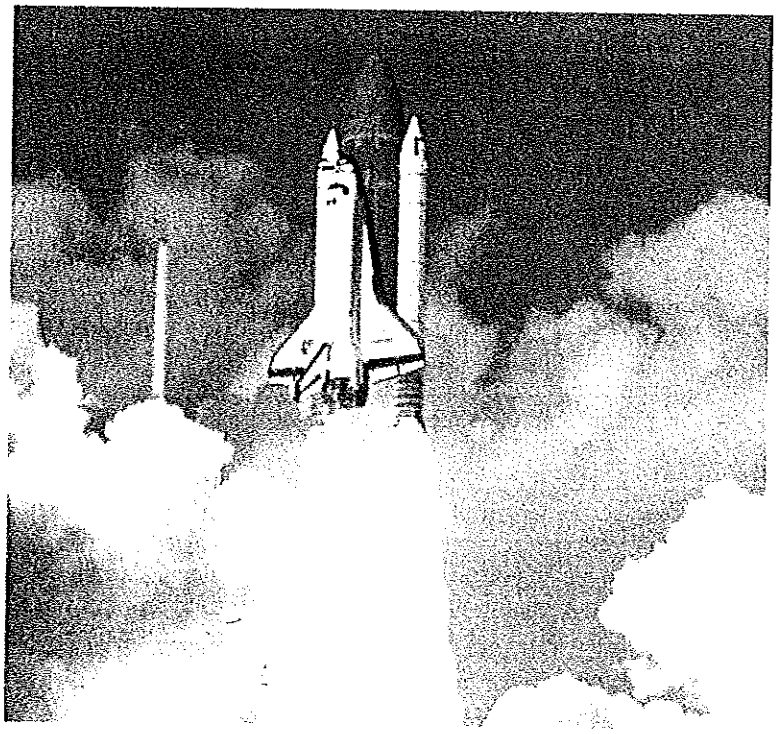
CHAPTER 1

Relativity



According to the theory of relativity, nothing can travel faster than light. Although today's spacecraft can exceed 10 km/s, they are far from this ultimate speed limit.

1.1 SPECIAL RELATIVITY
All motion is relative; the speed of light in free space is the same for all observers

1.2 TIME DILATION
A moving clock ticks more slowly than a clock at rest

1.3 DOPPLER EFFECT
Why the universe is believed to be expanding

1.4 LENGTH CONTRACTION
Faster means shorter

1.5 TWIN PARADOX
A longer life, but it will not seem longer

1.6 ELECTRICITY AND MAGNETISM
Relativity is the bridge

1.7 RELATIVISTIC MOMENTUM
Redefining an important quantity

1.8 MASS AND ENERGY
Where $E_0 = mc^2$ comes from

1.9 ENERGY AND MOMENTUM
How they fit together in relativity

1.10 GENERAL RELATIVITY
Gravity is a warping of spacetime

APPENDIX I: THE LORENTZ TRANSFORMATION

APPENDIX II: SPACETIME

In 1905 a young physicist of twenty-six named Albert Einstein showed how measurements of time and space are affected by motion between an observer and what is being observed. To say that Einstein's theory of relativity revolutionized science is no exaggeration. Relativity connects space and time, matter and energy, electricity and magnetism—links that are crucial to our understanding of the physical universe. From relativity have come a host of remarkable predictions, all of which have been confirmed by experiment. For all their profundity, many of the conclusions of relativity can be reached with only the simplest of mathematics.

1.1 SPECIAL RELATIVITY

All motion is relative; the speed of light in free space is the same for all observers

When such quantities as length, time interval, and mass are considered in elementary physics, no special point is made about how they are measured. Since a standard unit exists for each quantity, who makes a certain determination would not seem to matter—everybody ought to get the same result. For instance, there is no question of principle involved in finding the length of an airplane when we are on board. All we have to do is put one end of a tape measure at the airplane's nose and look at the number on the tape at the airplane's tail.

But what if the airplane is in flight and we are on the ground? It is not hard to determine the length of a distant object with a tape measure to establish a baseline, a surveyor's transit to measure angles, and a knowledge of trigonometry. When we measure the moving airplane from the ground, though, we find it to be shorter than it is to somebody in the airplane itself. To understand how this unexpected difference arises we must analyze the process of measurement when motion is involved.

Frames of Reference

The first step is to clarify what we mean by motion. When we say that something is moving, what we mean is that its position relative to something else is changing. A passenger moves relative to an airplane; the airplane moves relative to the earth; the earth moves relative to the sun; the sun moves relative to the galaxy of stars (the Milky Way) of which it is a member; and so on. In each case a **frame of reference** is part of the description of the motion. To say that something is moving always implies a specific frame of reference.

An **inertial frame of reference** is one in which Newton's first law of motion holds. In such a frame, an object at rest remains at rest and an object in motion continues to move at constant velocity (constant speed and direction) if no force acts on it. Any frame of reference that moves at constant velocity relative to an inertial frame is itself an inertial frame.

All inertial frames are equally valid. Suppose we see something changing its position with respect to us at constant velocity. Is it moving or are we moving? Suppose we are in a closed laboratory in which Newton's first law holds. Is the laboratory moving or is it at rest? These questions are meaningless because all constant-velocity motion is relative. There is no universal frame of reference that can be used everywhere, no such thing as "absolute motion."

The **theory of relativity** deals with the consequences of the lack of a universal frame of reference. **Special relativity**, which is what Einstein published in 1905, treats problems that involve inertial frames of reference. **General relativity**, published by Einstein a decade later, describes the relationship between gravity and the geometrical structure of space and time. The special theory has had an enormous impact on much of physics, and we shall concentrate on it here.

Postulates of Special Relativity

Two postulates underlie special relativity. The first, the **principle of relativity**, states:

> The laws of physics are the same in all inertial frames of reference.

This postulate follows from the absence of a universal frame of reference. If the laws of physics were different for different observers in relative motion, the observers could find from these differences which of them were "stationary" in space and which were "moving." But such a distinction does not exist, and the principle of relativity expresses this fact.

The second postulate is based on the results of many experiments:

> The speed of light in free space has the same value in all inertial frames of reference.

This speed is 2.998×10^8 m/s to four significant figures.

To appreciate how remarkable these postulates are, let us look at a hypothetical experiment basically no different from actual ones that have been carried out in a number of ways. Suppose I turn on a searchlight just as you fly past in a spacecraft at a speed of 2×10^8 m/s (Fig. 1.1). We both measure the speed of the light waves from the searchlight using identical instruments. From the ground I find their speed to be 3×10^8 m/s as usual. "Common sense" tells me that you ought to find a speed of $(3 - 2) \times 10^8$ m/s, or only 1×10^8 m/s, for the same light waves. But you also find their speed to be 3×10^8 m/s, even though to me you seem to be moving parallel to the waves at 2×10^8 m/s.

