen, light is emitted (as in a television). In this way, it is possible to determine the distribution in space and time of the electrons striking the screen. Inserted between the source of the beam and the screen is a partition with two tiny slits cut in it a short distance apart. Because the partition is opaque to electrons, only the electrons that pass through one slit or the other can reach the screen. In our experiment the intensity of the electron beam is reduced until only one electron at a time is passing through the apparatus.

First, close one of the slits, so that the electrons can pass only through a single slit. If the electrons act strictly as particles, they should pass in straight lines through the open slit to the screen, forming a sharp image of the slit on the screen. If they act strictly as waves, diffraction effects should cause the waves to spread out after passing through the slit, forming a pattern on the screen that is larger than the straight-line image of the slit (see Figure 13.8). When the experiment is actually done, the electron interacts with the screen as a particle would, and a single flash of light is emitted from a point on the screen. However, the point is not necessarily on a straight line from the source through the slit! Later, other electrons strike the screen after passing through the same slit, but each electron strikes at a different point, also not necessarily on a straight line from the source through the slit. If we take a time exposure of the light flashes from the screen, we find after a large number of electrons have had time to pass through the slit, the electrons are striking the screen in a definite pattern, and further that the pattern is exactly the one predicted by the wave model!

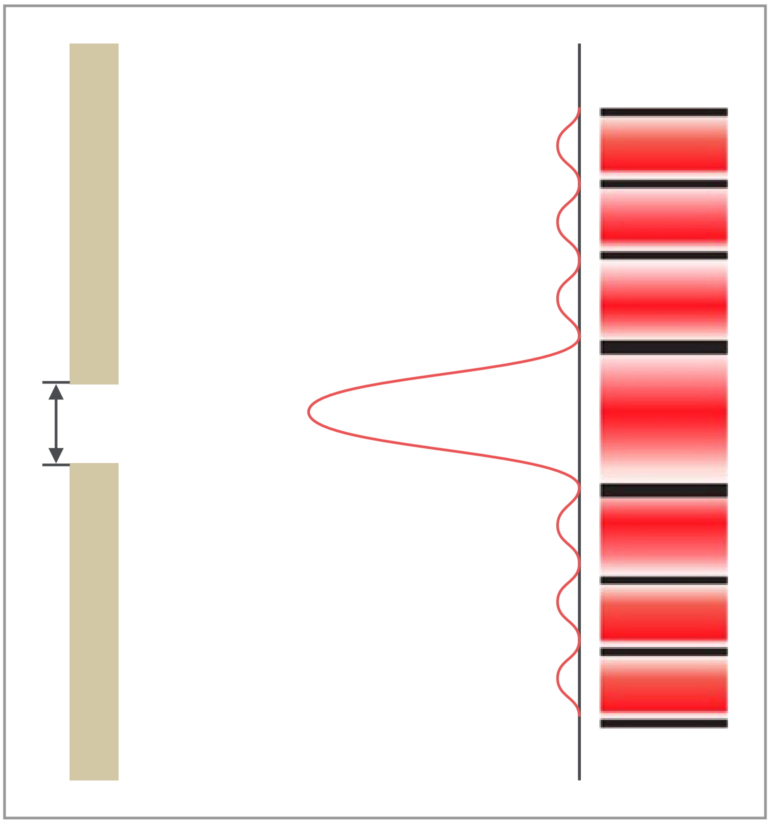


Figure 13.8 The single-slit experiment

Now let us perform the experiment using both slits. We might logically expect that because each electron has to pass through one slit or the other, the pattern we would get after a large number of electrons have had time to strike the screen would be a combination of the two individual one-slit patterns. (Think about this; make sure it makes sense.) If the slits are opened one at a time, even if they are alternated back and forth very quickly, this expected pattern is exactly what we do get. However, if both slits are open at the same time, and time is allowed for a large number of electrons to hit the screen, a completely different pattern results -- the one shown in Figure 13.9 and 13.10!

Double - Slit Experiment

If the difference in path length, x, is an integer multiple of the wavelength, the two light beams will be exactly in phase with each other and will constructively interfare. If the difference is 1/2, 3/2, etc., they will be exactly out of phase and will cancel out. For other values of x, there will be varying phase differences and the two light beams will produce alternating regions of brightness and darkness

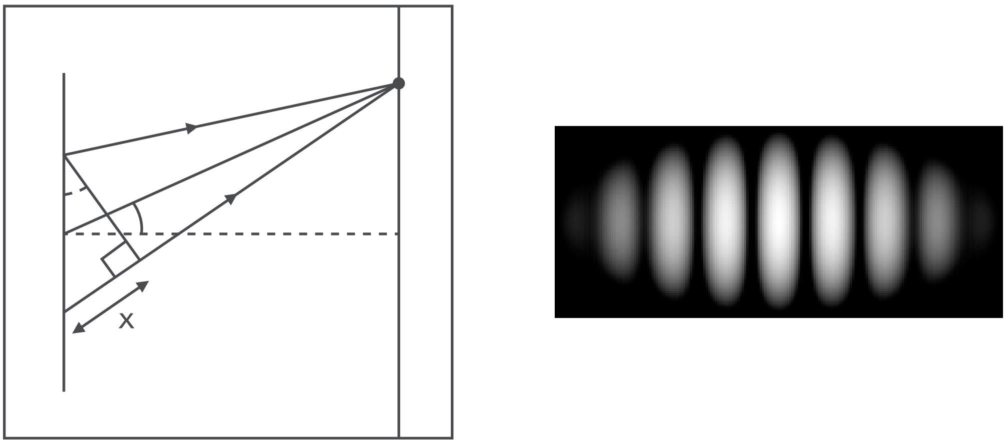


Figure 13.9 The double-slit experiment

Think about what Figure 13.10 must mean. The electrons are shot through one at a time, and yet; the accumulated pattern of electrons striking the screen depends on whether the two slits are open simultaneously or alternately. It is as if an electron somehow “knows” when passing through one slit whether the other slit is open or closed, and this fact influences the direction of its path after passing through the slit. Apparently, the presence or absence of the other slit, the one through which the electron seemingly did not pass, influences the electron’s motion! Remember, because the electrons are shot out one at a time, it is not possible that this is a collective effect with electrons passing through different slits interfering with or otherwise influencing each other.

Low Intensity, Double - Slit Experiment

The electrons strike the screen as a particle would, seeming at random points. However, as the number of electrons hitting the screen increases, a pattern begins to emerge, the familiar double-slit pattern.

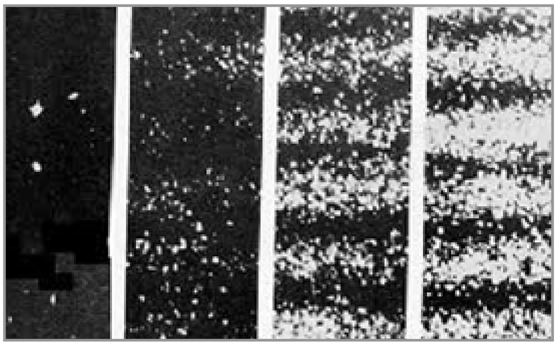


Figure 13.10 The double-slit experiment

What do these strange results have to say about Schrodinger's or Born's interpretation of the wave function? Schrodinger's concept that the wave function represents matter spread out over a region of space seems to be ruled out by this experiment. The interaction between the electron and the screen is localized; the electron exhibits clear particle-like behavior that is inconsistent with Schrodinger's view. Born's interpretation of the wave function is consistent with the single-slit experiment. Each individual electron behaves like a particle moving from the source to a single small spot on the screen, but the probability of the electron striking a given location on the screen is proportional to the wave function squared, with the wave function being spread out in wave-like fashion after passing through the slit. Born's interpretation is also able to explain the double-slit experiment if only one slit is open at a time.

However, the pattern produced when both slits are open at the same time is not consistent with the Born interpretation. Even though only one electron passes through the apparatus at a time, the time exposure of the individual impacts shows interference effects between the two slits. It is clear that the wave function associated with each particle somehow interferes with itself and therefore must represent some physical phenomena and not merely the mathematical probability of finding the electron at a given point. Thus, the experimental results support neither the Schrodinger interpretation nor the Born interpretation. The concept of wave-particle duality appears to be a necessary component of modern physical theory, whether we like it or not. Neither a wave model alone nor a particle model alone can account for the double-slit experimental observations.

Quantum Mechanics Is Not a Deterministic Theory

Newtonian mechanics is a deterministic physical theory. Given a set of initial conditions, we can determine with absolute certainty all future states of a system. The deterministic world view based upon Newtonian mechanics is dramatically summarized in a statement made by the mathematical physicist Pierre Laplace near the beginning of the nineteenth century.

> We must thus envisage the present state of the universe as the effect of its previous state, and as the cause of that which will follow. An intelligence that could know, at a given instant, all the forces governing the natural world, and the respective positions of the entities which compose it, if in addition it was great enough to analyze all this information, would be able to embrace in a single formula the movements of the largest bodies in the universe and those of the lightest atom: nothing would be uncertain for it, and the future, like the past, would be directly present to its observation.

However, Born's interpretation of quantum mechanics introduced into physics the concept that reality might be best described in terms of probabilities, thus initiating a philosophical debate that has continued to this day. According to classical physical theories, probability plays a useful role in physics only for the description of situations where the physicist has incomplete information. The behavior of a gas must be described in terms of probabilities simply because it is not possible to measure the position and velocity of every individual gas molecule at some given moment, but (as Laplace stated) a completely deterministic description of the behavior of every atom would be possible if we only had the capacity to achieve it.

Born's interpretation suggests a new and far more basic role for probability in physical theory. Even if the initial conditions are completely known, so that the wave function can be completely calculated and its behavior over time completely described according to the wave equation, still the best we can do is to compute the probabilities of detecting the particles at each future point in space and time. Thus, according to the Born interpretation, quantum mechanics is a non-deterministic theory. Of course, as we expect from the correspondence principle, the results of experiments on the macroscopic scale can be predicted in a deterministic fashion because they represent the summation of huge numbers of events on the atomic scale. Although we can predict only the probability of an individual electron striking the screen at a given point, we can predict with considerable certainty the pattern that will be formed by a large number of electrons striking the screen.

Many physicists were profoundly disturbed by the picture of physical reality that seemed to be emerging from quantum mechanics. They found it inconceivable that the ultimate nature of the universe could involve such illogical concepts as discontinuity (quantum jumps) and nondeterminism. At Niels Bohr's Institute for Theoretical Physics in Copenhagen, the atmosphere was intense and exciting, with constant debates on the new theories and the meaning of the new ideas. Among the most intense debaters at Copenhagen were Heisenberg and Bohr himself. These discussions often continued around the clock. 'Discussions' is perhaps too weak a word. Heisenberg was often reduced to tears by their intensity. Schrodinger was invited to Copenhagen specifically to discuss the differences of opinion between his interpretation and that of Bohr and Heisenberg. After a few days of intense confrontation, Schrodinger became ill, but the debate did not end. Bohr would sit on the edge of Schrodinger's bed constantly questioning, probing, and expounding his own point of view.

Werner Heisenberg and Niels Bohr

Heisenberg met Bohr when he was still a graduate student. Bohr soon became his mentor. Discussions, often arguments, between them played a significant role in both the development and interpretation of quantum mechanics.

Figure 13.11 Bohr and Heisenberg

13.6 THE COPENHAGEN INTERPRETATION

While the mathematical formalism of quantum mechanics was being developed by Heisenberg, Born, Dirac, Schrodinger, and others, Niels Bohr was struggling with the philosophical significance of the theory. As a result of his intense debates, Bohr was able to reach a coherent formulation of certain ideas that had concerned him for a long time.

Finally, in September of 1927 at a conference in Como, Italy, Bohr formally presented what is now known as the Copenhagen Interpretation. Bohr was a disaster as a public speaker. He generally spoke in a low tone, made worse by a strong accent, and would often switch from language to language. The main problem, however, was the fact that his sentence structure was convoluted. Some claimed that listening to Bohr was a lot like reading James Joyce's *Ulysses*, more a stream of consciousness than a coherent, structured presentation. The wife of a visiting professor at Bohr's institute, after listening to a welcoming lecture that ended to enthusiastic applause, turned to her neighbor and said she looked forward to hearing the English translation. He looked at her puzzled and said, "That was the English translation."

Como was no exception. His tortured and tortuous remarks left most of the audience utterly baffled. Others thought that Bohr had taken some well-known physics and clothed it in mysterious philosophical language. It is not clear that anyone outside Bohr's inner circle realized the significance of what Bohr was trying to say. Both Max Born and Heisenberg stood to say that they agreed with Bohr, this, only months after Heisenberg's fierce, tense standoff with Bohr over many of these same ideas. Despite much internal disagreement, the Bohr camp presented a united front at the meeting. Some later saw this as some sort of conspiracy to stifle criticism.

The Copenhagen Interpretation is in many ways vague and has sometimes been interpreted differently by different proponents. However, there are several basic principles that are generally accepted as being part of the interpretation.

1. A system is completely described by its wave function. There are no hidden variables. Any uncertainty that exists within the wave function is true uncertainty, not just a representation of our lack of knowledge about the system.
2. If the wave function is a mixed state with respect to a particular variable (that is if several different values of the variable are possible results of a measurement of the variable), the square of the wave function gives the probabilities of obtaining each of the possible results.
3. The result of a measurement on a mixed state is completely non-deterministic. There is no factor within either the wave function or the measurement device that will determine which of the allowed values will result, only the probability with which they might result. Quantum mechanics is a non-deterministic theory.

4. When a variable is measured, one of the possibilities is actualized. The wave function immediately collapses to one that is a pure state for the value that results from the measurement. All other possibilities cease to exist.
5. Measurements are not passive determinations of an objective world but active interactions with the thing measured. The way in which one chooses to measure the system becomes a part of the system and influences the outcome.

The Copenhagen Interpretation is based on the assumption that nothing is real until it is measured. Consider the wave function associated with the hydrogen atom. Each of the allowed energy states has its own wave function, which represents, among other things, information on the location of the electron with respect to the nucleus. When a measurement is made, the location is determined. Before the measurement, it may have been equally likely to be found on the opposite side of the atom. Does that mean that since it was found on one side that it could not have been on the opposite side? In the Copenhagen Interpretation the answer is no. The reason is that the measurement is not a passive determination of the location of the electron, as it would be in classical physics. In classical physics, the electron would have been at the location it was determined to be at whether or not a measurement is made. In the Copenhagen Interpretation, the measurement created the location of the electron and could just as likely have created it on the opposite side. The electron has no location until it is measured.

This renunciation of reality in the intuitive sense of the word profoundly disturbed Einstein. He remarked that, if Bohr's ideas were correct, he would rather be an employee in a gambling establishment than a physicist. The resulting clash between Bohr and Einstein, two of the greatest minds of their time (or any other time), is one of the most significant and interesting in the history of thought.

At the Fifth Solvay Congress in 1927, and later at the Sixth Solvay Congress in 1930, Einstein attempted to refute the Copenhagen Interpretation by proposing thought experiments in which it would be possible to provide a complete space-time description of the transfer of energy and momentum, thus refuting quantum mechanics. Every time Einstein proposed such a thought experiment, Bohr was able to show that Einstein had overlooked details of the measuring procedure that would introduce uncertainties consistent with the predictions of quantum mechanics. Einstein's most ingenious and famous paradox, a thought experiment called the photon box, was presented at the Sixth Solvay Congress. George Gamow described the confrontation as follows: "The argument seemed very persuasive and Bohr had nothing to say. But the next morning, after an almost sleepless night, Bohr, his face radiant, appeared at the meeting hall with an explanation." What Bohr had done was to use Einstein's own general theory of relativity to prove that the Copenhagen Interpretation of the thought experiment is consistent and satisfactory.

The Bohr - Einstein Debates

For both Bohr and Einstein, philosophy played a significant role in shaping their concepts of physical reality. Einstein was strongly committed to realism and determinism, while Bohr accepted quantum mechanism as a complete theory that ruled out both. Their most famous debates tookplace at the 1927 and 1930 Solvay conferences. Despite their differences of opinion, Bohr and Einstein remained close personal friends throughout their lives.