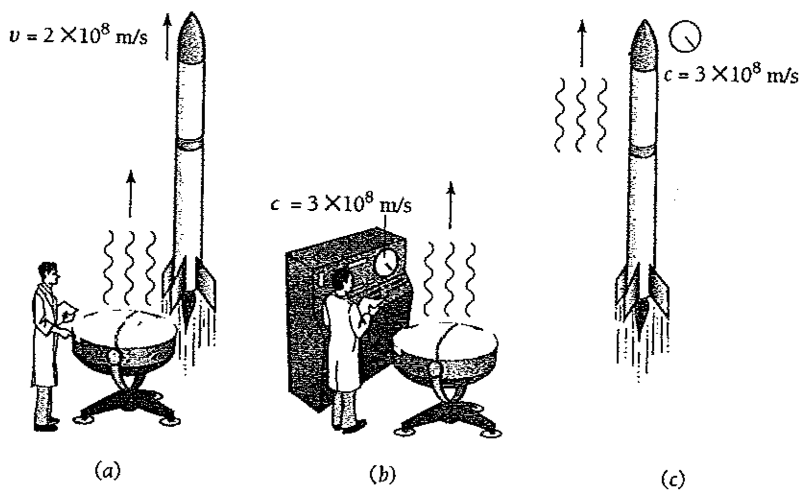


Figure 1.1 The speed of light is the same to all observers.

Albert A. Michelson (1852–1931) was born in Germany but came to the United States at the age of two with his parents, who settled in Nevada. He attended the U.S. Naval Academy at Annapolis where, after two years of sea duty, he became a science instructor. To improve his knowledge of optics, in which he wanted to specialize, Michelson went to Europe and studied in Berlin and Paris. Then he left the Navy to work first at the Case School of Applied Science in Ohio, then at Clark University in Massachusetts, and finally at the University of Chicago, where he headed the physics department from 1892 to 1929. Michelson's speciality was high-precision measurement, and for many decades his successive figures for the speed of light were the best available. He redefined the meter in terms of wavelengths of a particular spectral line and devised an interferometer that could determine the diameter of a star (stars appear as points of light in even the most powerful telescopes).

Michelson's most significant achievement, carried out in 1887 in collaboration with Edward Morley, was an experiment to measure the motion of the earth through the "ether," a hypothetical medium pervading the universe in which light waves were supposed to occur. The notion of the ether was a hangover from the days before light waves were recognized as electromagnetic, but nobody at the time seemed willing to discard the idea that light propagates relative to some sort of universal frame of reference.

To look for the earth's motion through the ether, Michelson and Morley used a pair of light beams formed by a half-silvered mirror, as in Fig. 1.2. One light beam is directed to a mirror along a path perpendicular to the ether current, and the other goes to a mirror along a path parallel to the ether current. Both beams end up at the same viewing screen. The clear glass plate ensures that both beams pass through the same thicknesses of air and glass. If the transit times of the two beams are the same, they will arrive at the screen in phase and will interfere constructively. An ether current due to the earth's motion parallel to one of the beams, however, would cause the beams to have different transit times and the result would be destructive interference at the screen. This is the essence of the experiment.

Although the experiment was sensitive enough to detect the expected ether drift, to everyone's surprise none was found. The negative result had two consequences. First, it showed that the ether does not exist and so there is no such thing as "absolute motion" relative to the ether: all motion is relative to a specified frame of reference, not to a universal one. Second, the result showed that the speed of light is the same for all observers, which is not true of waves that need a material medium in which to occur (such as sound and water waves).

The Michelson-Morley experiment set the stage for Einstein's 1905 special theory of relativity, a theory that Michelson himself was reluctant to accept. Indeed, not long before the flowering of relativity and quantum theory revolutionized physics, Michelson announced that "physical discoveries in the future are a matter of the sixth decimal place." This was a common opinion of the time. Michelson received a Nobel Prize in 1907, the first American to do so.

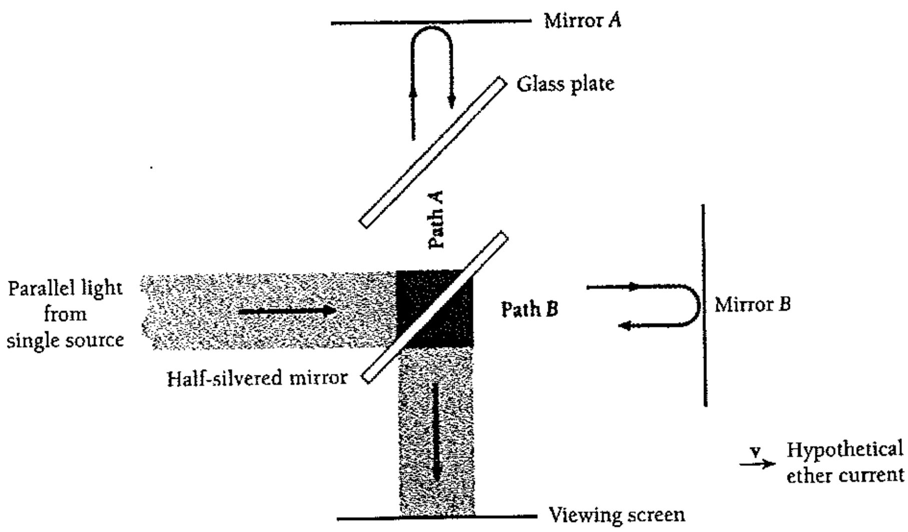


Figure 1.2 The Michelson-Morley experiment.

There is only one way to account for these results without violating the principle of relativity. It must be true that measurements of space and time are not absolute but depend on the relative motion between an observer and what is being observed. If I were to measure from the ground the rate at which your clock ticks and the length of your meter stick, I would find that the clock ticks more slowly than it did at rest on the ground and that the meter stick is shorter in the direction of motion of the spacecraft. To you, your clock and meter stick are the same as they were on the ground before you took off. To me they are different because of the relative motion, different in such a way that the speed of light you measure is the same 3×10^8 m/s I measure. Time intervals and lengths are relative quantities, but the speed of light in free space is the same to all observers.

Before Einstein's work, a conflict had existed between the principles of mechanics, which were then based on Newton's laws of motion, and those of electricity and magnetism, which had been developed into a unified theory by Maxwell. Newtonian mechanics had worked well for over two centuries. Maxwell's theory not only covered all that was then known about electric and magnetic phenomena but had also predicted that electromagnetic waves exist and identified light as an example of them. However, the equations of Newtonian mechanics and those of electromagnetism differ in the way they relate measurements made in one inertial frame with those made in a different inertial frame.

Einstein showed that Maxwell's theory is consistent with special relativity whereas Newtonian mechanics is not, and his modification of mechanics brought these branches of physics into accord. As we will find, relativistic and Newtonian mechanics agree for relative speeds much lower than the speed of light, which is why Newtonian mechanics seemed correct for so long. At higher speeds Newtonian mechanics fails and must be replaced by the relativistic version.

1.2 TIME DILATION

A moving clock ticks more slowly than a clock at rest

Measurements of time intervals are affected by relative motion between an observer and what is observed. As a result, a clock that moves with respect to an observer ticks more slowly than it does without such motion, and all processes (including those of life) occur more slowly to an observer when they take place in a different inertial frame.

If someone in a moving spacecraft finds that the time interval between two events in the spacecraft is t_0, we on the ground would find that the same interval has the longer duration t. The quantity t_0, which is determined by events that occur *at the same place* in an observer's frame of reference, is called the **proper time** of the interval between the events. When witnessed from the ground, the events that mark the beginning and end of the time interval occur at different places, and in consequence the duration of the interval appears longer than the proper time. This effect is called **time dilation** (to dilate is to become larger).