Figure 13.13 Bohr-Einstein debates

The discussions with Einstein were very helpful to Bohr in clarifying his own ideas. They continued for many years -- both in actual fact and in Bohr's mind. Even after Einstein's death, Bohr continued to ask himself how Einstein would have responded to ideas that Bohr was considering. In fact, when Bohr died in 1962, he left upon the blackboard in his study a drawing of Einstein's photon-box experiment that he had been considering the night before.

Although Einstein was defeated in his debates with Bohr, he was never convinced of the validity of the Copenhagen Interpretation. Finally abandoning hope of finding an internal inconsistency in the interpretation, he then attempted to show that the interpretation is incomplete. In 1935, Einstein and two younger colleagues, Boris Podolsky and Nathan Rosen, devised a thought experiment they believed demonstrated it is meaningful to attribute well-defined properties to elementary particles in the absence of a measurement.

The original Einstein-Podolsky-Rosen experiment was a thought experiment that could not actually be carried out. However, in 1969, John Bell, an Irish physicist, developed a theorem that could lead to a testable version of the EPR experiment These experiments have been carried out and the evidence is clear. Einstein was wrong. The results were completely consistent with predictions based on the Copenhagen Interpretation.

Einstein died in 1955, still believing the nature of reality must ultimately be deterministic -- that "God does not throw dice."

The period between the introduction of Bohr's atomic model in 1913 and the late 1920s was one of intense theoretical and experimental effort in the area of atomic physics. The result was that by 1930 most of the major questions about atomic phenomena had been satisfactorily answered. Physicists were now ready to direct their attention even more deeply into the micro-world, to the nucleus of the atom. In the next three chapters, we will see what they learned.

Summary

In the classical view of physics, matter is made up of discrete particles such as electrons, protons, and neutrons. Their behavior was understood in terms of the model treating them as tiny balls of matter. In 1924, Louis de Broglie suggested that the wave-particle duality already known to apply to electromagnetic radiation should also be applicable to material particles. On theoretical grounds, he suggested that the wavelength associated with an elementary particle is equal to Planck's constant divided by the momentum of the particle. This relationship was verified experimentally in 1927 when it was realized that the behavior of an electron beam scattered from a crystal surface could be understood only by treating the electron beam as a wave with the wavelength given by the de Broglie relationship.

In 1926, Erwin Schrodinger created a wave equation whose solution matched the properties of the de Broglie wave. Applied to the model of the hydrogen atom, Schrodinger's theory of wave mechanics completely reproduces and considerably extends all the achievements of the Bohr model. Another physical theory was formulated slightly earlier by Werner Heisenberg; his theory of matrix mechanics was equally successful in treating the hydrogen atom. It soon became clear that the two theories are equivalent, differing only in mathematical formalism. Together, they are now known as quantum mechanics. Quantum mechanics is a more general physical theory than classical mechanics. It gives the same physical results as classical mechanics in the size range where the laws of classical mechanics are known to be correct (the macroscopic scale), but (unlike classical mechanics) it also gives correct physical laws on the atomic scale.

Although the physical laws of quantum mechanics were quickly accepted as adequately explaining the physical data on the atomic scale, a controversy soon developed about the interpretation of the theory. What is the real nature of matter waves? Are the laws of physics inherently probabilistic, or are the probabilistic laws simply a result of our incomplete knowledge? Are there actually events that are discontinuous in space-time? The Copenhagen Interpretation proposed by Niels Bohr addresses some of these questions and generally speaking, comes down in each case on the side counter to our intuitive sense of the way nature "should" be. Experiments have been done that show predictions based on the traditional concept of reality are wrong but are consistent with predictions based on the Copenhagen Interpretation. Although the Copenhagen Interpretation is widely accepted among physicists today, the debate about the proper interpretation of quantum mechanics will undoubtedly continue for some time.

Important concepts

Wave-particle duality of matter; matter waves; Schrodinger's wave equation; wave mechanics; matrix mechanics; quantum mechanics; the Copenhagen interpretation.

Questions

1. An electron moves with a speed of 10^8 m/s. What is the de Broglie wavelength associated with the electron?
2. Explain how the de Broglie concept of matter waves provides a more satisfactory explanation of the quantized electron orbits in the Bohr model of the atom.
3. Explain how the results of the Davisson-Germer experiment supported de Broglie's concept of matter waves.

4. An automobile of mass 2000 kg is traveling at a speed of 30 m/s. What is the de Broglie wavelength associated with the automobile?
5. At what speed must an electron be traveling in order to have a wavelength of 5000 Å (the approximate wavelength of visible light)? What is the energy of such an electron in electron-volts?
6. In their electron-scattering experiment, Davisson and Germer used electrons with a velocity of 0.0015 c. What was the de Broglie wavelength associated with the electron beam?
7. What is the purpose of Schrodinger's wave equation?
8. Compare and contrast the interpretations of the wave function by Schrodinger and by Born.
9. What is a wave packet?

   The following questions are of a more general nature. They have no single correct answer and are just something for you to think about. When possible, they are best answered in conversation with others.

10. Is an electron a wave or a particle? Explain your answer.
11. Briefly discuss Einstein's opinions about quantum mechanics and the Copenhagen Interpretation?

12. Quantum mechanics is an example of a fundamental physical theory as defined in Chapter 1. Comment on the properties of physical theories as they apply to quantum mechanics.
13. Discuss how the Copenhagen Interpretation relates to the classical goal of obtaining a complete and rational description of physical processes.

Erwin Schrodinger (1887 – 1961)

Erwin Schrodinger's father was professor of chemistry at the Vienna Institute of Technology. Although Schrodinger's early interests included poetry and the grammatical structures of ancient languages as well as science, he received his degree in physics in 1910 from the University of Vienna. His academic career was interrupted by service as an artillery officer during World War I. After brief stays at Stuttgart and Breslau in Germany, Schrodinger in 1921 became professor of physics at the University of Zurich in Switzerland. There he developed his famous wave equation and formulated the theory of wave mechanics.

Somewhat unconventional in his personal habits, Schrodinger when traveling always carried his belongings in a rucksack on his back, so that he could wander off at a moment's notice. When arriving for a conference, he would walk from the train station using the time to organize his thinking.

In September 1926, Schrodinger visited Bohr's institute in Copenhagen to give a lecture on wave mechanics. He argued his viewpoint that the wave function represents the distribution of matter and charge in space-time, so that an electron or other elementary particle actually is "smeared out" in space. After the lecture, Schrodinger and Bohr debated this matter for days. Bohr argued for a probabilistic interpretation of the wave equation and for discontinuous quantum jumps made by discrete particles. At one point, the exasperated Schrodinger exclaimed, "If one has to stick to this damned quantum jumping, then I regret ever having been involved in this thing." Although Bohr did not persuade Schrodinger, neither did Schrodinger convince Bohr. After Schrodinger left, the intense debates continued between Bohr and Werner Heisenberg, who together developed the "Copenhagen Interpretation" that has dominated quantum mechanics since 1930.

In 1927, Schrodinger was invited to become Max Planck's successor at the University of Berlin. He accepted the post, but he left Germany for Oxford as soon as Hitler came to power in 1933, even though he had no reason to fear racial persecution himself. In that same year, he shared the Nobel prize in physics with P. A. M. Dirac. After a few years in England, Schrodinger returned to his native Austria to take a position at the University of Graz. When the Nazis annexed Austria in 1938, he fled to Italy before he could be arrested as an “unfriendly citizen.” The School for Theoretical Physics was established at The Institute of Advanced Studies in Dublin, primarily to provide a suitable post for Schrodinger, who became its director. He remained in Dublin until his retirement in 1955, when he returned to Vienna.

Schrodinger’s genius was many sided, in physics, his interests ranged from statistical mechanics to the unification of theories of gravitation and electromagnetism. He was particularly interested in the philosophical interpretation of quantum mechanics. Like Einstein, lie remained an opponent of the Copenhagen Interpretation until his death. Schrodinger was also very interested in biology, particularly in the concept of life. He published in 1944 a book titled *What is Life* that greatly influenced the later development of molecular biology.

Werner Karl Heisenberg (1901 - 1976)

Werner Heisenberg’s father was a professor of Byzantine history. Werner’s secondary education was interrupted by his service with the troops fighting the revolution in postwar Germany. When he did graduate from the Gymnasium in 1920, he immediately entered the University of Munich with the intention of studying theoretical physics. Three years later, and not yet 23 years old, Heisenberg earned his Ph.D. in physics. Only two years after that, he published his theory of matrix mechanics, which marked the beginning of the development of quantum mechanics. After some years with Bohr in Copenhagen, Heisenberg in 1929 became professor of theoretical physics at the University of Leipzig. There he published many papers on quantum mechanics and its applications. He also became interested in nuclear physics and made several significant contributions to that field.

In 1934, he received the Nobel prize in physics for his role in the development of quantum mechanics. He became director of the Max Planck Institute of Physics in Berlin in 1942 and remained in Germany during the war. As the leader of the German attempt to develop an atomic bomb, Heisenberg remained, at least on the surface, a supporter of the Nazis. However, it seems likely that he and many of his colleagues failed to pursue their work with much effort or enthusiasm, and the German project never came close to creating a functional atomic bomb. (However, some controversy remains on this point. See the play Copenhagen about a meeting between Heisenberg and Bohr that occurred during the war.) After the war, Heisenberg continued his theoretical research as director of the Max Planck Institute at Gottingen in West Germany. During this time, he became interested in elementary-particle physics and introduced the scattering-matrix (S-matrix) theory that is a very important part of modern research in particle physics.

Throughout his career, Heisenberg was interested in the history and philosophy of physics. He grew up in a household devoted to the humanities and steeped in culture. As a youth, he had an intense interest in the classics, especially in early Greek science. He read Plato, Democritus, and Thales of Miletus in the original Greek. Late in his life, he published a book called *Physics and Beyond* setting forth his recollections of the history of quantum mechanics and his speculations about its significance. In addition to his accomplishments as a leading theoretical physicist, Heisenberg was an enthusiastic outdoorsman and mountain climber, as well as an accomplished classical pianist.

14 THE NUCLEUS

Rutherford's alpha-scattering experiments demonstrated that most of the mass of the atom is contained in a tiny, incredibly dense nucleus at its center. Physicists almost immediately turned their attention to the question of the nature and structure of the nucleus. As is usually the case in physics, this new frontier of knowledge had many surprises in store.

14.1 THE COMPOSITION OF THE NUCLEUS

By 1920, it was well known that the masses of atomic nuclei are very nearly integer multiples of the mass of the hydrogen nucleus. Thus it seemed reasonable to believe that hydrogen nuclei – or protons, as Rutherford called them -- are the building blocks of the nuclei of the various elements. The other known elementary particle at the time was the electron whose charge was negative and equal in magnitude to the positive charge of the proton, and whose mass was extremely small compared to the mass of the proton ($m_p = 1840\ m_e$). By combining these two particles in appropriate numbers, one could account for the mass and charge of any atomic nucleus. For example, the helium nucleus was known to have a mass four times that of the proton. Its charge, however, was known to be only twice the charge of a proton. Thus it seemed possible to model the helium nucleus as a combination of four protons (providing the necessary mass) and two electrons (reducing the total charge to two positive units without appreciably affecting the mass).

With the advent of quantum mechanics, certain difficulties arose with the proton-electron model of the nucleus. By 1930, there were several very strong arguments against the existence of electrons in the nucleus. Fortunately, experimental physicists about this time were beginning to turn their attention to the study of the nucleus. Their efforts were rewarded in 1932 by the discovery of a third elementary particle, the neutron.

In 1930, two German physicists, Walther Bothe and Hans Becker, discovered that very penetrating and uncharged radiation is emitted when the very light element beryllium is bombarded with alpha particles. Quite naturally, they assumed that this radiation was gamma rays. Later, Irene Joliot-Curie (Marie Curie's daughter) and her husband Frederic repeated these experiments and found that the uncharged radiation is capable of ejecting high-energy protons from paraffin (a hydrocarbon compound that is rich in hydrogen). Again, they assumed that the uncharged radiation is gamma rays, even though conservation of energy and momentum required that the gamma rays have photon energies of 55 MeV, much higher than any previously known.

James Chadwick (1891 – 1974), an associate of Rutherford, decided in 1932 to repeat the experiments in greater detail, using nitrogen as well as hydrogen as target for the uncharged radiation. Chadwick was able to measure the recoil energies of the nitrogen and hydrogen nuclei after the collisions. From conservation of energy and momentum, he was able to determine that the radiation consists of uncharged particles of mass slightly greater than the proton. Chadwick called these new particles neutrons. Chadwick's discovery of the neutron immediately led to a solution of the problems facing the proton-electron nuclear model. The arguments that electrons could not exist in the nucleus did not apply to neutrons. A proton-neutron nuclear model is completely consistent with all experimental and theoretical results. The helium-4 nucleus consists of two protons and two neutrons. The helium-4 nucleus is also known as an alpha particle.

14.2 QUARKS

The search for the basic constituents of matter began with the Greeks. Empedocles suggested that the material world was made up of four elements: earth, fire, water, and air and Democritus suggested that small, indivisible particles, atoms, were the fundamental building blocks of matter. By the 1930s, it was known that atoms were composed of protons, neutrons, and electrons. But by then physicists realized it was more complicated than that. The positron had been discovered, the neutrino had been postulated along with additional antiparticles. By the 1960s, many more supposedly "elementary particles" had been discovered. The hope that matter could be explained in terms of a small number of truly elementary particles was beginning to fade. Then in 1964, the concept of quarks was independently introduced by physicists Murray Gell-Mann and George Zweig

In the quark model, matter is divided into two categories: quarks and leptons. Electrons and neutrinos are leptons. Protons and neutrons are composite particles composed of quarks. To account for the hundreds of "elementary particles" know in 1964 and the many more discovered since then, six leptons and six quarks, along with their corresponding antiparticles, are needed. Like the electron and neutrino, quarks are point particles. That is, they have no extension in space. They also have fractional charges, $\pm \frac{1}{3}$ or $\pm \frac{2}{3}$ times the charge on an electron.

For example, a pair of quarks, the up quark and the down quark, are needed to account for the proton and neutron. A proton is composed of two up quarks ($+\frac{2}{3}e$) and one down quark ($-\frac{1}{3}e$) for a total charge of +1 e. A neutron is composed of two down quarks and one up quark for a total charge of zero.

In addition to the quarks and leptons, there are guage particles that mediate three of the four fundamental forces in nature – the strong nuclear force results from an exchange of gluons between quarks, the electromagnetic force from an exchange of photons, and the weak nuclear force from an exchange of W and Z bosons (see Figure 14.1)

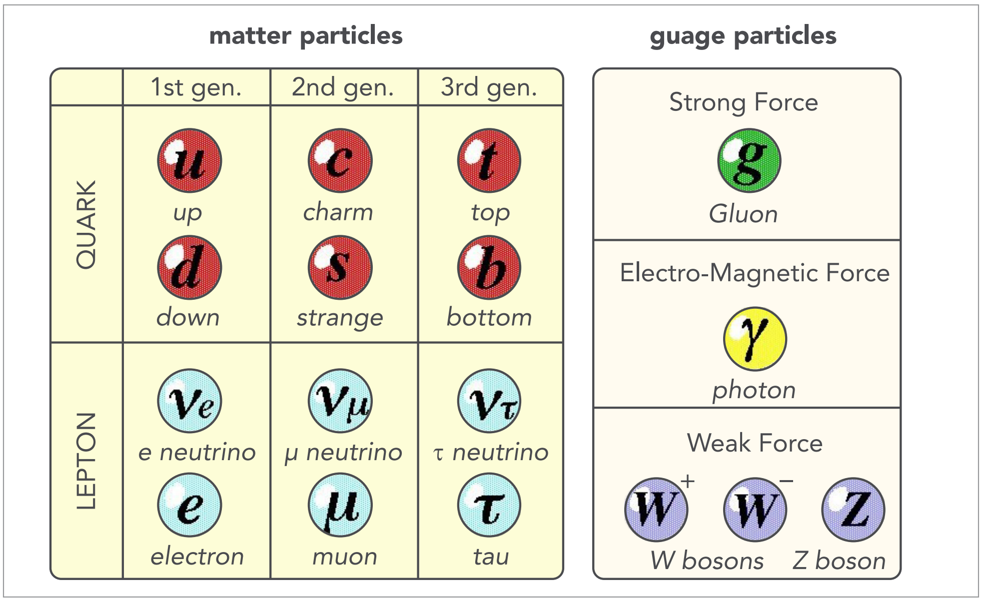


Figure 14.1 Quarks and leptons

14.3 OBSERVING SUBATOMIC PARTICLES

Science began with the study of those physical objects and events that are directly accessible to human senses. Twentieth-century physics, however, is characterized by the extension of science to those domains that lie outside the range of human perception, both in the large and the small scales. Because we obtain information through our senses, it has been necessary for experimental scientists to extend our natural senses with ingeniously designed instruments, so we can gather information about nature on the scale of the cosmos and on the scale of subatomic particles. The particles listed in Table 14.1 were discovered using instruments designed to respond in a macroscopic (human-scale) way to the submicroscopic interactions of the particles with the measuring instruments.

One of the earliest and most significant of these instruments is the Wilson cloud chamber. It was invented in 1896 by C. T. R. Wilson, who designed it as a tool to study the process of cloud formation. However, it has been used since 1911 as a basic tool in the study of elementary particles. When a charged particle such as an electron or a proton passes through the cloud chamber, it ionizes some of the molecules along its path. The temperature, pressure, and amount of water vapor in the cloud chamber are adjusted so that water vapor will condense along the ionized trail to make the path of the charged particle visible (See Figure 14.2.) This effect is similar to the one that produces a water-vapor trail behind a jet plane at high altitude.

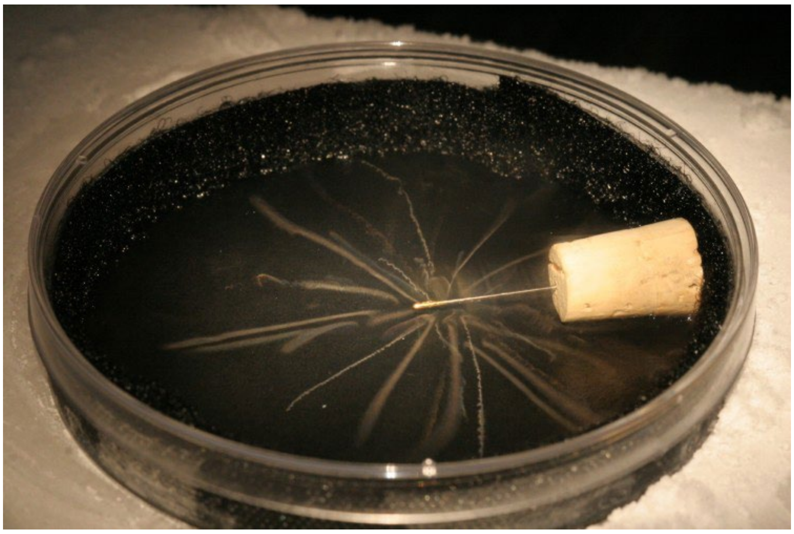


Figure 14.2 Cloud chamber

The tracks in the cloud chamber are photographed and studied to determine the properties of the subatomic particles that produced them. The cloud chamber usually is placed in a magnetic field, because a charged particle moving through a magnetic field experiences a force that causes it to move in a circular path whose radius is related to the kinetic energy and charge of the particle. Such properties of the particle as its charge, mass, and energy can be calculated from measurements of such track properties as thickness, length, and radius of curvature. Carl Anderson's 1932 discovery of the positron was made from a cloud-chamber photograph. The thickness of the track and its curvature in the magnetic field indicated a particle of the same mass as an electron but with the opposite charge.

An uncharged particle leaves no track in a cloud chamber. Nonetheless, the presence of an uncharged particle can be detected if it collides with a charged particle, causing the sudden track in the chamber. The properties of the uncharged particle can then be calculated from its effects upon the charged particle of known properties, using the laws of conservation of energy and momentum.

An even more efficient device for detecting elementary particles was invented in 1952 by Donald Glaser. In the bubble chamber, charged particles passing through a superheated liquid leave visible tracks of bubbles along their ionized paths. Glaser is said to have gotten the idea for the bubble chamber while watching bubbles rise in a glass of beer. (It's hard to tell when a physicist is working!)

14.4 NUCLEAR NOTATION

The nucleus of an atom is composed of neutrons and protons. The composition of the nucleus is described by an atomic number Z and a mass number A.

The atomic number Z is defined as the number of protons in the nucleus.

If the atom is neutral, then the charge +Ze on the nucleus must be balanced by an equal negative charge of Z orbital electrons surrounding the nucleus.

The periodic table (see Figure 14.3) is arranged in order of increasing atomic number, from hydrogen with Z = l to uranium with Z = 92. No element with Z1arger than 92 has ever been found in nature. However, in 1940, a nucleus with Z = 93 was produced in a laboratory reaction, and given the name neptunium. Since then, twelve other trans-uranium elements have been produced and identified, so the periodic table now contains 110 elements.

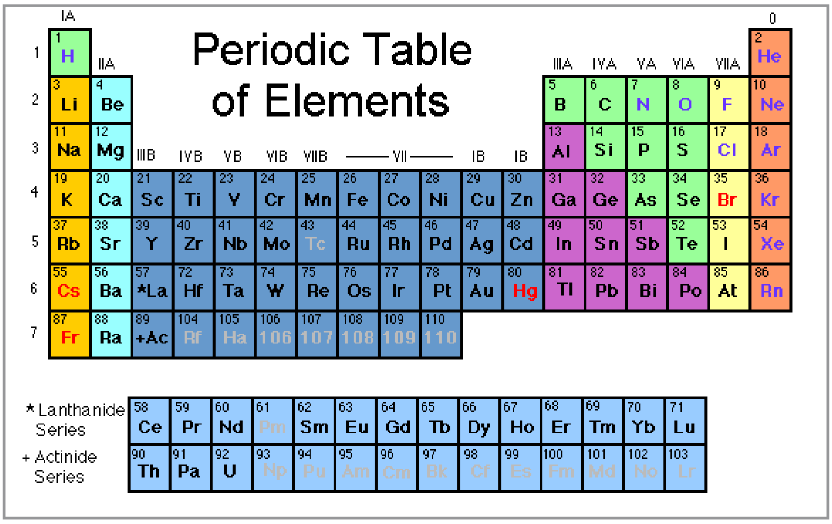


Figure 14.3 The periodic table.

In many ways, the neutron and the proton behave as if they are two different states of the same particle. For this reason, the word nucleon is often used to include both particles. Thus, the mass number can be defined as the number of nucleons in the nucleus. The mass of a nucleon (either a proton or a neutron) is approximately l u, The mass of an electron is very small compared to the mass of a nucleon (me = 0.00055 u). Therefore, the mass of a neutral atom in atomic mass units (u) is very close to the value A -- hence the name, mass number. The number, N, of neutrons in the nucleus obviously is just the difference between the mass number A and the atomic number Z.

The mass number A is defined as the number of neutrons plus the number of protons in the nucleus.

The notation used to specify a particular nucleus is $_ZX^A$, where X is the chemical symbol of the element. The atomic number Z appears as a subscript preceding the symbol and the mass number A appears as a superscript. For example, the helium nucleus (or alpha particle) is represented as $_2He^4$. The element fluorine (chemical symbol F) has Z = 9, so $_9F^{19}$ represents a fluorine nucleus with 19 nucleons -- 9 protons and 10 neutrons.

Not every possible combination of neutrons and protons forms a stable nucleus. In fact, most possible combinations prove to be unstable. That is, they break apart when they are formed. The neutron to proton ratio for stable nuclei increases gradually from one for the very light nuclei (generally Z less than 20) to approximately 1.5 for the heaviest nuclei.

Nuclei of the same element (the same atomic number) having different mass numbers are called isotopes of that element. For example, $_8O^{16}$, $_8O^{17}$, and $_8O^{18}$ are isotopes of oxygen. Similarly, $_{92}U^{235}$ and $_{92}U^{238}$ are isotopes of uranium. These isotopes are called uranium-235 and uranium-238. Isotopic identification is very important when discussing nuclear reactions. The number of neutrons in the nucleus has just as great an influence on the behavior of the nucleus as the number of protons.

The isotopes of an element are quite distinct from one another in regard to their nuclear properties. However, the isotopes behave almost identically in chemical reactions. The chemical properties of an atom depend solely on the number of orbital electrons, which in turn is usually equal to the atomic number of the nucleus. The only difference among isotopes in chemical reactions is a very slight difference in reaction rates due to the differences in mass (and hence in kinetic energy) among the isotopes. In practice, it is almost impossible to detect any difference in chemical properties among the isotopes of a single element. Therefore, we speak of chemical reactions involving oxygen, whereas when discussing nuclear reactions, we must specify that the reaction involves oxygen-l6 or oxygen-18.

14.5 THE FORCE BETWEEN NUCLEONS

The nucleus of a stable atom is very difficult to break apart into separate nucleons. The energy required to remove a nucleon from a stable nucleus is far greater than would be needed if the nucleons were held together only by gravitational forces. In fact, the protons and neutrons are held together by an extremely strong force that is called, logically enough, the strong nuclear force. The strong nuclear force has the following general properties:

1. Of the four forces we discuss in this book (gravitational, weak nuclear, electromagnetic, and strong nuclear) it is by far the strongest. The only charge in the nucleus is the positive charge of the closely clustered protons, so there is a large repulsive force acting to push the protons apart and disrupt the nucleus. The attractive nuclear force must be considerably stronger than this repulsive electric force because the stable combinations of protons and neutrons are very strongly bound together. The gravitational force between the nucleons is also is attractive, but its magnitude is completely negligible in comparison to the electric and nuclear forces.
2. The nuclear force is significant only over an extremely short distance. Two nucleons must be within a distance of 10^{-15} m from each other before the nuclear force between them becomes significant, whereas both the electric and the gravitational forces are significant at much larger distances. The diameter of the nucleus is on the order of 10^{-14} m, so the effect of the nuclear force does not extend very far beyond the nucleus itself.
3. The nuclear force is not related to charge. The nuclear force holding a pair of neutrons together is the same as the nuclear force holding a pair of protons or a neutron-proton pair together. Of course, the repulsive electric force decreases the net force holding together the pair of protons.
4. At very small distances, the nuclear force is repulsive. Nucleons separation at a distance of less than about 0.5×10^{-15} m will repel each other. Thus the nucleons in a nucleus are prevented from getting too close to one another. This feature of the nuclear force prevents the nucleus from collapsing down to the size of a single nucleon.

Coulomb's law is a simple and exact quantitative expression for the electric force that exists between two charged particles (see Chapter 12), There is no analogous quantitative expression for the nuclear force between two nucleons. The nuclear force is a much more complex force. Therefore, we can assume the nuclear force is not in itself a fundamental force. As an analogy, consider the extremely complex force between the atoms of a molecule or between the molecules of a solid or liquid. No simple force law can be created for such chemical forces either. However, the complex chemical force can be explained as a summation of the electric forces between all the electrons and nuclei of the individual atoms. Similarly, the complex nuclear force is the summation of more fundamental forces acting between the quarks comprising the nucleons.

14.6 THE DEUTERON

The simplest atom found in nature is the $_1H^1$ isotope of hydrogen, containing one proton as its nucleus and one orbital electron. However, hydrogen has two stable isotopes. The other is $_1H^2$, called heavy hydrogen, or deuterium. The deuterium atom contains one proton and one neutron bound together by the nuclear force and one orbital electron. The nucleus of a deuterium atom is called a deuteron.

It is clear that positive work must be done against the strong nuclear force in order to separate the deuteron into its component proton and neutron; in other words, energy must be added to the system, If the initial system is a deuteron at rest and the final system is a separated neutron and proton also at rest, then conservation of energy requires that the total mass of the neutron plus proton must be greater than the mass of the deuteron. Let ΔE be the energy added to separate the neutron and proton. Then conservation of energy requires that

$$\Delta E + m_d c^2 = m_p c^2 + m_n c^2$$

where m_d is the mass of the deuteron, m_p is the mass of the proton, and m_n is the mass of the neutron. We can rearrange this equation to write

$$\Delta E = (m_p + m_n - m_d)c^2$$

Because we know that ΔE is positive (we must do positive work on the system to separate the nucleons), we see that

$$m_p + m_n > m_d$$

To see if this prediction is valid, let us compare the experimentally measured masses of the deuteron, neutron, and proton,

$$m_d = 2.013553 \text{ u}$$
$$m_n = 1.008665 \text{ u}$$
$$m_p = 1.007276 \text{ u}$$

and, in fact, $m_p + m_n$ is greater than m_d

$$\Delta E = (1.007276 \text{ u} + 1..008665 \text{ u} - 2.013553 \text{ u})(931 \text{ MeV/u}) = 2.015941 \text{ u } (931 \text{ MeV/u})$$
$$\Delta E = 2.22 \text{ MeV}$$

That is, 2.22 MeV of energy must be added to the deuteron to separate the proton and neutron. If deuterium gas is irradiated with a beam of gamma rays, we would expect free neutrons and protons to begin to appear only when the energy of the individual photons in the beam reaches 2.22 MeV or more. Any photon energy exceeding 2.22 MeV should appear as kinetic energy of the separated neutron and proton. This experiment has been performed, and the results are exactly in accord with the prediction. Once again, Einstein's mass-energy relationship is confirmed by experiment.

Note the similarities of the experiment we have just described to the photoelectric effect. In the photoelectric effect, photon energies below a certain threshold value fail to produce free electrons from a metal surface. For photon energies above the threshold energy, free electrons are produced having kinetic energies provided by the excess photon energy above the threshold energy. The threshold energy is a measure of the work needed to overcome the electric force that holds the electrons to the metallic surface. In general, this threshold energy is on the order of several electron-volts. However, the proton and neutron are bound together in the deuteron by the much stronger nuclear force, so a much greater amount of work is needed to separate them. In general, threshold energies of several million electron-volts are required to overcome the strong nuclear force. It is only when such relatively large energies are involved that measurable changes in mass are produced.

14.7 MASS DEFECT AND BINDING ENERGY

Energy must be added to a nucleus in order to separate it into its component nucleons. Therefore, conservation of energy requires that the mass of the nucleus be less than the sum of the masses of the nucleons from which it is constructed. We have just seen one example of this with the deuterium nucleus. The difference in mass between the nucleus and its component nucleons is called the mass defect. If we use M to represent the mass of the nucleus, we can write

$$\Delta m = Zm_p + (A - Z)m_n - M.$$

Δm is called the mass defect of the nucleus.

In words, the mass defect equals the sum of the masses of the component nucleons minus the mass of the nucleus.

Example 14.1

Calculate the mass defect of $_3Li^7$, lithium-7.

Solution

The $_3Li^7$ nucleus contains 3 protons and (7 – 3) = 4 neutrons. From Appendix B, we find the mass of the $_3Li^7$ nucleus to be M = 7.014357 u.
Thus we can find

$$\Delta m = 3 \times (1.007276\ u) + 4 \times (1008665\ u) - (7.014357\ u)]$$
$$\Delta m = (3.021828\ u) + 4.034660\ u) - (7.014357\ u)]$$
$$\Delta m = 0.042131\ u$$

In order to separate a lithum-7 nucleus into its component nucleons, we must supply an energy equivalent to 0.042131 u of mass. From Einstein's mass-energy relationship,

$$\Delta E = (0.042131\ u) \times (931\ MeV/u) = 39.2\ MeV$$

We must add a minimum energy of 39.2 MeV to a lithium-7 nucleus in order to separate it into three protons and four neutrons. This energy is called the binding energy of the lithium-7 nucleus.

The binding energy can be thought of in any of the equivalent ways:

1. The binding energy is the energy equivalent to the mass defect: $E_b = \Delta mc^2$.
2. The binding energy is the energy that must be added to the nucleus in order to separate the nucleus into its component nucleons.
3. The binding energy is the energy released from the nucleus when the nucleus is formed from its component nucleons.