To see how time dilation comes about, let us consider two clocks, both of the particularly simple kind shown in Fig. 1.3. In each clock a pulse of light is reflected back and forth between two mirrors L_0 apart. Whenever the light strikes the lower mirror, an electric signal is produced that marks the recording tape. Each mark corresponds to the tick of an ordinary clock.

One clock is at rest in a laboratory on the ground and the other is in a spacecraft that moves at the speed v relative to the ground. An observer in the laboratory watches both clocks: does she find that they tick at the same rate?

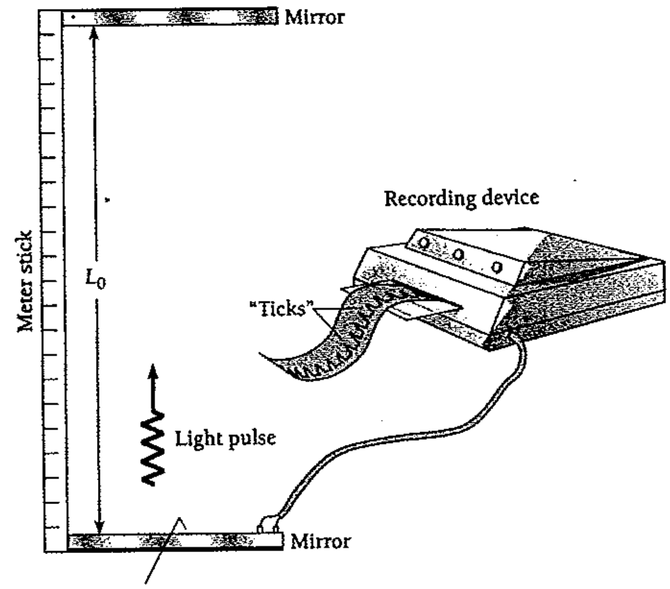


Figure 1.3 A simple clock. Each "tick" corresponds to a round trip of the light pulse from the lower mirror to the upper one and back.

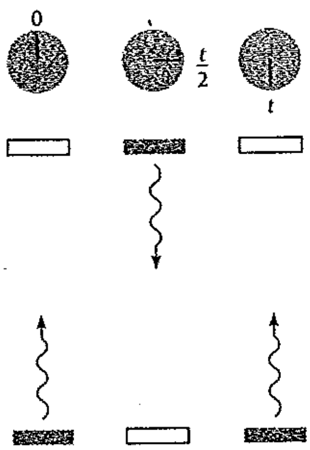


Figure 1.4 A light-pulse clock at rest on the ground as seen by an observer on the ground. The dial represents a conventional clock on the ground.

Figure 1.4 shows the laboratory clock in operation. The time interval between ticks is the proper time t_0 and the time needed for the light pulse to travel between the mirrors at the speed of light c is $t_0/2$. Hence $t_0/2 = L_0/c$ and

$$t_0 = \frac{2L_0}{c} \tag{1.1}$$

Figure 1.5 shows the moving clock with its mirrors perpendicular to the direction of motion relative to the ground. The time interval between ticks is t. Because the clock is moving, the light pulse, as seen from the ground, follows a zigzag path. On its way from the lower mirror to the upper one in the time $t/2$, the pulse travels a horizontal distance of $v(t/2)$ and a total distance of $c(t/2)$. Since L_0 is the vertical distance between the mirrors,

$$\left(\frac{ct}{2}\right)^2 = L_0^2 + \left(\frac{vt}{2}\right)^2$$

$$\frac{t^2}{4}(c^2 - v^2) = L_0^2$$

$$t^2 = \frac{4L_0^2}{c^2 - v^2} = \frac{(2L_0)^2}{c^2(1 - v^2/c^2)}$$

$$t = \frac{2L_0/c}{\sqrt{1 - v^2/c^2}} \tag{1.2}$$

But $2L_0/c$ is the time interval t_0 between ticks on the clock on the ground, as in Eq. (1.1), and so

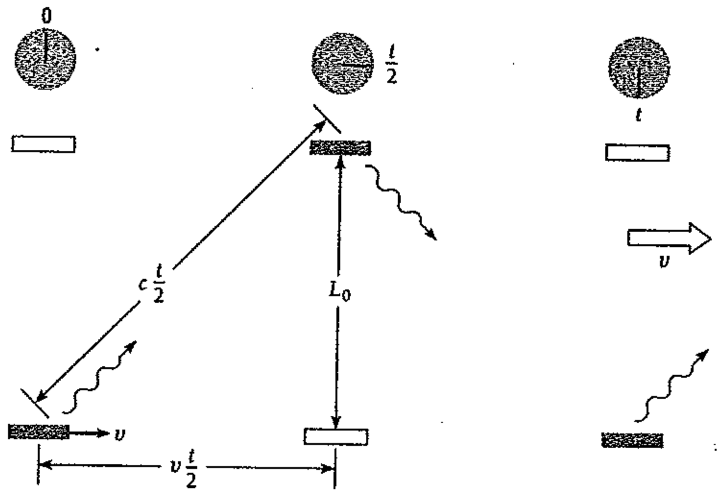


Figure 1.5 A light-pulse clock in a spacecraft as seen by an observer on the ground. The mirrors are parallel to the direction of motion of the spacecraft. The dial represents a conventional clock on the ground.

Time dilation
$$t = \frac{t_0}{\sqrt{1 - v^2/c^2}} \qquad (1.3)$$

Here is a reminder of what the symbols in Eq. (1.4) represent:

t_0 = time interval on clock at rest relative to an observer = proper time
t = time interval on clock in motion relative to an observer
v = speed of relative motion
c = speed of light

Because the quantity $\sqrt{1 - v^2/c^2}$ is always smaller than 1 for a moving object, t is always greater than t_0. The moving clock in the spacecraft appears to tick at a slower rate than the stationary one on the ground, as seen by an observer on the ground.

Exactly the same analysis holds for measurements of the clock on the ground by the pilot of the spacecraft. To him, the light pulse of the ground clock follows a zigzag path that requires a total time t per round trip. His own clock, at rest in the spacecraft, ticks at intervals of t_0. He too finds that

$$t = \frac{t_0}{\sqrt{1 - v^2/c^2}}$$

so the effect is reciprocal: *every* observer finds that clocks in motion relative to him tick more slowly than clocks at rest relative to him.

Our discussion has been based on a somewhat unusual clock. Do the same conclusions apply to ordinary clocks that use machinery—spring-controlled escapements, tuning forks, vibrating quartz crystals, or whatever—to produce ticks at constant time intervals? The answer must be yes, since if a mirror clock and a conventional clock in the spacecraft agree with each other on the ground but not when in flight, the disagreement between then could be used to find the speed of the spacecraft independently of any outside frame of reference—which contradicts the principle that all motion is relative.