It should be clear that the binding energy of the nucleus must increase with increasing values of A. Because some work is required to remove each nucleon from the nucleus, a greater amount of work will be required to separate a greater number of nucleons. However, as we have mentioned, the nuclear force is a complex one. It turns out that the binding energy depends on several other factors in addition to the number of nucleons in the nucleus. As a result, the average binding energy per nucleon (the binding energy divided by the number of nucleons in the nucleus) is not constant. Figure 14.4 plots the average binding energy per nucleon against the mass number A. This plot will be very helpful later when nuclear reactions such as fission and fusion are discussed.

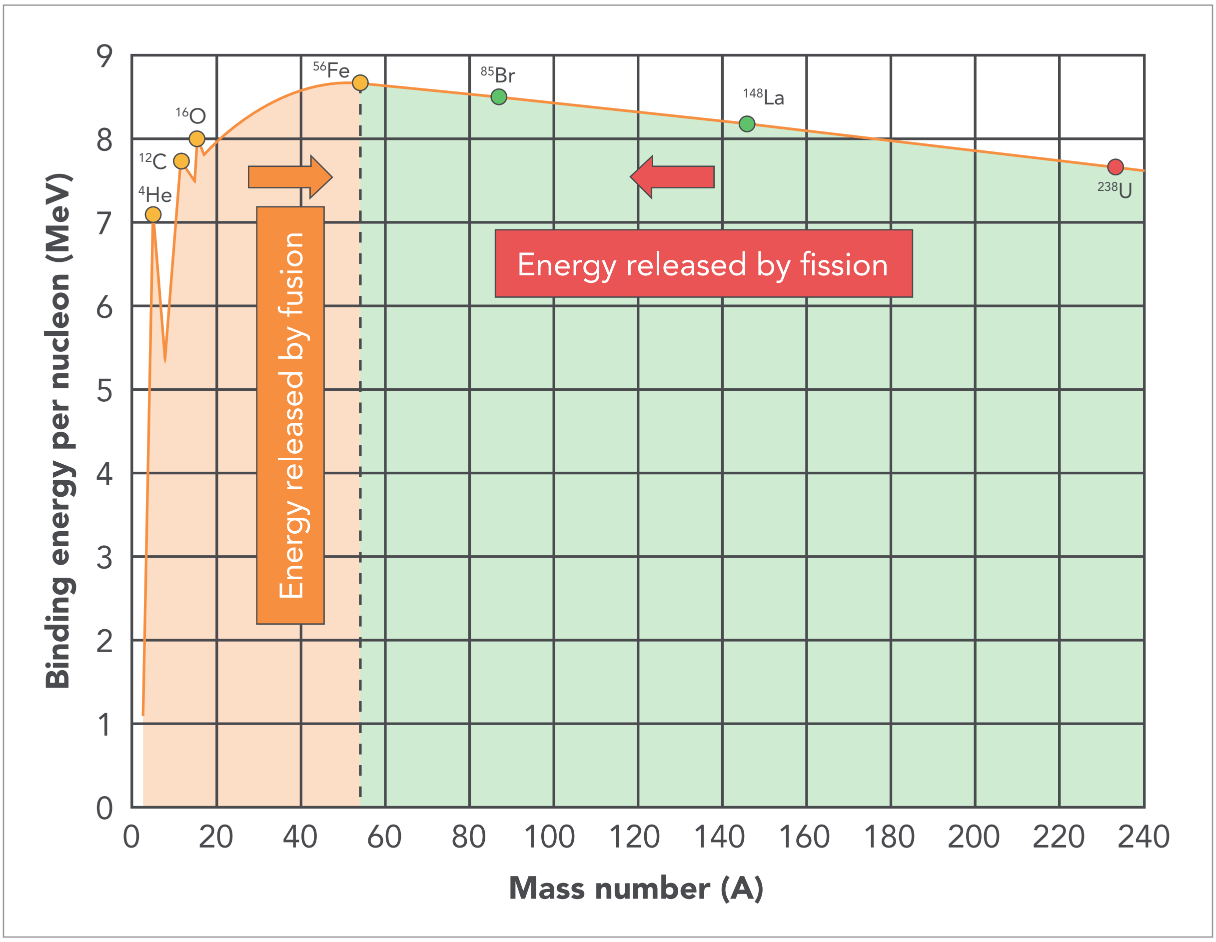


Figure 14.4 Average binding energy per nucleon

Summary

After the discovery of the neutron in 1932, it was clear that the atomic nucleus is composed of neutrons and protons. The gravitational force of attraction between these nucleons is far too weak to overcome the electric force of repulsion between the positively charged protons. Therefore, the existence of a strong nuclear force of attraction between nucleons was postulated to explain the stability of the atomic nucleus. From the beginning, it was realized that this strong nuclear force does not obey a simple force law like those that describe the electric and gravitational forces. It is now believed that the strong nuclear force between nucleons is actually the summation of the net effects of more fundamental forces exerted between the quarks of which the individual nucleons are composed.

Because the nuclear force is so strong, the positive work that must be done (energy that must be added) to separate nucleons is sufficiently great to produce a measurable change in the rest mass of the system. When a stable nucleus is broken down into its individual nucleons, the sum of the masses of the separate nucleons is measurably greater than the mass of the original nucleus. This difference in mass is called the mass defect of the nucleus. The energy required to separate the nucleus into its component nucleons is called the binding energy of the nucleus.

Important concepts

Neutron; cloud chamber; atomic number Z; atomic mass number A; nucleon; strong nuclear force; deuteron; mass defect Dm; binding energy E_b.

Questions

1. What is a nucleon?
2. Describe the properties of a neutron. How do they compare with the properties of the proton and the electron?
3. What is an atomic mass unit? How is the quantity l u defined?
4. Estimate the density of a nucleus. How do nuclear densities compare with the densities of ordinary substances such as water?
5. Explain how the Wilson cloud chamber is used to study elementary particles.
6. Explain the significance of the atomic number Z and the mass number A. What do these numbers represent in terms of nuclear and atomic structure?
7. What is an isotope? Explain how isotopes of the same element differ from one another. How are they similar?
8. List the numbers of neutrons, protons, and electrons in a neutral atom of each of the following isotopes: hydrogen-l, helium-4, iron-56, uranium-238 (You can refer to information at the back of the book to determine the atomic number of these elements.)
9. Explain why it is necessary to postulate the existence of a nuclear force that is neither gravitational nor electromagnetic in nature. Summarize the main properties of the strong nuclear force.
10. What is a deuteron? What is deuterium?
11. Explain why the mass of a nucleus is always less than the sum of the masses of its component nucleons. What happens to this missing mass?
12. Define and explain the term mass defect.
13. What is binding energy? How is it related to the mass defect?

14. Calculate the mass defect and the binding energy for each of the following nuclei: hydrogen-3, helium-3, neon-20.
15. Explain how you would calculate the average binding energy per nucleon for a particular nucleus.
16. Suppose we wish to use radioactive dating on sample whose age we believe is approximately t. Should we choose an isotope whose half-life is much greater than, approximately equal to, or much less than t? Why?

15 INDUCED NUCLEAR REACTIONS

Nuclear physics had its beginnings in 1896 when Henri Becquerel accidentally discovered the phenomenon of radioactivity. The experimental verification of Rutherford's nuclear model in 1913 made it clear that radioactivity is a nuclear phenomenon. We will look in detail at radioactivity in the following chapter.

15.1 INDUCED NUCLEAR REACTIONS

Radioactivity is a spontaneous reaction. A radioactive (or unstable) nucleus can emit radiation without any kind of disruption by external agents. However, in 1919 Rutherford discovered that he could produce an artificial nuclear reaction by bombarding nitrogen with alpha particles. This reaction led to the production of oxygen. Thus transmutation of elements can be accomplished either naturally by radioactivity or artificially by bombardment. The alchemists' dream has finally been realized. In general, an artificial nuclear reaction is produced by bombarding one nucleus with another nucleus or elementary particle. Such a nuclear reaction is called an induced nuclear reaction. It typically results in the production of two or more nuclei, or some combination of nuclei and elementary particles. Using ${}_{Z}X^{A}$ to represent an arbitrary nucleus or elementary particle, we can write a general equation for an induced nuclear reaction as:

$${}_{Z1}X_1^{A1} + {}_{Z2}X_2^{A2} \rightarrow {}_{Z3}X_3^{A3} + {}_{Z4}X_4^{A4} + ,,,,$$

The term reactants is used to refer to the nuclei or other particles that exist before the nuclear reaction takes place (${}_{Z1}X_1^{A1} + {}_{Z2}X_2^{A2}$ in the general equation). The term products refers to the nuclei or other particles that exist after the reaction takes place (${}_{Z3}X_3^{A3} + {}_{Z4}X_4^{A4} + ,,,,$). Radioactive reactions occur spontaneously, so there is only a single reactant in that case. Two reactants are involved in an induced nuclear reaction. In most low-energy nuclear reactions there are two product nuclei.

As an example, consider the induced nuclear reaction in which nitrogen is transmuted to carbon through bombardment by a deuteron.

$${}_{1}H^{2} + {}_{7}N^{14} \rightarrow {}_{6}C^{12} + {}_{2}He^{4}$$

The reactants are hydrogen-2 and nitrogen-l4, and the products are carbon-12 and helium-4: The helium-4 produced in this reaction is identical to that produced by alpha-decay. However, this induced reaction is not called alpha decay; that term is used only to describe the spontaneous (radioactive) emission of an alpha particle.

The laws of conservation of charge, energy, and momentum must be obeyed in a nuclear reaction. In addition, there are several other conservation laws that apply only to nuclear physics. One of these is the law of conservation of nucleons. That is, the total number of nucleons must be the same before and after a nuclear reaction; nucleons can be neither created nor destroyed. In some reactions, a neutron can change into a proton, or vice versa, but the total number of nucleons remains constant.

In a nuclear reaction, it is very unlikely that the rest mass of the reactants will exactly equal the rest mass of the products. If the rest-mass energy changes during a nuclear reaction, then conservation of energy requires that there must be a corresponding change in the kinetic energy and/or radiant energy of the system. The physical quantity associated with this change is called the Q of the reaction. It is defined as the difference between the rest-mass energy of the reactants and the rest-mass energy of the products. If M_{re} is the rest mass of the reactants and M_{pr} is the rest mass of the products, then

$$Q = (M_{re} - M_{pr})c^2$$

If the mass of the reactants is greater than the mass of the products, then the reaction is said to be exoergic. In an exoergic reaction, some rest-mass energy is converted to kinetic energy and/or radiant energy. The induced reaction considered earlier in which nitrogen-14 is bombarded by a deuteron is an example of an exoergic reaction:

$$_1H^2 + {}_7N^{14} \rightarrow {}_6C^{12} + {}_2He^4$$

The masses of reactants and products are

$$M_{re} = (2.013553 \text{ u}) + (13.999231 \text{ u}) = 16.012784 \text{ u}$$
$$M_{pr} = (11.96706 \text{ u}) + (4.001505 \text{ u}) = 15.998211 \text{ u}$$

Therefore, $Q = (16.012784 \text{ u} - 15.998211 \text{ u} \times (931 \text{ MeV/u}) = +13.57 \text{ MeV}$

The positive value for Q means that the rest-mass energy of the system decreases as a result of the reaction. In this case, it decreases by 13.57 MeV, and, therefore, the kinetic energy of the products is increases by 13.57 MeV.

If the mass of the reactants is less than the mass of the products, then Q is negative, and the reaction is said to be endoergic. In an endoergic reaction, some radiant or kinetic energy is converted to rest-mass energy. This energy must be supplied by the bombarding particle or photon, and therefore a spontaneous nuclear reaction (radioactive decay) cannot be endoergic.

As an example of an endoergic reaction, consider the reverse of the transmutation reaction we have just been discussing: Instead of bombarding nitrogen-14 with a deuteron, we are now bombarding carbon-l2 with an alpha particle.

$${}_{2}He^{4} + {}_{6}C^{12} \rightarrow + {}_{7}N^{14} + {}_{1}H^{2}$$

Using the masses from above, the Q for this reaction is -l3 57 MeV. This means that 13.57 MeV of kinetic energy must be converted to rest-mass energy in the reaction. If the bombarding alpha particle has less than 13.57 MeV of kinetic energy, then this reaction cannot take place. In fact, the alpha particle must have more than this minimum amount because momentum will not be conserved if all of the kinetic energy is converted to rest-mass energy. More complicated calculations based on conservation of momentum would be required to find the actual threshold energy for this reaction.

Example 15.2

Calculate the Q for the following reaction, which involves magnesium (Mg) and sodium (Na):

$$_0n^1 + {}_{12}Mg^{24} \rightarrow + {}_{11}Na^{24} + {}_1H^1$$

The symbol $_0n^1$ represents a neutron. (The element sodium is represented by the symbol Na because it was commonly known by its German name natrium when the system of chemical symbols was developed.)

Solution

$$M_{re} = (1.008665 \text{ u}) + (23.978495 \text{ u}) = 24.987160 \text{ u}$$
$$M_{pr} = (23.984886 \text{ u}) + (1.007276 \text{ u}) = 24.992142 \text{ u}$$

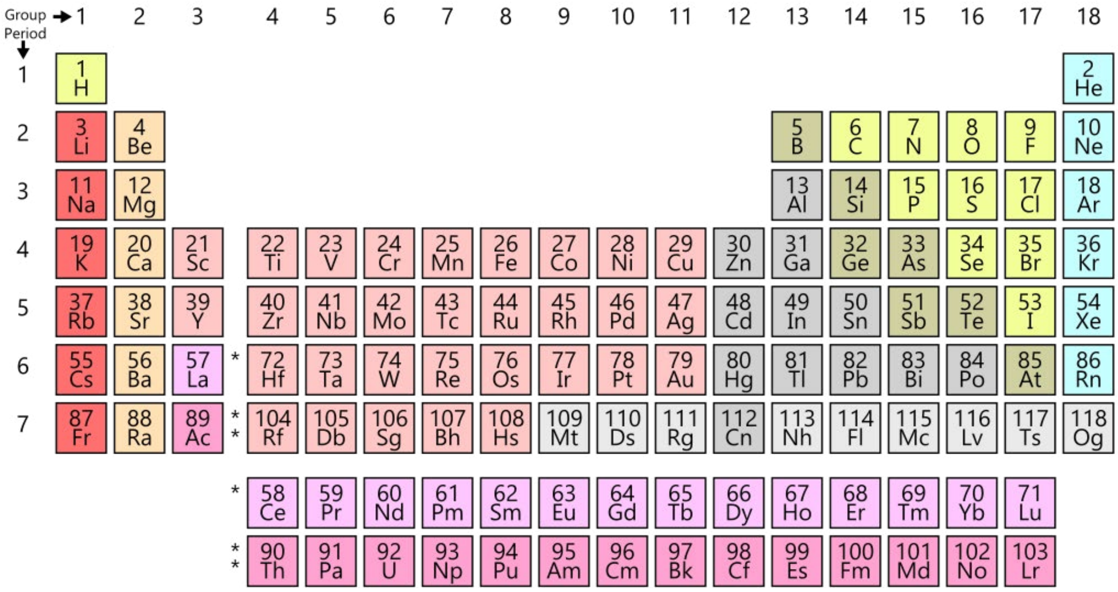
The mass of the products is greater than the mass of the reactants, so the reaction is endoergic and Q is negative. We find

$$(M_{re} - M_{pr}) = (24.987160 \text{ u}) - (24.992142 \text{ u}) = -0.004932 \text{ u}$$
$$Q = (-0.004932 \text{ u}) \times (931 \text{ MeV/u}) = -4.64 \text{ MeV}$$

The rest-mass energy of the system increases by 4.64 MeV. This energy (plus a little more to conserve momentum) must be supplied by the kinetic energy of the neutron in order for this reaction to occur.