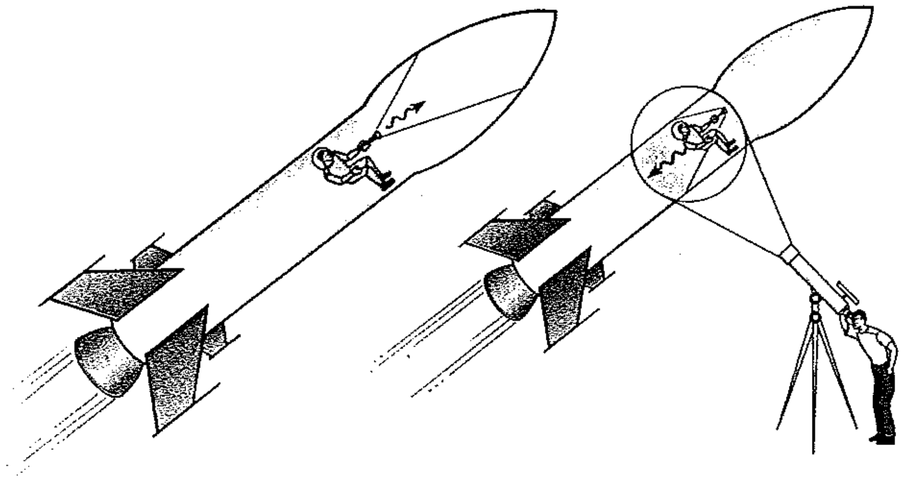
The Ultimate Speed Limit

The earth and the other planets of the solar system seem to be natural products of the evolution of the sun. Since the sun is a rather ordinary star in other ways, it is not surprising that other stars have been found to have planetary systems around them as well. Life developed here on earth, and there is no known reason why it should not also have done so on some of these planets. Can we expect ever to be able to visit them and meet our fellow citizens of the universe? The trouble is that nearly all stars are very far away—thousands or millions of light-years away. (A light-year, the distance light travels in a year, is 9.46×10^{15} m.) But if we can build a spacecraft whose speed is thousands or millions of times greater than the speed of light c, such distances would not be an obstacle.

Alas, a simple argument based on Einstein's postulates shows that nothing can move faster than c. Suppose you are in a spacecraft traveling at a constant speed v relative to the earth that is greater than c. As I watch from the earth, the lamps in the spacecraft suddenly go out. You switch on a flashlight to find the fuse box at the front of the spacecraft and change the blown fuse (Fig. 1.6*a*). The lamps go on again.

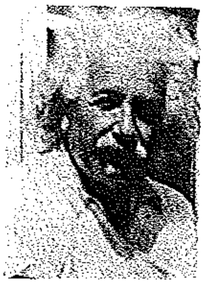
From the ground, though, I would see something quite different. To me, since your speed v is greater than c, the light from your flashlight illuminates the *back* of the spacecraft (Fig. 1.6*b*). I can only conclude that the laws of physics are different in your inertial frame from what they are in my inertial frame—which contradicts the principle of relativity. The only way to avoid this contradiction is to assume that nothing can move faster than the speed of light. This assumption has been tested experimentally many times and has always been found to be correct.

The speed of light c in relativity is always its value in free space of 3.00×10^8 m/s. In all material media, such as air, water, or glass, light travels more slowly than this, and atomic particles are able to move faster in such media than does light. When an electrically charged particle moves through a transparent substance at a speed exceeding that of light in the substance, a cone of light waves is emitted that corresponds to the bow wave produced by a ship moving through the water faster than water waves do. These light waves are known as **Cerenkov radiation** and form the basis of a method of determining the speeds of such particles. The minimum speed a particle must have to emit Cerenkov radiation is c/n in a medium whose index of refraction is n. Cerenkov radiation is visible as a bluish glow when an intense beam of particles is involved.



(*a*) (*b*)

Figure 1.6 A person switches on a flashlight in a spacecraft assumed to be moving relative to the earth faster than light. (*a*) In the spacecraft frame, the light goes to the front of the spacecraft. (*b*) In the earth frame, the light goes to the back of the spacecraft. Because observers in the spacecraft and on the earth would see different events, the principle of relativity would be violated. The conclusion is that the spacecraft cannot be moving faster than light relative to the earth (or relative to anything else).



(AIP Niels Bohr Library)

Albert Einstein (1879–1955), bitterly unhappy with the rigid discipline of the schools of his native Germany, went at sixteen to Switzerland to complete his education, and later got a job examining patent applications at the Swiss Patent Office. Then, in 1905, ideas that had been germinating in his mind for years when he should have been paying attention to other matters (one of his math teachers called Einstein a "lazy dog") blossomed into three short papers that were to change decisively the course not only of physics but of modern civilization as well.

The first paper, on the photoelectric effect, proposed that light has a dual character with both particle and wave properties. The subject of the second paper was Brownian motion, the irregular zigzag movement of tiny bits of suspended matter, such as pollen grains in water. Einstein showed that Brownian motion results from the bombardment of the particles by randomly moving molecules in the fluid in which they are suspended. This provided the long-awaited definite link with experiment that convinced the remaining doubters of the molecular theory of matter. The third paper introduced the special theory of relativity.

Although much of the world of physics was originally either indifferent or skeptical, even the most unexpected of Einstein's conclusions were soon confirmed and the development of what is now called modern physics began in earnest. After university posts in Switzerland and Czechoslovakia, in 1913 he took up an appointment at the Kaiser Wilhelm Institute in Berlin that left him able to do research free of financial worries and routine duties. Einstein's interest was now mainly in gravitation, and he started where Newton had left off more than two centuries earlier.

Einstein's general theory of relativity, published in 1916, related gravity to the structure of space and time. In this theory the force of gravity can be thought of as arising from a warping of spacetime around a body of matter so that a nearby mass tends to move toward it, much as a marble rolls toward the bottom of a saucer-shaped hole. From general relativity came a number of remarkable predictions, such as that light should be subject to gravity, all of which were verified experimentally. The later discovery that the universe is expanding fit neatly into the theory. In 1917 Einstein introduced the idea of stimulated emission of radiation, an idea that bore fruit forty years later in the invention of the laser.

The development of quantum mechanics in the 1920s disturbed Einstein, who never accepted its probabilistic rather than deterministic view of events on an atomic scale. "God does not play dice with the world," he said, but for once his physical intuition seemed to be leading him in the wrong direction.

Einstein, by now a world celebrity, left Germany in 1933 after Hitler came to power and spent the rest of his life at the Institute for Advanced Study in Princeton, New Jersey, thereby escaping the fate of millions of other European Jews at the hands of the Germans. His last years were spent in an unsuccessful search for a theory that would bring gravitation and electromagnetism together into a single picture, a problem worthy of his gifts but one that remains unsolved to this day.

Example 1.1

A spacecraft is moving relative to the earth. An observer on the earth finds that, between 1 P.M. and 2 P.M. according to her clock, 3601 s elapse on the spacecraft's clock. What is the spacecraft's speed relative to the earth?

Solution

Here $t_0 = 3600$ s is the proper time interval on the earth and $t = 3601$ s is the time interval in the moving frame as measured from the earth. We proceed as follows:

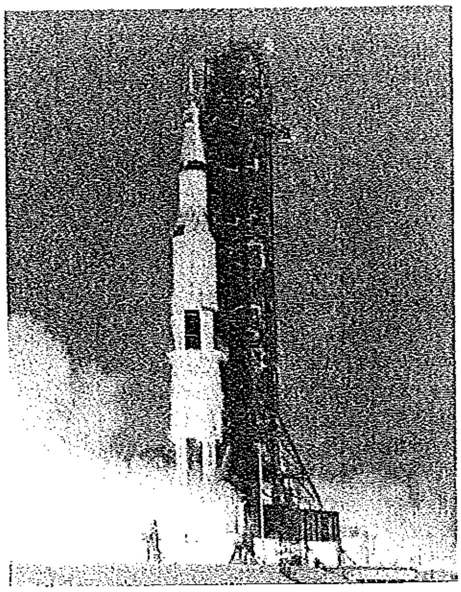
$$t = \frac{t_0}{\sqrt{1 - v^2/c^2}}$$

$$1 - \frac{v^2}{c^2} = \left(\frac{t_0}{t}\right)^2$$

$$v = c\sqrt{1 - \left(\frac{t_0}{t}\right)^2} = (2.998 \times 10^8 \text{ m/s})\sqrt{1 - \left(\frac{3600 \text{ s}}{3601 \text{ s}}\right)^2}$$

$$= 7.1 \times 10^6 \text{ m/s}$$

Today's spacecraft are much slower than this. For instance, the highest speed of the Apollo 11 spacecraft that went to the moon was only 10,840 m/s, and its clocks differed from those on the earth by less than one part in 10^9. Most of the experiments that have confirmed time dilation made use of unstable nuclei and elementary particles which readily attain speeds not far from that of light.



Apollo 11 lifts off its pad to begin the first human visit to the moon. At its highest speed of 10.8 km/s relative to the earth, its clocks differed from those on the earth by less than one part in a billion.

Although time is a relative quantity, not all the notions of time formed by everyday experience are incorrect. Time does not run backward to *any* observer, for instance. A sequence of events that occur at some particular point at $t_1, t_2, t_3, \ldots$ will appear in the same order to all observers everywhere, though not necessarily with the same time intervals $t_2 - t_1, t_3 - t_2, \ldots$ between each pair of events. Similarly, no distant observer, regardless of his or her state of motion, can see an event before it happens—more precisely, before a nearby observer sees it—since the speed of light is finite and signals require the minimum period of time L/c to travel a distance L. There is no way to peer into the future, although past events may appear different to different observers.

1.3 DOPPLER EFFECT

Why the universe is believed to be expanding

We are all familiar with the increase in pitch of a sound when its source approaches us (or we approach the source) and the decrease in pitch when the source recedes from us (or we recede from the source). These changes in frequency constitute the doppler effect, whose origin is straightforward. For instance, successive waves emitted by a source moving toward an observer are closer together than normal because of the advance of the source; because the separation of the waves is the wavelength of the sound, the corresponding frequency is higher. The relationship between the source frequency ν_0 and the observed frequency ν is

Doppler effect in sound

$$\nu = \nu_0\left(\frac{1 + v/c}{1 - V/c}\right) \tag{1.4}$$

where c = speed of sound
v = speed of observer (+ for motion toward the source, − for motion away from it)
V = speed of the source (+ for motion toward the observer, − for motion away from him)

If the observer is stationary, $v = 0$, and if the source is stationary, $V = 0$.

The doppler effect in sound varies depending on whether the source, or the observer, or both are moving. This appears to violate the principle of relativity: all that should count is the relative motion of source and observer. But sound waves occur only in a material medium such as air or water, and this medium is itself a frame of reference with respect to which motions of source and observer are measurable. Hence there is no contradiction. In the case of light, however, no medium is involved and only relative motion of source and observer is meaningful. The doppler effect in light must therefore differ from that in sound.

We can analyze the doppler effect in light by considering a light source as a clock that ticks ν_0 times per second and emits a wave of light with each tick. We will examine the three situations shown in Fig. 1.7.

1 *Observer moving perpendicular to a line between him and the light source.* The proper time between ticks is $t_0 = 1/\nu_0$, so between one tick and the next the time $t = t_0/\sqrt{1 - v^2/c^2}$ elapses in the reference frame of the observer. The frequency he finds is accordingly

$$\nu(\text{transverse}) = \frac{1}{t} = \frac{\sqrt{1 - v^2/c^2}}{t_0}$$

Transverse doppler effect in light

$$\nu = \nu_0\sqrt{1 - v^2/c^2} \tag{1.5}$$

The observed frequency ν is always lower than the source frequency ν_0.

2 *Observer receding from the light source.* Now the observer travels the distance vt away from the source between ticks, which means that the light wave from a given tick takes

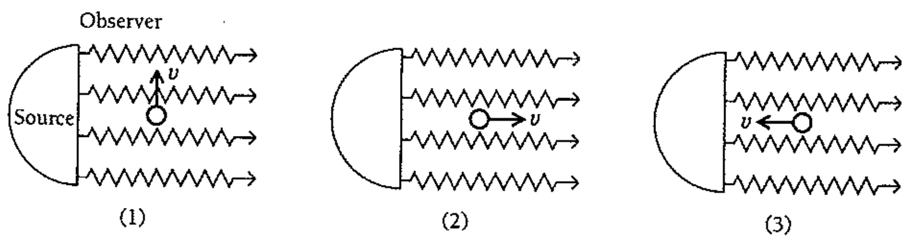


Figure 1.7 The frequency of the light seen by an observer depends on the direction and speed of the observer's motion relative to its source.

vt/c longer to reach him than the previous one. Hence the total time between the arrival of successive waves is

$$T = t + \frac{vt}{c} = t_0 \frac{1 + v/c}{\sqrt{1 - v^2/c^2}} = t_0 \frac{\sqrt{1 + v/c}\sqrt{1 + v/c}}{\sqrt{1 + v/c}\sqrt{1 - v/c}} = t_0 \sqrt{\frac{1 + v/c}{1 - v/c}}$$

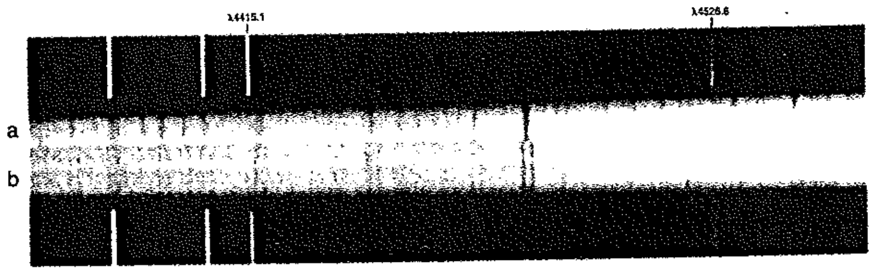
and the observed frequency is

$$\nu(\text{receding}) = \frac{1}{T} = \frac{1}{t_0}\sqrt{\frac{1 - v/c}{1 + v/c}} = \nu_0 \sqrt{\frac{1 - v/c}{1 + v/c}} \qquad (1.6)$$

The observed frequency ν is lower than the source frequency ν_0. Unlike the case of sound waves, which propagate relative to a material medium it makes no difference whether the observer is moving away from the source or the source is moving away from the observer.

3 *Observer approaching the light source:* The observer here travels the distance vt toward the source between ticks, so each light wave takes vt/c less time to arrive than the previous one. In this case $T = t - vt/c$ and the result is

$$\nu(\text{approaching}) = \nu_0 \sqrt{\frac{1 + v/c}{1 - v/c}} \qquad (1.7)$$



Spectra of the double star Mizar, which consists of two stars that circle their center of mass, taken 2 days apart. In *a* the stars are in line with no motion toward or away from the earth, so their spectral lines are superimposed. In *b* one star is moving toward the earth and the other is moving away from the earth, so the spectral lines of the former are doppler-shifted toward the blue end of the spectrum and those of the latter are shifted toward the red end.