Origins of the Elements

In your high school chemistry class, you learned about the Periodic Table of the Elements, elements ranging from hydrogen, Z = 1, whose atoms have the lightest atomic weight to uranium, Z = 92, with atoms having the greatest atomic weight found naturally on earth, and beyond to the transuranic elements produced in the laboratory.



But how did these elements come to be in our universe. The answer is induced nuclear reactions that took place long before our existence and allowed for the possibility of our existence.

Early in the history when the universe was extremely hot, the only elements that existed were hydrogen and helium, the two lightest elements. When the universe expanded and became sufficiently cool to allow electrons to combine with the hydrogen and helium nuclei, gravity, the weakest force in nature, began to assemble the hydrogen and helium gas into stars and galaxies. The gravitational contraction of the stars reheats the matter, especially in the center of the star. The temperature in the cores of stars became sufficiently high for the proton to combine (converting some of them to neutrons in the process) to produce more helium nuclei. Nucleosynthesis begins again where it had left off in the early universe. When the hydrogen fuel is depleted, the core contracts and heats up allowing helium to fuse into carbon.

In the more massive stars, the process continues beyond carbon. Carbon is fused into even heavier elements. When the carbon is exhausted, the core contracts and heats -- still heavier elements are produced. This sequence is repeated again and again. Each successive round of fusion produces heavier nuclei. Finally, the star develops an iron core. surrounded by layers of the other elements produced over its lifetime. More or less like an onion.

The iron core however is the beginning of the end for the star. The death of a high mass star is called a supernova, an explosion of unimaginable proportions. The supernova hurls the elements created in the star plus elements created by the explosion itself (elements ranging from iron to uranium and beyond) out into the space between the stars. There it mixes with the hydrogen and helium gas left behind in the initial phase of star formation. The interstellar matter is said to be enriched. In this way the concentration of elements heavier than helium slowly increases with time. This enriched interstellar matter, with its carbon, oxygen, iron and other heavier elements, is the raw material from which later generations of star systems, such as our solar system, and in fact we ourselves formed.

15.2 FISSION

In 1934, Enrico Fermi (1901 – 1954) and his students began systematically bombarding the various elements with neutrons and observing the results. In general, the result of neutron bombardment of a heavy element was its transmutation to a still heavier element. Therefore, Fermi was not particularly surprised when neutron bombardment of uranium, (the element of highest atomic number (Z = 92) then known to exist, resulted in reaction: products whose chemical properties did not match those of any of the known heavy elements. Preliminary calculations suggested that the product was a new element with Z = 93. However, before Fermi was satisfied with this conclusion sufficiently to publish it, the word leaked to the Italian press. The discovery of a new "artificial" element was trumpeted as a great "fascist victory." During the following four years, similar experiments were conducted in Paris, Cambridge, Zurich, and Berlin --in each case, with an unidentified substance produced through the bombardment of uranium.

Most scientists at the time were certain that the products of an induced nuclear reaction involving uranium would include a heavy nucleus with atomic number slightly larger than 92, but there was one notable exception. The German chemist Ida Noddack stated in I934 that

> "It would be equally possible to assume that when a nucleus is demolished in this novel way by neutrons, nuclear reactions occur which may differ considerably from those hitherto observed. .. . It would be conceivable that . . . the nuclei in question might break into a number of larger pieces which would no doubt be isotopes of known elements but not neighbors of the elements subject to radiation."
> Quoted in Ruth Howes and Caroline Herzenberg, Their Day in the Sun: Women of the Manhattan Project.

Two years later, her husband asked a noted nuclear scientist why he never referenced her work in his lectures and publications. The scientist replied that he did not want to make her look ridiculous by calling attention to her absurd speculation.

Then in the fall of 1938, the German chemist Otto Hahn, at the urging of his assistant Fritz Strassman, began a careful study of earlier bombardment experiments done in Paris by Irene Joliot-Curie. The chemical evidence unmistakably indicated that barium (a medium-sized nucleus) is one of the elements produced when uranium is bombarded by neutrons. Hahn was baffled. Where does the barium come from? Noddack's speculation had long been forgotten, and Hahn was too surprised by his results to have enough confidence to publish them. However, he did send an account of them to his former colleague, Lise Meitner, who had fled from Germany to Sweden earlier that year because of her Jewish ancestry. Meitner's nephew Otto Frisch, a physicist at Bohr's institute, was visiting for the Christmas holidays when Hahn's letter arrived. They soon realized the true significance of the experimental results, and together Meitner and Frisch worked out a preliminary interpretation involving the splitting, or fission, of the uranium nucleus.

In their published report, Meitner and Frisch showed that the nuclear fission of uranium should be a highly exoergic reaction with a Q of approximately 200 MeV. Other physicists were quick to recognize the awesome significance of obtaining such huge amounts of energy from a relatively simple induced nuclear reaction. Thus began the chain of events that would eventually lead to the atomic bomb.

A typical fission reaction is the following:

$$_{0}n^{1} + _{92}U^{235} \rightarrow _{41}Nb^{99} + _{51}Sb^{134} + 3\ _{0}n^{1} + \text{gamma rays}$$

Example 15.3

Calculate the Q for the fission reaction given above.

Solution

Using the rest masses from Appendix B, the rest mass of the reactants is

$$M_{re} = 1.008665 \text{ u} + 234.993393 \text{ u} = 236.002058 \text{ u}$$

The rest mass of the products is

$$M_{pr} = 98.888648 \text{ u} + 133.893191 \text{ u} + (3 \text{ x } 1.008665 \text{ u}) = 235.807834 \text{ u}$$

$$Q = (M_{re} - M_{pr})\ c^2 = (236.002058 \text{ u} - 235.807834 \text{ u}) \text{ x } 931\text{Mev/u} = 181 \text{ MeV}$$

This is a tremendous amount of energy for a single nuclear reaction. Looking back at Figure 14.3 at the binding energy per nucleon curve, splitting a very heavy nucleus into two medium sized nuclei greatly increases the binding energy per nucleon which is then released in the reaction.

The Fission Chain Reaction

The report by Meitner and Frisch on the discovery of nuclear fission by Hahn and Strassmann was published in Sweden in January 1939. In early February, Leo Szilard (a Hungarian physicist who had fled from Nazi Germany to the United States) wrote to the Joliot-Curies in France:

> When Hahn's paper reached this country about a fortnight ago, a few of us at once got interested in the question whether neutrons are liberated in the disintegration of uranium. Obviously if more than one neutron were liberated, a sort of chain reaction would be possible. In certain circumstances this might then lead to the construction of bombs which would be extremely dangerous in general and particularly in the hands of certain governments.
> Quoted in Robert Jungk, *Brighter than a Thousand Suns.*

One of the "interested few" in the United States was Enrico Fermi. He had traveled to Stockholm in 1938 to accept the Nobel Prize in physics. Rather than return to Fascist Italy, with the help of Niels Bohr, he fled with his family (his wife was Jewish) to the United States. Together, Fermi and Szilard soon discovered an average of 2.5 neutrons are emitted in each fission of a uranium-235 nucleus. The possibility of constructing an explosive device appeared very real indeed.

By the summer of 1939, Szilard and his colleagues had received two disturbing items of confidential information from Germany. First, they learned that physicists in Germany were already working on the problem of uranium fission with the knowledge and support of the Nazi government. Second, they learned that the Germans had suddenly forbidden all exports of uranium ore from Czechoslovakia, which they had recently occupied. This information seemed to confirm the scientists' worst fears. Apparently, the Third Reich was actively involved in a program to build nuclear weapons.

The physicists drafted a letter to President Roosevelt, warning of the danger and encouraging government support for atomic research. Albert Einstein (who had left Germany in 1932) was convinced to add his prestige by signing the letter, although he was not involved in the research in any way. The letter was delivered in October 1939, and almost immediately the

United States embarked on the Manhattan Project, a top secret, crash program to develop an atomic bomb. On December 2, 1942, Fermi and his colleagues (in a squash court under the University of Chicago stadium) achieved the first self-sustaining controlled release of nuclear energy.

In the spring of 1943, scientists began assembling in Los Alamos, New Mexico, for the actual construction of atomic bombs. A wide variety of practical problems had to be solved in the design of the weapons and in the purification of necessary quantities of uranium-235 and plutonium-238, an artificially produced element (Z = 94) that also undergoes fission reactions suitable for the production of a chain reaction. Under the direction of the American physicist J. Robert Oppenheimer (1904 – 1967), a group of the leading scientists in America and brilliant young newcomers to the scientific community proved remarkably adept at solving theoretical and practical problems. On July 16, 1945, the first atomic bomb was exploded in the New Mexico desert.

The motivating force behind most scientists' participation in the atomic-bomb project was the belief that the Germans were on the verge of developing such weapons themselves. The thought of Adolf Hitler having such weapons at his disposal was terrifying. Many of the top scientists working on the project were European refugees who had experienced the horrors of Nazism at first hand. Their moral qualms about helping to create such an awesome weapon were more than offset by the dread of what Hitler could do to the world if he alone possessed it.

Toward the end of the war, it became increasingly clear the Germans had failed to develop nuclear weapons. In fact, the German project had never progressed beyond some preliminary research into the possibilities of such a weapon. Szilard and many of the other scientists who had helped to develop the bomb now turned their efforts toward the prevention of its use. To the bitter disappointment and horror of Szilard and others, on August 6 and August 9, 1945, the cities of Hiroshima and Nagasaki were destroyed by nuclear weapons.

Today many countries use controlled fission chain reactions to generate electricity Opponents of the nuclear power industry point out that the reactors are generating tons and tons of nuclear waste – radioactive fissions fragments that will be toxic for tens of thousands of years. There is some hope that one day nuclear reactors and coal burning power plants will replaced by plants utilizing fusion as their source of energy.

15.3 FUSION

From as far back as the Greeks, humans have tried to explain the tremendous energy released by the sun. In 1929, the correct solution was finally found. In the hot interior of stars, light nuclei are fuse and energy is produced. It wasn't until 1938, however, that the exact nature of the thermonuclear reactions was determined – the fusion of hydrogen into helium. This occurs through a sequence of reactions, but the net effect is

$$4\ ({}_1H^1) \rightarrow {}_2He^4 + 2\ ({}_{+1}\beta^0) + 2(\ {}_0\nu^0)$$

That is, four protons are fused into one helium-4 nucleus with the emission of two positrons (to convert two of the protons into neutrons) and two neutrinos (to be discussed in the following chapter).

Example 15.4

Calculate the Q for the fusion reaction given above.

Solution

Using the rest masses from Appendix D, the rest mass of the reactants is

$$M_{re} = 4 \text{ x } 1007276 \text{ u} = 4.29104 \text{ u}$$

The rest mass of the products is

$$M_{pr} = 4.001505 \text{ u} + (2 \text{ x } 0.000549 \text{ u}) = 4.002603 \text{ u (neutrinos have negligible mass)}$$

$$Q = (M_{re} - M_{pr})\ c^2 = (4.029104 \text{ u} - 4.002603 \text{ u}) \text{ x } 931\text{Mev/u} = 24.7 \text{ MeV}$$

Looking back at Figure 14.3, the binding energy per nucleon curve, combining very light nuclei increases the binding energy per nucleon, which is then released in the reaction.

Nuclear fusion releases about ten times more energy per kilogram of fuel than dose fission, and about 70 million times more than combustion. Furthermore, the hydrogen needed as fuel can readily be obtained from water, and the fusion reaction produced far less radioactive waste than does fission. All compelling reasons for trying to design power plants that use fusion as their energy source. The extremely high temperatures required for fusion to occur (the protons repel each other and must be moving at very high speeds in order to initiate the reaction), however, poses formidable problems to making fusion a cost-effective source of energy.

Summary

Reactions involving the nuclei of atoms are of two types. An unstable nucleus can spontaneously disintegrate by ejecting one or more particles. This type of nuclear reaction is called radioactive decay and is the subject of the next chapter. A second type of nuclear reaction is an induced nuclear reaction that may occur when one nucleus is bombarded with another nucleus or with an elementary particle.

In general, the rest mass of the system after a nuclear reaction (the sum of the rest masses of the products) is not equal to the rest mass of the system before the reaction (the sum of the rest masses of the reactants). The change in the rest mass of the system, expressed in energy units, is called the Q of the reaction. In an exoergic reaction, Q is positive and rest-mass energy is converted to kinetic and/or radiant energy during the reaction. In an endoergic reaction, Q is negative and kinetic and/or radiant energy is converted to rest-mass energy.

Nuclear fission is the splitting a very heavy nucleus by bombarding it with a neutron. This is an extremely exoergic reaction. Nuclear fusion is the combining of very light nuclei into a single heavier nucleus and is the source of energy for our sun and the other stars. The net effect of fusion in the sun is the combining of four protons (hydrogen nuclei) into one helium-4 nuclei.

Important concepts

Spontaneous nuclear reaction; induced nuclear reaction; reactants; products; exoergic reaction; endoergic reaction, fission, fusion.

Questions

1. List the physical quantities that are conserved in any nuclear reaction.
2. Explain the significance of the sign (plus or minus) of the Q of a nuclear reaction.
3. What is the difference between an endoergic reaction and an exoergic reaction?
4. Suggest a possible nuclear reaction in which mercury-200 would be transmuted into gold-197.
5. Explain why every naturally occurring radioactive-decay reaction must be exoergic.
6. Why is a neutron a more useful particle for bombardment experiments than either a proton or an alpha particle?
7. Calculate the Q for the following nuclear reaction:
   ${}_1H^2 + {}_1H^3 \rightarrow {}_2He^4 + {}_0n^1$
8. Explain why splitting a very heavy nucleus and combining very light nuclei both give off energy.
9. Why were physicists confused at first about the product nuclei in the reaction where uranium was bombarded with neutrons?
10. Estimate the number of fissions per second that must occur in a 1000 mega-Watt power plant, assuming a 30 percent efficiency of energy conversion.
11. Explain why fusion reactors require extremely high temperatures.
12. Write all possible reactions resulting from the absorption of a neutron by a helium-4 nucleus.
13. Assume that U-235 splits into two fragments with mass numbers 90 and 145 with each fragment having roughly the same ratio of Z/A as U-235. Explain why neutrons are emitted in fission.
14. Why is it necessary for the first reaction in the hydrogen fusion chain to result in the conversion of a proton into a neutron?

Enrico Fermi (1901 – 1954)

Fermi received his doctor's degree at the University of Pisa in 1922, just a few months before Benito Mussolini seized power in Italy. By 1926 he was a professor of physics at the University of Rome.

Fermi grew interested in the neutron, as soon as it was discovered in 1932. In a lecture explaining the difference between Chadwick's neutron and Pauli's neutron, he called Pauli's neutron the neutrino (the little neutron in Italian). The name stuck.

When a neutron is absorbed by the nucleus of a particular atom, the new nucleus sometimes becomes an atom of the next higher element. In 1934, he bombarded uranium with neutrons hoping to form an artificial element above uranium in the periodic table. At first, he thought he had done this. Much to his embarrassment, this was announced by the Fascist press before Fermi had a chance to confirm this.

These were hard times for the Fermis. Fermi was anti-Fascist and his wife was Jewish. Their chance to escape came in 1938 when they were allowed to travel to Stockholm for him to receive the Nobel Prize in physics.

They didn't return to Italy. After receiving the award, he and his family sailed to the United States and remained there until his death in 1954. Fermi worked on the Manhattan project and, on December 2, 1942, was the first to construct a nuclear reactor capable of a sustained nuclear fission chain reaction. It was announced among those in the know by a cryptic telegram, "The Italian navigator has entered the new world."

In 1945, Fermi accepted a professorship at the Institute for Nuclear Studies at the University of Chicago. He died young of stomach cancer. Element 100, discovered the year after his death, was named fermium in his honor.

16 RADIOACTIVITY

The phenomenon of radioactivity was discovered in 1896 by the French physicist Henri Becquerel (1852 – 1908). In January of that year, Becquerel learned of an amazing discovery made by the German physicist Wilhelm Roentgen (1845 – 1923). While doing research with cathode-ray tubes, he noticed when the cathode rays struck glass, they cause the glass to emit visible light. This phenomenon is called fluorescence. What Roentgen discovered is that, in addition to the visible light, the fluorescent areas of the glass also emit an unexpected, extremely penetrating radiation. Because the nature of this radiation was unknown, Roentgen simply called it X-rays. News of this mysterious radiation spread rapidly, and physicists all over the world began to study the properties of X-rays. The popular press splashed the story over the front pages, particularly after it was discovered that the X-rays can be used to photograph the bones inside the living body.

When Becquerel learned about Roentgen's discovery, he immediately set out to try to discover whether the X-rays are simply a peculiar feature of cathode-ray tubes or whether they are associated with fluorescence in general. Becquerel knew that certain minerals will glow (fluoresce) when irradiated with ultraviolet light, so he set out to discover whether X-rays also are associated with this fluorescence.

Becquerel carefully wrapped a photographic plate with black paper to block visible and ultraviolet light and placed the fluorescent mineral on the wrapped plate. He then irradiated the mineral with ultraviolet light to cause fluorescence. Then he developed the photographic plate to see whether any X-rays had penetrated the black paper to expose the film. His early experiments produced no exposure of the film. Then he happened to use some fluorescent uranium minerals, which did cause exposure of the film, indicating that a very penetrating radiation was emitted by the fluorescing uranium minerals. Naturally, Becquerel assumed that this radiation was X-rays. One day, Becquerel happened to develop some photographic plates that had been left in a drawer with samples of the uranium minerals. The plates had been wrapped in black paper, and the minerals had not been exposed to ultraviolet light, so there was no reason to expect any exposure of the plates. Fortunately, Becquerel for some reason decided to check this straightforward prediction. To his surprise, the photographic plates had been exposed by the non-fluorescing uranium minerals. Subsequent experiments showed that the uranium minerals spontaneously emit the penetrating radiation, whether or not they are fluorescing.

Henri Becquerel
(1852 – 1908 * France)

1896 – discovered that uranium spontaneously emitted radiation, later identifies as alpha radiation.

Thesis advisor to Marie Curie.

1903 - Nobel Prize in Physics.

Figure 16.1 Henri Becquerel

Further research showed that the radiation is emitted by the uranium atoms in the minerals. Any sample of uranium spontaneously emits this radiation without any external energy supplier. This phenomenon is quite different from that observed by Roentgen where X-rays are emitted only when glass is bombarded by cathode rays (which were shown in 1897 to be streams of high-energy electrons). Becquerel's new phenomenon was named radioactivity. A substance that emits this spontaneous radiation is said to be radioactive.

Because they did not produce pictures of bones, Becquerel's rays did not cause the sensation that had accompanied Roentgen's discovery. However, near the end of 1897, Becquerel's rays attracted the interest of a brilliant young Polish-born French physicist, Marie Sklodowska Curie (1867 – 1934). She found another radioactive element, thorium. More importantly, she discovered that a uranium mineral called pitchblende emits radiation at a rate far too great to be explained by the amount of uranium in the mineral. Curie concluded that pitchblende must contain small quantities of some unknown substance far more radioactive than either uranium or thorium. After a long and tedious series of chemical separations, Marie Curie and her husband Pierre succeeded in isolating small quantities of two new, highly radioactive elements. These were named polonium (in honor of Poland) and radium (because of the great intensity of radiation emitted by this substance). The study of radioactivity soon became an important area of research in physics, providing many clues about the nature of the atomic nucleus. For their contributions to this research, Marie and Pierre Curie shared the 1903 Nobel Prize in physics with Henri Becquerel. Marie Curie also won the 1911 Nobel Prize in chemistry for her discovery of the two new elements, thus becoming the first person to receive two Nobel Prizes.

16.1 THE RADIOACTIVE DECAY LAW

Certain naturally occurring isotopes are radioactive. That is, they are unstable. After some time interval, a nucleus of such an isotope will decay by ejecting some particle or particles. As Becquerel discovered, the isotopes of uranium are radioactive. It has since been found that all naturally occurring isotopes with atomic number greater than 82 (and a few with smaller atomic numbers) are radioactive. In addition to the 66 naturally occurring radioactive isotopes, a tremendous number (around 1500) of radioactive isotopes have been produced artificially in the laboratory by nuclear bombardment.

The law of radioactive decay was first formulated in 1902 by Ernest Rutherford and the English chemist Frederick Soddy. They found that the fraction of radioactive nuclei decaying in a given time interval is constant for any sample of a particular radioactive isotope. (The larger this fraction, the more unstable or radioactive the isotope is considered to be.) For instance, suppose that one-sixth of the nuclei in a sample of a particular radioactive substance decay during a period of one year. Then, during the following year, one-sixth of the remaining nuclei will decay. As long as a statistically significant number of atoms remain, one-sixth of the radioactive nuclei present will decay in any one-year period. Obviously, the rate at which decay reactions occur will gradually decrease as the number of radioactive nuclei decreases. Eventually, all of the radioactive nuclei will have decayed, and the sample will no longer be radioactive.

A useful way to discuss radioactivity quantitatively is in terms of the half-life.

> **The half-life of a particular radioactive isotope is defined as the time required for one-half of the atoms initially present in a sample to decay**

As an example, consider the radioactive isotope sodium-24, which has a half-life of 15 hours. If we start at time t = 0 with an arbitrary number N_o of sodium-24 atoms, then $N_o/2$ sodium-24 atoms will remain after 15 hours. Over the next 15-hour interval, one-half of these remaining atoms will decay. Thus, at t = 30 hours, the number of remaining sodium-24 atoms will be $N_o/4$. Similarly, at t = 45 hours, there will be $N_o/8$ of the radioactive atoms left. Figure 15.1 graphs this relationship. Sodium-24 decays to form magnesium-24, so that the number of magnesium-24 atoms increases at the same rate that the number of sodium-24 atoms decreases. The total number of atoms in the sample remains unchanged.

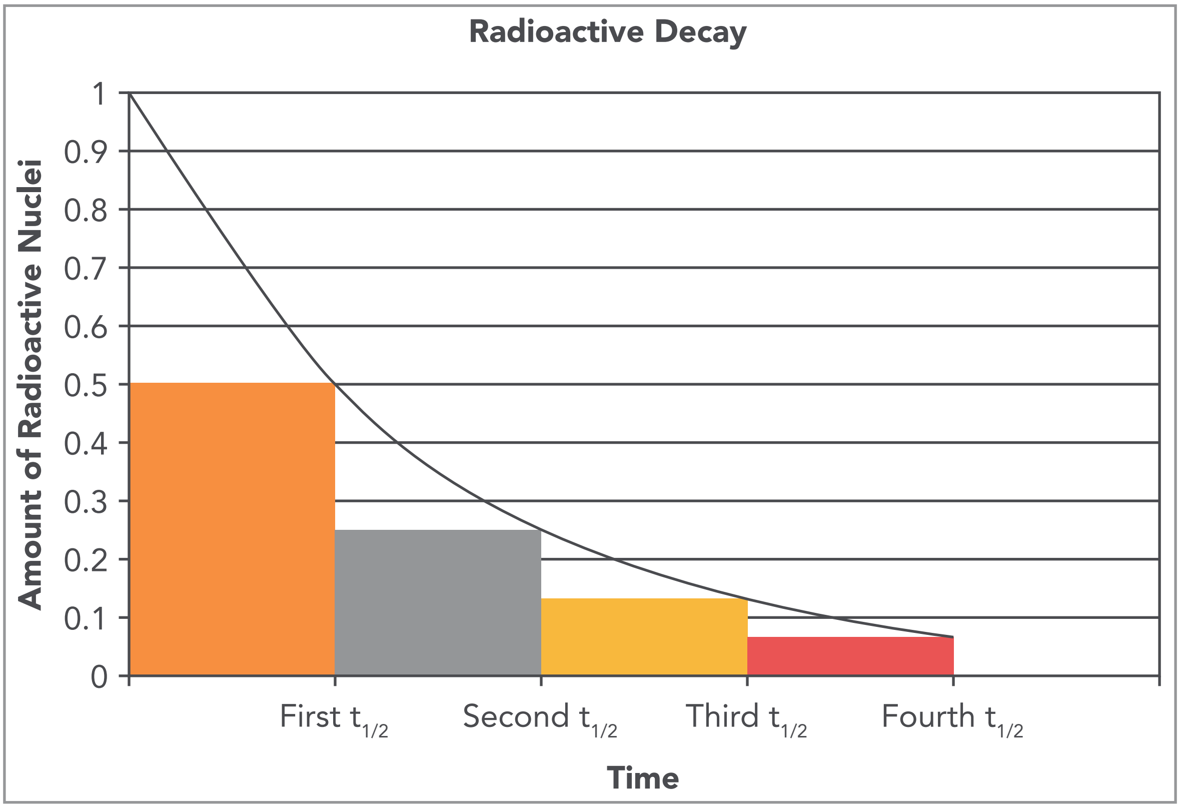


Figure 16.2 Radioactive decay law.

A smooth curve has been drawn through the points on the graph in Figure 16.2. This curve represents the number of radioactive atoms present as a function of time. This curve turns out to represent the mathematical relationship

$$N = N_o \exp(-0.693\ t/t_{1/2})$$

where N_o is the number of radioactive atoms present at time t = 0, N is the number present at time t, and $\exp(-0.693\ t/t_{1/2})$ represents the mathematical constant, e, that is the base of the natural logarithm that is, e raised to the $(-0.693\ t/t_{1/2})$ power. The graph of Figure 16.2 is said to be exponentially decreasing with time. Appendix B lists the half-lives of some radioactive isotopes.

Example 16.1

A sample contain 4.0×10^{20} atoms of phosphorus-30, ${}_{15}P^{30}$. If the half-life of phosphorus-30 is 2.5 minutes, how many of these atoms will be left after 15 minutes?

Solution

$$N = N_o \exp(-0.693\ t/t_{1/2}) = N_o \exp(-0.693 \times 15 \text{ minutes}/ 2.5 \text{ minutes}) = N_o \exp(-0.693 \times 6)$$

$$N = N_o \exp(-4.16) = N_o \times 0.0156 = 4.0 \times 10^{20} \times 0.0156 = 6.25 \times 10^{18}.$$

Another way to do this is to recognize that 15 minutes is exactly 6 half-lives of phosphorus-30 Each half-life reduces the sample by one half so 6 half-lives reduces it by $(1/2)^6$ or 1/64.

Example 16.2

Tritium is an isotope of hydrogen whose nucleus includes one proton and two neutrons, ${}_1H^3$. A sample is prepared containing 4.2×10^{16} atoms. How long will it take for the sample to be reduced to 1.1×10^{15} atoms? The half-life of tritium is 12.33 years.

Solution

$$N = N_o \exp(-0.693\ t/t_{1/2}) \text{ or}$$

$$N/N_o = 1.1 \times 10^{15} / 4.\ 2 \times 10^{16} = 0.262 \times 10^{-1} = 0.0262 = \exp(-0.693\ t/12.33 \text{ yrs})$$

$$\ln [0.0262] = \ln [\exp(-0.693\ t/12.33 \text{ yrs})]$$

where ln is the natural logarithm. Recall that $\ln[\exp (x)] = x$

$$\ln[-0.0262] = -3.64 = (-0.693\ t/12.33 \text{ yrs})$$

$$0.693\ t / 12.33 \text{ yrs} = 3.64 \text{ or } 0.693\ t = 12.33 \text{ yrs} \times 3.64 = 44.9$$

$$t = 44.9 / 0.693 = 64.8 \text{ years.}$$

One way to check this answer is to notice that 64.8years is about 5.2 half-lives. Five half-lives would have reduced the sample to $(1/2)^5$ of the original tritium atoms. $(1/2)^5 = 0.0313$ and

$$0.0313 \times 4.2 \times 10^{16} = 1.31 \times 10^{15} \text{ atoms.}$$

Because 64.8 years is a little more than five half-lives, we would expect the sample to contain slightly less than 1.31×10^{15} atoms, which is true.

The decay of radioactive nuclei is a purely statistical process. What this means is the behavior of very large numbers of atoms satisfies certain relationships such as the radioactive-decay law, but that the actual behavior of any individual nucleus cannot be predicted. For example, if we start with a sample containing a very large number of sodium-24 atoms, we can predict with great confidence that almost exactly one-half of those atoms will remain after an interval of 15 hours (one half-life). However, if we begin with only two sodium-24 atoms, we cannot be sure that exactly one atom will remain after 15 hours; we may well have both atoms left or none left. We can compare this situation to the statistical behavior of tossed coins. If you toss a very large number of fair coins, you can be certain that almost exactly 50 percent of the results will be heads. However, you cannot be certain at all about the results that will be obtained for two tosses. It would not be unusual to obtain a pair of heads or a pair of tails.

It is important to understand that the age of the atom has no influence on the probability that it will decay within a given time interval. If a nucleus of sodium-24 is produced in an induced nuclear reaction, there is a fifty-fifty chance that this nucleus will decay in the first 15 hours after its formation. The probability that the nucleus will survive for three days is very small (less than 1 in 20) but, if it does not decay in the first two days and nine hours, then the chance that it will decay in the following 15 hours is still fifty-fifty. Past history cannot influence this probability in any way. Again, the coin-tossing process provides a useful analogy. Although the chance of obtaining heads of five successive tosses is small (again less than 1 in 20), if you obtain four heads in a row, then the chances of getting heads on the fifth toss are still fifty-fifty. To think that the probability of getting a fifth head after getting four in a row is less than fifty-fifty is known as the gambler's fallacy.

The radioactive-decay law is valid for any sample of a given radioactive isotope, no matter what the environmental conditions. The rate of decay is not affected at all by such factors as temperature, pressure, or chemical reactions. Only nuclear reactions produced by bombardment experiments can have any effect upon the natural process of radioactive decay.

We should also note one difference between the coin-tossing analogy and the process of radioactive decay. At least in theory, it should be possible to predict the result of an individual coin toss if we knew all of the forces acting on the coin. This macroscopic process obeys statistical laws only because various factors (such as the direction and force of the toss and air currents) vary randomly from one toss to the next. We describe the process by statistical laws because we do not have detailed knowledge of the particular conditions on a given toss. In contrast, the theory of quantum mechanics indicates the process of radioactive decay is statistical by its very nature. According to the Copenhagen interpretation, it is impossible even in theory to predict when an individual nucleus will decay. There is no property of the nucleus that could be used to make such a prediction. Even if we knew everything there is to know about every individual nucleus in a radioactive sample, we still would know nothing more than the probability that a given nucleus will decay within a given time interval.

16.2 AGE-DATING WITH RADIOACTIVE ISOTOPES

Naturally occurring radioactive isotopes exist on earth for two reasons. Some isotopes with relatively short half-lives exist in nature because they are continually being formed by natural processes. For example, hydrogen-3 (tritium) and carbon-l4 are continuously being formed in the upper atmosphere through nuclear reactions that occur as cosmic rays (high-energy particles arriving from space) bombard the nuclei in the air. Other radioactive isotopes (such as uranium-238 and potassium-40) exist in nature because they have half-lives that are comparable to the age of the solar system. These isotopes were formed in nuclear reactions that occurred under very different conditions earlier in the history of the universe. Due to their long half-lives, a significant fraction of the original population still exists.

The naturally occurring radioactive isotopes are used in a variety of ways to date astronomical, geological, and archeological events. This procedure is commonly called age-dating. Uranium-238 is one example of a radioactive isotope used for dating geological events. The half-life of uranium-238 is 4.47×10^9 years. Alter a series of radioactive decays, the final product is the stable isotope, lead-206 ($_{82}Pb^{206}$). The symbol Pb represents the old Latin name (plumbum) for lead. The half-lives for all of the other decays in this series of reactions are far shorter than the half-life for the initial decay of the uranium-238, so we can approximate the situation quite closely by saying that uranium-238 decays to form lead-206 with a half-life of 4.47 x 10^9 years. Now consider a rock sample that contains some uranium-238. It will also contain some lead-206 that has been formed by the decay of the radioactive uranium. Let N be the number of atoms of uranium-238 in the sample and let N_1 be the number of atoms of lead-206. Assume there was no lead-206 in the rock when it formed, so that all of the lead-206 now present has been formed by decay of uranium-238. Also, assume no uranium or lead has entered or left the sample over the time since the rock first formed. Then we know that the number of atoms of uranium-238 present when the rock formed was $N_o = N_1 + N$. That is, the number of uranium-238 atoms still remaining plus the number that have decayed into lead-206. We now know the values of No, N, and the half-life, so we can use the radioactive-decay law to solve for the time that has passed since the rock was formed.