The observed frequency is higher than the source frequency. Again, the same formula holds for motion of the source toward the observer.

Equations (1.6) and (1.7) can be combined in the single formula

Longitudinal doppler effect in light

$$\nu = \nu_0 \sqrt{\frac{1 + v/c}{1 - v/c}} \qquad (1.8)$$

by adopting the convention that v is + for source and observer approaching each other and − for source and observer receding from each other.

Example 1.2

A driver is caught going through a red light. The driver claims to the judge that the color she actually saw was green ($\nu = 5.60 \times 10^{14}$ Hz) and not red ($\nu_0 = 4.80 \times 10^{14}$ Hz) because of the doppler effect. The judge accepts this explanation and instead fines her for speeding at the rate of $1 for each km/h she exceeded the speed limit of 80 km/h. What was the fine?

Solution

Solving Eq. (1.8) for v gives

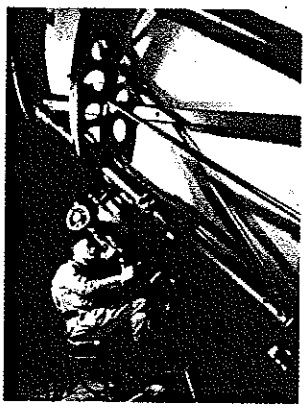
$$v = c\left(\frac{\nu^2 - \nu_0^2}{\nu^2 + \nu_0^2}\right) = (3.00 \times 10^8 \text{ m/s})\left[\frac{(5.60)^2 - (4.80)^2}{(5.60)^2 + (4.80)^2}\right]$$
$$= 4.59 \times 10^7 \text{ m/s} = 1.65 \times 10^8 \text{ km/h}$$

since 1 m/s = 3.6 km/h. The fine is therefore $\$(1.65 \times 10^8 - 80) = \$164{,}999{,}920$.

Visible light consists of electromagnetic waves in a frequency band to which the eye is sensitive. Other electromagnetic waves, such as those used in radar and in radio communications, also exhibit the doppler effect in accord with Eq. (1.8). Doppler shifts in radar waves are used by police to measure vehicle speeds, and doppler shifts in the radio waves emitted by a set of earth satellites formed the basis of the highly accurate Transit system of marine navigation.

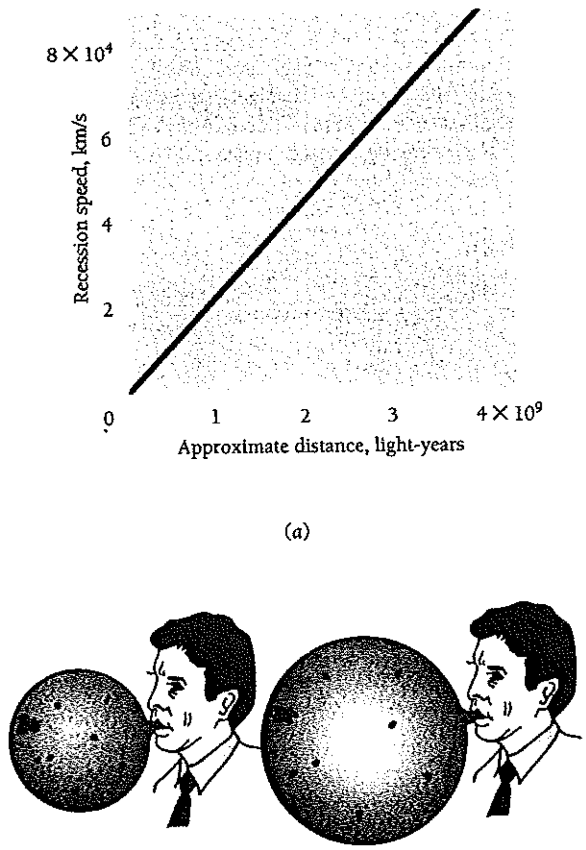
The Expanding Universe

The doppler effect in light is an important tool in astronomy. Stars emit light of certain characteristic frequencies called spectral lines, and motion of a star toward or away from the earth shows up as a doppler shift in these frequencies. The spectral lines of distant galaxies of stars are all shifted toward the low-frequency (red) end of the spectrum and hence are called "red shifts." Such shifts indicate that the galaxies are receding from us and from one another. The speeds of recession are observed to be



Edwin Hubble (1889–1953) was born in Missouri and, although always interested in astronomy, pursued a variety of other subjects as well at the University of Chicago. He then went as a Rhodes Scholar to Oxford University in England where he concentrated on law, Spanish, and heavyweight boxing. After two years of teaching at an Indiana high school, Hubble realized what his true vocation was and returned to the University of Chicago to study astronomy.

At Mt. Wilson Observatory in California, Hubble made the first accurate measurements of the distances of spiral galaxies which showed that they are far away in space from our own Milky Way galaxy. It had been known for some time that such galaxies have red shifts in their spectra that indicate motion away from the Milky Way, and Hubble joined his distance figures with the observed red shifts to conclude that the recession speeds were proportional to distance. This implies that the universe is expanding, a remarkable discovery that has led to the modern picture of the universe. Hubble was the first to use the 200-inch telescope, for many years the world's largest, at Mt. Palomar in California, in 1949. In his later work Hubble tried to determine the structure of the universe by finding how the concentration of remote galaxies varies with distance, a very difficult task that only today is being accomplished.



(*a*)

(*b*)

Figure 1.8 (*a*) Graph of recession speed versus distance for distant galaxies. The speed of recession averages about 21 km/s per million light-years. (*b*) Two-dimensional analogy of the expanding universe. As the balloon is inflated, the spots on it become farther apart. A bug on the balloon would find that the farther away a spot is from its location, the faster the spot seems to be moving away; this is true no matter where the bug is. In the case of the universe, the more distant a galaxy is from us, the faster it is moving away, which means that the universe is expanding uniformly.

proportional to distance, which suggests that the entire universe is expanding (Fig. 1.8). This proportionality is called **Hubble's law.**

The expansion apparently began about 13 billion years ago when a very small, intensely hot mass of primeval matter exploded, an event usually called the **Big Bang.** As described in Chap. 13, the matter soon turned into the electrons, protons, and neutrons of which the present universe is composed. Individual aggregates that formed during the expansion became the galaxies of today. Present data suggest that the current expansion will continue forever.

Example 1.3

A distant galaxy in the constellation Hydra is receding from the earth at 6.12×10^7 m/s. By how much is a green spectral line of wavelength 500 nm (1 nm $= 10^{-9}$ m) emitted by this galaxy shifted toward the red end of the spectrum?

Solution

Since $\lambda = c/\nu$ and $\lambda_0 = c/\nu_0$, from Eq. (1.6) we have

$$\lambda = \lambda_0 \sqrt{\frac{1 + v/c}{1 - v/c}}$$

Here $v = 0.204c$ and $\lambda_0 = 500$ nm, so

$$\lambda = 500 \text{ nm} \sqrt{\frac{1 + 0.204}{1 - 0.204}} = 615 \text{ nm}$$

which is in the orange part of the spectrum. The shift is $\lambda - \lambda_0 = 115$ nm. This galaxy is believed to be 2.9 billion light-years away.