In the actual age-dating technique used by geologists, more complicated calculations are used to estimate the amount of lead-206 that was present initially and to test the assumption that none of these isotopes has entered or left the sample during its lifetime. Other age-dating techniques based on other isotopes with long half-lives can be used for an independent check on the age computed from the uranium-238 decay. The age of a rock calculated from this technique represents the time that has passed since the rock solidified from a molten state. Thus, most of the rocks found near the earth's surface today are considerably younger than the earth itself. Rocks did not begin to solidify until well after the initial formation of the earth, and most rock materials have subsequently been re-melted and re-solidified at least once by processes occurring in the earth's crust. The oldest rocks found on the earth thus far are about 4 x 10^9 years old. Some rocks brought back from the moon have been age-dated at 4.4 x 10^9 years, and meteorites have been dated at about 4.6 x 10^9 years. These results, together with those obtained from other age-dating methods, indicate that our solar system (including the earth) probably formed around 4.6 x 10^9 years ago. This makes our solar system relatively young in the universe, which data indicate to have existed for 13.7 billion years.

Age-dating techniques of the kind just discussed can be applied only to mineral materials that contain long-lived radioactive isotopes. Because uranium-238 decays so slowly, the number of atoms decaying over a period of less than several million years is too small to permit accurate age determinations. Therefore, other techniques must be used to age-date samples of archeological interest. One such technique uses the isotope carbon-14 ($_6C^{14}$), which has a half-life of 5730 years.

Obviously, all of the carbon-14 that was present when the earth formed has long since decayed. However, carbon-14 does exist naturally on the earth because it is continuously formed as high-energy cosmic rays (mostly protons) bombard nuclei in the upper atmosphere. As radioactive carbon is formed in the upper atmosphere, it combines with oxygen to form the gas carbon dioxide. During photosynthesis, this gas is taken in by plant tissues and eventually finds its way into all living plants and animals through the food chains. The percentage of all carbon atoms that are carbon-14 is approximately the same in any living organism; about 1 in 10^{12} carbon atoms is a carbon-14 atom. If we assume that the rate of cosmic-ray bombardment has remained roughly constant over the period for which this age-dating technique is useful (about 60,000 years), then we can assume that every living organism during this interval has had about one carbon-14 atom for every 10^{12} carbon atoms in its body. This proportion of carbon-14 remains constant only as long as the organism is alive. The living organism constantly takes in carbon atoms through photosynthesis or feeding, so that new carbon-14 atoms are taken in to replace those that decay. When the organism dies, this exchange of carbon atoms ceases, and the proportion of carbon-14 begins to decrease as the carbon-14 atoms decay. At the time of death (call it $t = 0$), there is one carbon-14 atom for every 10^{12} carbon atoms. After one half-life (5730 years) there will be one-half of the carbon-14 atoms remaining, or one carbon-14 atom for every 2×10^{12} carbon atoms. After two half-lives (11,460 years), only one-quarter of the original carbon-14 atoms will remain. The decay process continues until all of the carbon-14 atoms have decayed. After around 60,000 years, the amount of carbon-14 in the sample is so small that it cannot be accurately measured.

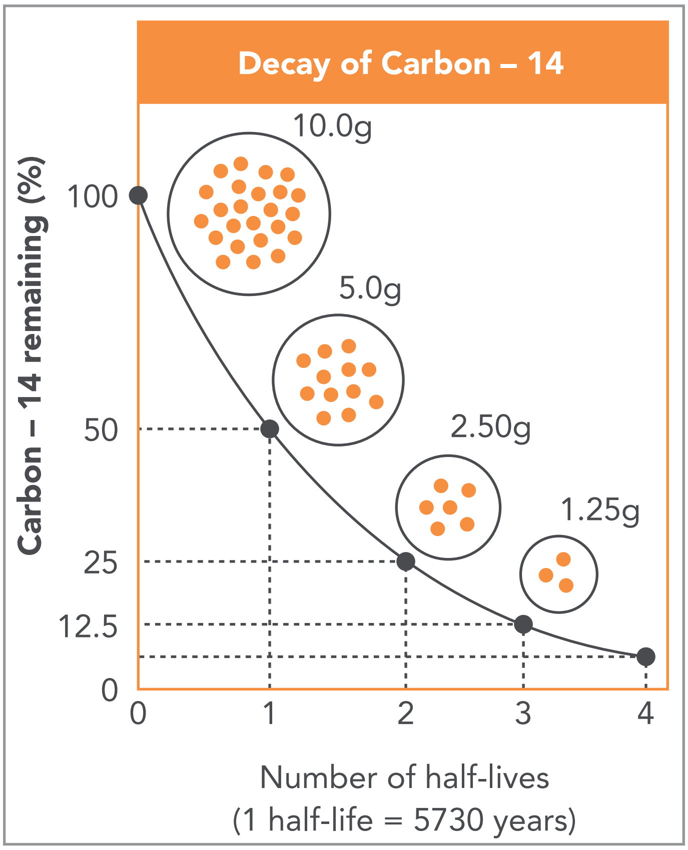


Figure 16.3 Decay of carbon-12

When an archeologist discovers some fragment of bone or wood, a physicist can determine the age of that sample by measuring the fraction of its carbon that.is carbon-l4. Again, tests are used to verify that the sample has not gained or lost carbon atoms during the period since the organism died. This age-dating technique has been tested by applying it to artifacts whose age is known through historical records, and the agreement is reasonably good. Other age-dating techniques have been developed to provide independent checks on the results of the carbon-14 method. Thus archeologists are able to speak of the age of various human remains with considerable confidence in the validity of these numbers.

16.3 GAMMA DECAY

Naturally occurring radioactive isotopes undergo three distinct types of radioactive decay: alpha decay, beta decay, and gamma decay. Each type corresponds to the emission of a particular type of radiation by the decaying nucleus. Alpha decay involves the ejection of an alpha particle (a helium-4 nucleus) from the nucleus. Beta decay involves the ejection of an electron by the nucleus. Gamma decay involves emission of a photon from the nucleus. The energies of the gamma-ray photons are the highest energy photons found in nature. Thus, gamma rays are found at the short-wavelength, high-frequency end of the electromagnetic spectrum.

In Chapter 12 we saw that the existence of discrete atomic emission and absorption spectra can be explained only by assuming quantized energy states for the atom. Analysis of gamma rays emitted or absorbed by a particular atomic nucleus also yields discrete spectra. Therefore, we must conclude that the nucleus exists only in quantized energy states. As in the case of atomic spectra, the wavelengths and intensities of the radiation are characteristic of the particular isotope involved. The principle difference between atomic and nuclear spectra is that the photons emitted or absorbed in nuclear reactions are of much greater energies than those involved in the atomic spectra. An excited atomic state is typically several electron-volts above the ground-state energy, so the photons emitted or absorbed in transitions between these states have corresponding energies. In contrast, the separations between energy states of the nucleus are on the order of hundreds of thousands or millions of electron-volts.

For the atom, we were able to visualize the higher energy states in terms of larger electron orbits. That is, the greater the average distance between the electron and the nucleus, the higher the energy of the atom. However, no such model is suitable for the excited states of the nucleus. If you wish to have a visualizible model for excited nuclear states, you might imagine that the protons and neutrons move within the nucleus at greater speeds for higher nuclear-energy states, just as the molecules of a contained gas move at higher speeds if heat (energy) is added to the system.

The nucleus of a given isotope can exist only in certain definite energy states. A transition from a higher (excited) state to a lower state is usually accomplished by the emission of a photon, which traditionally has been called a gamma (γ) ray. This is the process called gamma decay. Most excited nuclei undergo gamma decay with very short half-lives (on the order of 10^{-14} s). A few excited nuclei do have much longer half-lives, on the order of hours in some cases. These long-lived excited nuclei are called isomers, and they usually are designated by an asterisk. Thus, the excited strontium-87 nucleus ${}_{38}Sr^{87*}$ with a half-life of 2.3 hours is an isomer of ${}_{38}Sr^{87}$.

The general nuclear equation for gamma decay is

$$ {}_{Z}X^{A*} \rightarrow {}_{Z}X^{A} + {}_{0}\gamma^{0} $$

where $_{Z}X^{A}$ designates an arbitrary nucleus. The nuclear notation, $_{0}\gamma^{0}$, for the gamma-ray photon simply indicates that the gamma ray is uncharged (Z = 0) and contains no nucleons (A = 0). This extension of the nuclear notation to elementary particles simply makes it easy to check the equation for conservation of charge and conservation of nucleon number. In a radioactive-decay reaction, the nucleus that decays is called the parent nucleus, and the resulting nucleus is called the daughter nucleus.

Consider the gamma decay of an excited nucleus whose mass is M*. From conservation of energy, it is clear that the mass M of the daughter nucleus must be smaller than M*. If we assume that the excited nucleus was at rest before the reaction, then there was only rest-mass energy before the reaction, whereas there is kinetic and radiant energy as well as rest-mass energy after the reaction. The rest-mass energy must decrease during the reaction by an amount equal to the increase in kinetic plus radiant energy. Applying conservation of energy quantitatively, we can write

$$M^*c^2 = Mc^2 + E_k + h\nu$$

where hν is the energy of the emitted photon. If a massive cannon fires a small bullet, the cannon recoils with negligibly small kinetic energy. The situation of a nucleus emitting a gamma ray is similar (see Question 15 at the end of this chapter), so we can emit the kinetic-energy term in the energy-conservation equation without significantly affecting the result. Thus we have

$$M^*c^2 = Mc^2 + h\nu \text{ or}$$

$$h\nu = (M^* - M)c^2 = \text{the Q of the reaction.}$$

Like every process of radioactive decay, gamma decay is always an exoergic reaction. That is, Q is always positive. Notice the similarity between this equation for the energy of the photon emitted in a transition between nuclear energy states and the Bohr equation for the energy of the photon emitted in a transition between atomic energy states,

$$h\nu = E_u - E_l$$

As examples of gamma decay, here are the nuclear equations for the gamma decay of isomers of copper-68 (half-life = 3.8 minutes) and zinc-69: (half-life = 13.9 minutes)

$${}_{29}Cu^{68*} \rightarrow {}_{29}Cu^{68} + {}_{0}\gamma^{0}$$

$${}_{30}Zn^{69*} \rightarrow {}_{30}Zn^{69} + {}_{0}\gamma^{0}$$

Example 16.5

An excited oxygen-16 nucleus decays to the ground state by emitting a photon of energy 1.300 MeV. Determine the mass of the excited nucleus. The mass of the ground state oxygen is 15.990523 u.

Solution

$${}_{8}O^{16*} \rightarrow {}_{8}O^{16} + {}_{0}\gamma^{0}$$

From energy conservation, $M^*c^2 = Mc^2 + h\nu$.

$$M^* = M + h\nu/ c^2 = 15{,}990523 \text{ u} + 1.300 \text{ MeV}/931 \text{ MeV/u}$$

$$M^* = 15{,}990523 \text{ u} + 0.001396 = 15.991928 \text{ u.}$$

16.4 ALPHA DECAY

The nucleus of an atom can exist only in certain discrete energy states. If the nucleus is not in its ground state, it is necessarily unstable, and it will usually decay by gamma emission. However, certain combinations of nucleons are unstable even in their ground states and will usually decay by alpha or beta decay.

Alpha decay is the spontaneous transition of an unstable nucleus (typically a very heavy nucleus) by the emission of an alpha particle (${}_2He^4$ nucleus). The general nuclear equation for alpha decay is

$$ {}_ZX^A \rightarrow {}_{Z-2}Y^{A-4} + {}_2He^4 $$

Note that alpha decay is a transmutation reaction. The daughter nucleus is of a different element (a different atomic number Z) than the parent nucleus. In contrast, gamma decay does not involve a change in the element. The daughter nucleus in alpha decay has two fewer protons and two fewer neutrons than the parent nucleus. (We have used the symbol Y in the general equation simply to indicate that the daughter element is of a different element than the parent nucleus.

Consider the alpha decay of a parent nucleus of rest mass M_p. If the parent nucleus is at rest before the reaction, then the total energy of the system is the rest-mass energy of the parent nucleus, M_pc^2. After the reaction, there is kinetic energy and rest-mass energy of the daughter nucleus and the alpha particle. The rest-mass energy of the system must decrease by an amount equal to the increase in kinetic energy. Therefore, the sum of the rest masses of the daughter nucleus (M_d) and the alpha particle (M_α) must be smaller than M_p.

Applying conservation of energy quantitatively, we can write

$$ M_pc^2 = (M_d + M_\alpha)\, c^2 + E_{kd} + E_{k\alpha} $$

Because of the relatively large mass of the alpha particle, we cannot neglect the recoil kinetic energy of the daughter as we did for gamma decay. Solving the energy-conservation equation for the kinetic energies, we obtain

$$ E_{kd} + E_{k\alpha} = (M_p - M_d - M_\alpha)\, c^2 = \text{the Q of the reaction.} $$

That is, the increase in kinetic energy is equal to the decrease in rest-mass energy. In any case where M_p is larger than the sum $M_d + M_\alpha$, alpha decay can occur as an exoergic reaction, and the parent nucleus is necessarily unstable. In any case where M_p is smaller than $M_d + M_\alpha$, alpha decay cannot occur spontaneously. As examples of alpha decay, consider the following nuclear equations for the alpha decay of uranium-238 (half-life = 4.47×10^9 years) and polonium-210 (half-life = 138.4 days)

$${}_{92}U^{238} \rightarrow {}_{90}Th^{234} + {}_{2}He^{4}$$

$${}_{84}Po^{210} \rightarrow {}_{82}Pb^{206} + {}_{2}He^{4}$$

The daughter nuclei in these reactions are thorium-234 and lead-206.

16.5 BETA DECAY

There are good reasons for believing that electrons cannot be a constituent of the nucleus. However, in some cases, an unstable nucleus undergoes a spontaneous transition in which a neutron is converted to a proton with the emission of an electron (beta particle). This process is called beta decay.

From conservation of energy,

$$M_p c^2 = (M_d + M_\beta)c^2 + E_{kd} + E_{k\beta}$$

where the subscript β refers to the beta particle, or electron, so that $M_\beta = m_{oe}$ (the rest mass of an electron). The mass of the daughter nucleus is very much greater than the mass of the electron, so in this case we can ignore the negligible recoil kinetic energy of the daughter nucleus. Thus we can write

$$M_p c^2 = (M_d + M_\beta)c^2 + E_{k\beta} \text{ or}$$

$$E_{k\beta} = (M_p - M_d - M_\beta)\, c^2 = \text{the Q of the reaction.}$$

As an example, consider the beta decay of carbon-14 to form nitrogen-14. The mass of the carbon-14 nucleus is 13.999948 u, the mass of the nitrogen-l4 nucleus is 13.999231 u, and the mass of the beta particle (electron) is 0.000549 u. Therefore, the expected kinetic energy of the beta particle is

$$E_{k\beta} = [(13.499948 \text{ u}) - (13.99923l \text{ u}) - (0.000549 \text{ u})] \text{ x } (931 \text{ MeV/u})$$

$$E_{k\beta} = (0.000162 \text{ u}) \text{ x } (931 \text{ MeV/u}) = 0.156 \text{ MeV}$$

Thus, we predict from conservation of energy and momentum that the beta particle in this reaction should be emitted with a kinetic energy of 0.156 MeV. (We used conservation of momentum to justify neglecting the kinetic energy of the daughter nucleus, and we used conservation of energy to obtain the equation for $E_{k\beta}$). Experimental results, however, do not confirm this prediction. Instead, it is found that the beta particles emitted by carbon-14 have a continuous range of kinetic energies. Only a very few have kinetic energies of 0.156 MeV; all others have less kinetic energy.

Suppose that a given carbon-14 nucleus emits a beta particle with 0.052 MeV of kinetic energy. This means that 0.104 MeV of energy has disappeared. Energy seems not to be conserved in this case. This problem of disappearing energy is not peculiar to the beta decay of carbon-14; the same problem arises in all beta decays.

As if this were not enough, there seem to be other problems with the beta-decay reaction. It is clear that the daughter nucleus must recoil in the direction opposite to the direction in which the beta particle is emitted if momentum is to be conserved. Again, experiments indicate that this is seldom the case. Therefore, although beta decay appears to be consistent with conservation of charge, it appears not to be consistent with conservation of either energy or momentum. Furthermore, other conservation laws that we have not discussed appear to be violated by the beta-decay process.

Should we abandon these conservation laws that have served us so faithfully in such a large range of physical phenomena? It may have occurred to you already that the explanation for the missing energy and the direction of recoil might be the involvement of a third particle in the beta-decay reaction. However, if such a third particle exists, it must be a very unusual particle indeed. Repeated and careful attempts to detect this particle were unsuccessful. Therefore, if such a particle exists, it must interact with matter far more weakly than any other known particle. Yet, if we are to save our conservation laws, we must postulate the existence of such a particle.

The existence of this undetected particle on was first postulated in 1930 by Wolfgang Pauli (1900 – 1958). Although the postulate was a radical one, it gained widespread support in 1934 when Enrico Fermi (1901 – 1954) successfully incorporated Pauli's particle into an explanation of beta decay. Fermi's model showed that the postulated properties of the particle were exactly those needed to make beta decay consistent with all the conservation laws. Physicists were greatly relieved to find a way to avoid having to modify the conservation laws that formed such a basic part of physical theories. In fact, they felt so strongly about the conservation laws that the existence of this undetected particle was never seriously doubted after 1934, even though no direct experimental evidence of its existence was obtained until 1956.

In informal conversation, Pauli called this particle the neutron, but he never used that name in his publications. When Chadwick discovered the uncharged nucleon in 1932, he used the name neutron for that particle in the paper announcing his discovery. Fermi, in explaining the difference between these two particles that had been given the same name, referred to Pauli's particle as the "little neutron" (in Italian, the "neutrino"). Because Chadwick had officially used the name neutron first for the nucleon, Pauli's particle came to be known officially as the neutrino. It is symbolized by the Greek letter nu (ν).

In some beta decays, the kinetic energy of the beta particle is (within the accuracy of measurement) equal to the decrease in rest-mass energy, so the neutrino can have little or no rest mass. The energy associated with the emitted neutrino must be almost completely in the form of kinetic energy. For theoretical reasons, it had long been believed that the rest-mass energy of the neutrino is exactly zero. However, some recent experimental evidence suggests that neutrinos have a mass of about mass of 0.06 eV, which is far less than a billionth of the mass of a proton. Another property of the neutrino is its extremely small probability of interacting with matter. If a beam of neutrinos were shot straight at the earth, almost all of them would pass straight through the earth without any interaction. Only one in 10^{12} neutrinos would interact in any way with the matter of the earth. As you can imagine, this makes it very difficult to devise any experiment that will detect the presence of neutrinos!

We now know that antineutrinos exist as particles distinct from neutrinos. Most properties of neutrinos are also properties of antineutrinos. Neutrinos and antineutrinos differ from each other only in a rather esoteric property called helicity, which we shall not discuss in this book. In fact, it turns out that the particles emitted .in beta decay are antineutrinos. The symbol for the antineutrino is nu overbar, $\bar{\upsilon}$.

Beta decay is now known to be the spontaneous transition of an unstable nucleus that results in the conversion of a neutron to a proton with the emission of a beta particle (electron) and an antineutrino. The general nuclear equation for beta decay is

$$ {}_{Z}X^{A} \rightarrow {}_{Z+1}Y^{A} + {}_{-1}\beta^{0} + {}_{0}\bar{\upsilon}^{0} $$

A neutron in the parent nucleus is converted to a proton, thus changing the atomic number (the element) but not the mass number. The beta particle is assigned a value of Z = -1 because of its negative unit charge. Now using the complete mechanism for beta decay, again consider the energetics of the decay. From conservation of energy. The rest mass of the antineutrino is so small that it can be neglected.

$$ M_p c^2 = (M_d + M_\beta)c^2 + E_{k\beta} + E_k \nu $$

$$E_{k\beta} + E_{kv} = (M_p - M_d - M_\beta)c^2 = Q \text{ of the reaction.}$$

This equation makes it clear that the energy that appeared to be missing in the beta-decay experiments was the kinetic energy of the undetected antineutrino. A simple example of a beta-decay reaction is

$${}_0n^1 \rightarrow {}_1H^1 + {}_{-1}\beta^0 + {}_0\bar{\upsilon}^0 \text{ with a half-life of 10.2 minutes}$$

The neutron, a fundamental constituent of the nucleus, is not a stable particle in its free state. After a relatively short time, it will decay to a proton, an electron, and an antineutrino. When a neutron combines with other neutrons and protons to form a nucleus, the configuration may or may not be stable. For example, the helium-4 nucleus with two neutrons and two protons is a stable configuration; the neutrons in this nucleus will not undergo beta decay. However, the carbon-14 nucleus with six protons and eight neutrons is an unstable configuration. Eventually, one of the neutrons will be converted to a proton, changing the carbon-14 to nitrogen-14, with the emission of an electron and an antineutrino Carbon-14 has a half-life of 5730 years.

$${}_6C^{14} \rightarrow {}_7N^{14} + {}_{-1}\beta^0 + {}_0\bar{\upsilon}^0$$

This is the reaction involved in carbon-14 age-dating. Another example of a beta-emitting nucleus is chlorine-38, which decays to argon-38 with a half-life of 37.2 minutes.

$${}_{17}Cl^{38} \rightarrow {}_{18}Ar^{38} + {}_{-1}\beta^0 + {}_0\bar{\upsilon}^0$$

The type of beta decay we have just discussed applies to all beta-decay reactions occurring in samples of radioactive nuclei found in nature. However, in the 1930s, Irene Joliot-Curie and her husband Frederic discovered that some radioactive nuclei produced by induced nuclear reactions undergo other types of beta decay. One of these types results in the conversion of a proton within the unstable nucleus to a neutron, with the emission of an antielectron (positron) and a neutrino. This type of decay is called beta-positive decay, and the type we have already discussed is more properly called beta-negative decay.

Summary

Radioactivity is the process in which an unstable atomic nucleus spontaneously disintegrates (decays) with the ejection of one or more particles from the nucleus. For any particular radioactive isotope, the fraction of radioactive nuclei that decay in a given time interval is constant. The half-life of an isotope is the time interval in which one-half of the radioactive nuclei of that isotope present in a sample will decay. During the next half-life interval, one-half of the remaining radioactive nuclei will decay, and so on until the number of radioactive nuclei is reduced to zero. Thus a radioactive sample is like a clock, with the percentage of atoms that have already decayed being a measure of the time that has passed since the sample was formed. Radioactive isotopes are used in a variety of ways to date astronomical, geological, and archeological events. Naturally occurring radioactive isotopes undergo three distinct types of decay: alpha decay, beta decay, and gamma decay. Gamma decay is the simplest process, involving the transition of a nucleus from an excited energy state to a lower state by the emission of a high-energy photon (a gamma ray). Alpha decay involves the emission of a helium-4 nucleus (an alpha particle) from the nucleus, reducing the atomic number by two and thus changing the chemical element. Beta decay involves the change of a neutron in the nucleus into a proton, with the emission of an electron (a beta particle) and an antineutrino; again, this results in a change in the atomic number and hence a chemical transmutation to a different element. The process of beta decay does not represent a separation of the neutron into three component parts; it is a discontinuous process in space-time -- at one instant there is a neutron, and at the next instant there is a proton, an electron, and an antineutrino that did not previously exist within the neutron.

Important concepts

Fluorescence; radioactivity; radioactive substance; spontaneous nuclear reaction; half-life; radioactive-decay law; age-dating; gamma decay; parent nucleus; daughter nucleus; alpha decay; beta decay; neutrino; antineutrino.

Questions

1. Explain what is meant by the half-life of a radioactive isotope.
2. How can we account for the existence of naturally occurring radioactive isotopes on the earth?
3. Use the nuclear masses in Appendix B to calculate the Q for the alpha decay of uranium-238.
4. A sample contains N_o atoms of a radioactive isotope with a half-life of one day. How many atoms of this isotope will remain in the sample after an interval of 6 days?
5. A 32-gram sample of a radioactive isotope is purified from its surroundings. After 30 days, only 1 gram of this isotope remains in the sample. What is the half-life of this isotope? What happened to the other 31 grams of the original sample?

6. The half-life of strontium-90 is 28 years. A soil sample contains 4 x 10^{-5} grams of this radioactive isotope. Estimate the mass of this isotope that will remain in the sample after 80 years have passed.
7. In any sample of natural uranium, 99.3 percent of the atoms are uranium-238 and 0.7 percent are uranium-235, How has each of these percentages changed over the past billion years (increased, decreased, or remained constant)? Explain your answer.
8. Outline the geological age-dating method that makes use of uranium-238.
9. In discussions of radioactive age-dating, uranium-238 and lead-206 are sometimes compared to the sand in the top and bottom of an hourglass. Explain this analogy.
10. Explain the radioactive age-dating technique that makes use of carbon-14. For what kinds of objects (and what range of ages) is this method particularly suitable?
11. Explain how carbon-14 is produced on the earth. What effects might cause the fraction of carbon-14 in the earth's atmosphere to change over time?
12. Explain why the carbon-14 radioactivity in any living organism is found to be virtually a constant

13. An archeologist extracts 10^{-6} grams of carbon from an ancient ax handle and finds that it is one-fourth as radioactive as 10^{-6} grams of carbon extracted from a freshly cut tree branch. How old is the ax handle?
14. What happens to the parent nucleus during the process of gamma decay?
15. The ground state of a zinc-67 nucleus has a mass of 66.910511 u. It has an isomeric state with a half-life of 9.3×10^{-6} seconds. The isomer decays with the emission of a photon (gamma ray) having an energy of 0.0933 7 MeV. What is the mass of the excited (isomeric) zinc-67 nucleus? What is the momentum of the gamma ray? Use conservation of momentum to calculate the recoil kinetic energy of the zinc-67 nucleus. Are we justified in neglecting E_{kd} in energy calculations for this gamma-decay reaction? Explain your answer.
16. Explain the difference between the terms isotope and isomer.
17. Is it possible for oxygen-16 to decay by alpha emission? Justify your answer.
18. Briefly explain the reasons for the postulation of the existence of the neutrino. What are the properties of this particle? Why did it take so long to actually detect the existence of the neutrino?
19. List the similarities and differences between the photon and the neutrino.
20. Describe the three types of radiation that are emitted by naturally occurring radioactive substances.
21. In general terms, discuss the nuclear changes that occur when each of the following particles is emitted from a nucleus: a) an alpha particle; b) an electron. c) a gamma ray.
22. Bismuth-214 may decay either by alpha decay or by beta-negative decay What is the daughter nucleus in each case?
23. Write the decay equations for hydrogen-3 and for uranium-235.
24. Write the general equation for beta-positive decay.
25. Would you expect fission fragments to decay by beta-positive or beta-negative decay? (Hint: look back at Figure 14.3.)
26. The natural decay chain for ${}_{92}U^{238}$ results in ${}_{82}Pb^{206}$ (lead). The chain is a sequence of alpha decays and beta-negative decays. How many of each type occurs in the chain?

The following questions are of a more general nature. They have no single correct answer and are just something for you to think about. When possible, they are best answered in conversation with others.

27. Many of the early investigators of radioactivity noted the large amounts of energy given off during radioactive decay and wondered about the source of this energy. Einstein's special theory of relativity was welcomed as a possible solution to this difficulty. Explain why.
28. There are very strong reasons to believe that electrons cannot exist in an atomic nucleus. However, during the process of beta decay, an electron is emitted from a nucleus. Explain how this can happen.

APPENDIX A: SCIENTIFIC NOTATION

Scientific notation is the way that scientists easily handle very large numbers or very small numbers. For example, instead of writing 0.0000000056, we write 5.6×10^{-9}. So, how does this work?

We can think of 5.6×10^{-9} as the product of two numbers: 5.6 (the digit term) and 10^{-9} (the exponential term).

Here are some examples of scientific notation.

$10000 = 1 \times 10^4$
$1000 = 1 \times 10^3$
$100 = 1 \times 10^2$
$10 = 1 \times 10^1$
$1 = 10^0$
$1/10 = 0.1 = 1 \times 10^{-1}$
$1/100 = 0.01 = 1 \times 10^{-2}$
$1/1000 = 0.001 = 1 \times 10^{-3}$
$1/10000 = 0.0001 = 1 \times 10^{-4}$

$24327 = 2.4327 \times 10^4$
$7354 = 7.354 \times 10^3$
$482 = 4.82 \times 10^2$
(two and one digit numbers are usually not written in scientific notation)
$0.053 = 5.3 \times 10^{-2}$
$0.0078 = 7.8 \times 10^{-3}$
$0.00044 = 4.4 \times 10^{-4}$

As you can see, the exponent of 10 is the number of places the decimal point must be shifted to give the number in long form. A positive exponent shows that the decimal point is shifted that number of places to the right. A negative exponent shows that the decimal point is shifted that number of places to the left.

In scientific notation, the digit term indicates the number of significant figures in the number. The exponential term only places the decimal point. As an example,

$46600000 = 4.66 \times 10^7$

This number only has 3 significant figures. The zeros are not significant; they are only holding a place. As another example,

$0.00053 = 5.3 \times 10^{-4}$

This number has 2 significant figures. The zeros are only place holders.

How to do calculations:

Addition and Subtraction:

- All numbers are converted to the same power of 10, and the digit terms are added or subtracted.
- Example: $(4.215 \times 10^{-2}) + (3.2 \times 10^{-4}) = (4.215 \times 10^{-2}) + (0.032 \times 10^{-2}) = 4.247 \times 10^{-2}$
- Example: $(8.97 \times 10^{4}) - (2.62 \times 10^{3}) = (8.97 \times 10^{4}) - (0.262 \times 10^{4}) = 8.71 \times 10^{4}$

Multiplication:

- The digit terms are multiplied in the normal way and the exponents are added. The end result is changed so that there is only one nonzero digit to the left of the decimal.
- Example: $(3.4 \times 10^{6})(4.2 \times 10^{3}) = (3.4)(4.2) \times 10^{(6+3)} = 14.28 \times 10^{9} = 1.4 \times 10^{10}$ (to 2 significant figures)
- Example: $(6.73 \times 10^{-5})(2.91 \times 10^{2}) = (6.73)(2.91) \times 10^{(-5+2)} = 19.58 \times 10^{-3} = 1.96 \times 10^{-2}$ (to 3 significant figures)

Division:

- The digit terms are divided in the normal way and the exponents are subtracted. The quotient is changed (if necessary) so that there is only one nonzero digit to the left of the decimal.
- Example: $(6.4 \times 10^{6})/(8.9 \times 10^{2}) = (6.4)/(8.9) \times 10^{(6-2)} = 0.719 \times 10^{4} = 7.2 \times 10^{3}$ (to 2 significant figures)
- Example: $(3.2 \times 10^{3})/(5.7 \times 10^{-2}) = (3.2)/(5.7) \times 10^{3-(-2)} = 0.561 \times 10^{5} = 5.6 \times 10^{4}$ (to 2 significant figures)

Powers of Exponentials:

- The digit term is raised to the indicated power and the exponent is multiplied by the number that indicates the power.
- Example: $(2.4 \times 10^{4})^{3} = (2.4)^{3} \times 10^{(4x3)} = 13.824 \times 10^{12} = 1.4 \times 10^{13}$ (to 2 significant figures)
- Example: $(6.53 \times 10^{-3})^{2} = (6.53)^{2} \times 10^{(-3)x2} = 42.64 \times 10^{-6} = 4.26 \times 10^{-5}$ (to 3 significant figures)