1.4 LENGTH CONTRACTION

Faster means shorter

Measurements of lengths as well as of time intervals are affected by relative motion. The length L of an object in motion with respect to an observer always appears to the observer to be shorter than its length L_0 when it is at rest with respect to him. This contraction occurs only in the direction of the relative motion. The length L_0 of an object in its rest frame is called its **proper length.** (We note that in Fig. 1.5 the clock is moving perpendicular to v, hence $L = L_0$ there.)

The length contraction can be derived in a number of ways. Perhaps the simplest is based on time dilation and the principle of relativity. Let us consider what happens to unstable particles called muons that are created at high altitudes by fast cosmic-ray particles (largely protons) from space when they collide with atomic nuclei in the earth's atmosphere. A muon has a mass 207 times that of the electron and has a charge of either $+e$ or $-e$; it decays into an electron or a positron after an average lifetime of 2.2 μs (2.2×10^{-6} s).

Cosmic-ray muons have speeds of about 2.994×10^8 m/s ($0.998c$) and reach sea level in profusion—one of them passes through each square centimeter of the earth's surface on the average slightly more often than once a minute. But in $t_0 = 2.2\ \mu$s, their average lifetime, muons can travel a distance of only

$$vt_0 = (2.994 \times 10^8 \text{ m/s})(2.2 \times 10^{-6} \text{ s}) = 6.6 \times 10^2 \text{ m} = 0.66 \text{ km}$$

before decaying, whereas they are actually created at altitudes of 6 km or more.

To resolve the paradox, we note that the muon lifetime of $t_0 = 2.2\ \mu$s is what an observer at rest with respect to a muon would find. Because the muons are hurtling toward us at the considerable speed of $0.998c$, their lifetimes are extended in our frame of reference by time dilation to

$$t = \frac{t_0}{\sqrt{1 - v^2/c^2}} = \frac{2.2 \times 10^{-6} \text{ s}}{\sqrt{1 - (0.998c)^2/c^2}} = 34.8 \times 10^{-6} \text{ s} = 34.8\ \mu\text{s}$$

The moving muons have lifetimes almost 16 times longer than those at rest. In a time interval of 34.8 μs, a muon whose speed is $0.998c$ can cover the distance

$$vt = (2.994 \times 10^8 \text{ m/s})(34.8 \times 10^{-6} \text{ s}) = 1.04 \times 10^4 \text{ m} = 10.4 \text{ km}$$

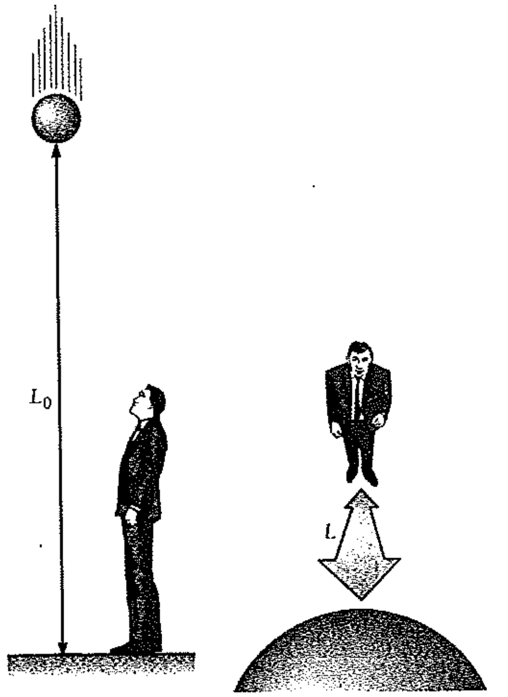


Figure 1.9 Muon decay as seen by different observers. The muon size is greatly exaggerated here; in fact, the muon seems likely to be a point particle with no extension in space.

Although its lifetime is only $t_0 = 2.2\ \mu s$ in its own frame of reference, a muon can reach the ground from altitudes of as much as 10.4 km because in the frame in which these altitudes are measured, the muon lifetime is $t = 34.8\ \mu s$.

What if somebody were to accompany a muon in its descent at $v = 0.998c$, so that to him or her the muon is at rest? The observer and the muon are now in the same frame of reference, and in this frame the muon's lifetime is only 2.2 μs. To the observer, the muon can travel only 0.66 km before decaying. The only way to account for the arrival of the muon at ground level is if the distance it travels, from the point of view of an observer in the moving frame, is shortened by virtue of its motion (Fig. 1.9). The principle of relativity tells us the extent of the shortening—it must be by the same factor of $\sqrt{1 - v^2/c^2}$ that the muon lifetime is extended from the point of view of a stationary observer.

We therefore conclude that an altitude we on the ground find to be h_0 must appear in the muon's frame of reference as the lower altitude

$$h = h_0 \sqrt{1 - v^2/c^2}$$

In our frame of reference the muon can travel $h_0 = 10.4$ km because of time dilation. In the muon's frame of reference, where there is no time dilation, this distance is abbreviated to

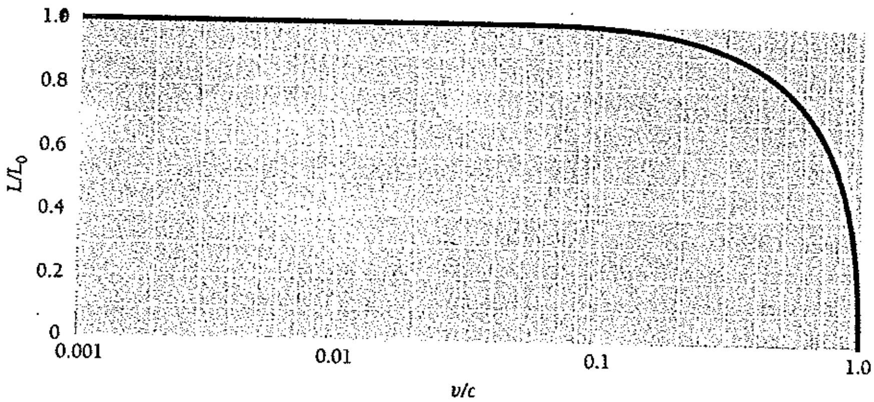


Figure 1.10 Relativistic length contraction. Only lengths in the direction of motion are affected. The horizontal scale is logarithmic.

$$h = (10.4 \text{ km}) \sqrt{1 - (0.998c)^2/c^2} = 0.66 \text{ km}$$

As we know, a muon traveling at $0.998c$ goes this far in 2.2 μs.

The relativistic shortening of distances is an example of the general contraction of lengths in the direction of motion:

Length contraction

$$L = L_0 \sqrt{1 - v^2/c^2} \tag{1.9}$$

Figure 1.10 is a graph of L/L_0 versus v/c. Clearly the length contraction is most significant at speeds near that of light. A speed of 1000 km/s seems fast to us, but it only results in a shortening in the direction of motion to 99.9994 percent of the proper length of an object moving at this speed. On the other hand, something traveling at nine-tenths the speed of light is shortened to 44 percent of its proper length, a significant change.

Like time dilation, the length contraction is a reciprocal effect. To a person in a spacecraft, objects on the earth appear shorter than they did when he or she was on the ground by the same factor of $\sqrt{1 - v^2/c^2}$ that the spacecraft appears shorter to somebody at rest. The proper length L_0 found in the rest frame is the maximum length any observer will measure. As mentioned earlier, only lengths in the direction of motion undergo contraction. Thus to an outside observer a spacecraft is shorter in flight than on the ground, but it is not narrower.

1.5 TWIN PARADOX

A longer life, but it will not seem longer

We are now in a position to understand the famous relativistic effect known as the twin paradox. This paradox involves two identical clocks, one of which remains on the earth while the other is taken on a voyage into space at the speed v and eventually is brought back. It is customary to replace the clocks with the pair of twins Dick and

Jane, a substitution that is perfectly acceptable because the processes of life—heartbeats, respiration, and so on—constitute biological clocks of reasonable regularity.

Dick is 20 y old when he takes off on a space voyage at a speed of $0.80c$ to a star 20 light-years away. To Jane, who stays behind, the pace of Dick's life is slower than hers by a factor of

$$\sqrt{1 - v^2/c^2} = \sqrt{1 - (0.80c)^2/c^2} = 0.60 = 60\%$$

To Jane, Dick's heart beats only 3 times for every 5 beats of her heart; Dick takes only 3 breaths for every 5 of hers; Dick thinks only 3 thoughts for every 5 of hers. Finally Dick returns after 50 years have gone by according to Jane's calendar, but to Dick the trip has taken only 30 y. Dick is therefore 50 y old whereas Jane, the twin who stayed home, is 70 y old (Fig. 1.11).

Where is the paradox? If we consider the situation from the point of view of Dick in the spacecraft, Jane on the earth is in motion relative to him at a speed of $0.80c$. Should not Jane then be 50 y old when the spacecraft returns, while Dick is then 70—the precise opposite of what was concluded above?

But the two situations are not equivalent. Dick changed from one inertial frame to a different one when he started out, when he reversed direction to head home, and when he landed on the earth. Jane, however, remained in the same inertial frame during Dick's whole voyage. The time dilation formula applies to Jane's observations of Dick, but not to Dick's observations of her.

To look at Dick's voyage from his perspective, we must take into account that the distance L he covers is shortened to

$$L = L_0\sqrt{1 - v^2/c^2} = (20 \text{ light-years})\sqrt{1 - (0.80c)^2/c^2} = 12 \text{ light-years}$$

To Dick, time goes by at the usual rate, but his voyage to the star has taken $L/v = 15$ y and his return voyage another 15 y, for a total of 30 y. Of course, Dick's life span has

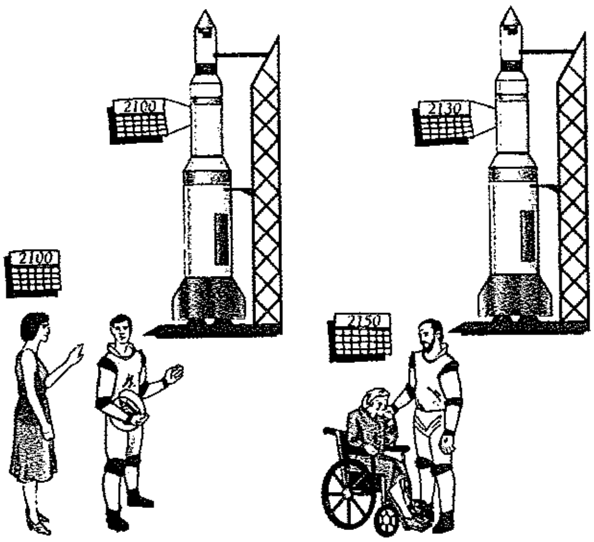


Figure 1.11 An astronaut who returns from a space voyage will be younger than his or her twin who remains on earth. Speeds close to the speed of light (here $v = 0.8c$) are needed for this effect to be conspicuous.

not been extended *to him*, because regardless of Jane's 50-y wait, he has spent only 30 y on the roundtrip.

The nonsymmetric aging of the twins has been verified by experiments in which accurate clocks were taken on an airplane trip around the world and then compared with identical clocks that had been left behind. An observer who departs from an inertial system and then returns after moving relative to that system will always find his or her clocks slow compared with clocks that stayed in the system.

Example 1.4

Dick and Jane each send out a radio signal once a year while Dick is away. How many signals does Dick receive? How many does Jane receive?

Solution

On the outward trip, Dick and Jane are being separated at a rate of 0.80*c*. With the help of the reasoning used to analyze the doppler effect in Sec. 1.3, we find that each twin receives signals

$$T_1 = t_0\sqrt{\frac{1+v/c}{1-v/c}} = (1\text{ y})\sqrt{\frac{1+0.80}{1-0.80}} = 3\text{ y}$$

apart. On the return trip, Dick and Jane are getting closer together at the same rate, and each receives signals more frequently, namely

$$T_2 = t_0\sqrt{\frac{1-v/c}{1+v/c}} = (1\text{ y})\sqrt{\frac{1-0.80}{1+0.80}} = \frac{1}{3}\text{ y}$$

apart.

To Dick, the trip to the star takes 15 y, and he receives $15/3 = 5$ signals from Jane. During the 15 y of the return trip, Dick receives $15/(1/3) = 45$ signals from Jane, for a total of 50 signals. Dick therefore concludes that Jane has aged by 50 y in his absence. Both Dick and Jane agree that Jane is 70 y old at the end of the voyage.

To Jane, Dick needs $L_0/v = 25$ y for the outward trip. Because the star is 20 light-years away, Jane on the earth continues to receive Dick's signals at the original rate of one every 3 y for 20 y after Dick has arrived at the star. Hence Jane receives signals every 3 y for 25 y + 20 y = 45 y to give a total of $45/3 = 15$ signals. (These are the 15 signals Dick sent out on the outward trip.) Then, for the remaining 5 y of what is to Jane a 50-y voyage, signals arrive from Dick at the shorter intervals of 1/3 y for an additional $5/(1/3) = 15$ signals. Jane thus receives 30 signals in all and concludes that Dick has aged by 30 y during the time he was away—which agrees with Dick's own figure. Dick is indeed 20 y younger than his twin Jane on his return.

1.6 ELECTRICITY AND MAGNETISM

Relativity is the bridge

One of the puzzles that set Einstein on the trail of special relativity was the connection between electricity and magnetism, and the ability of his theory to clarify the nature of this connection is one of its triumphs.

Because the moving charges (usually electrons) whose interactions give rise to many of the magnetic forces familiar to us have speeds far smaller than *c*, it is not obvious that the operation of an electric motor, say, is based on a relativistic effect. The idea becomes less implausible, however, when we reflect on the strength of electric forces. The electric attraction between the electron and proton in a hydrogen atom, for instance, is 10^{39} times greater than the gravitational attraction between them. Thus even a small change in the character of these forces due to relative motion, which is what magnetic forces represent, may have large consequences. Furthermore, although the effective speed of an individual electron in a current-carrying wire (<1 mm/s) is less than that of a tired caterpillar, there may be 10^{20} or more moving electrons per centimeter in such a wire, so the total effect may be considerable.

Although the full story of how relativity links electricity and magnetism is mathematically complex, some aspects of it are easy to appreciate. An example is the origin of the magnetic force between two parallel currents. An important point is that, like the speed of light,

Electric charge is relativistically invariant.

A charge whose magnitude is found to be Q in one frame of reference is also Q in all other frames.

Let us look at the two idealized conductors shown in Fig. 1.12*a*. They contain equal numbers of positive and negative charges at rest that are equally spaced. Because the conductors are electrically neutral, there is no force between them.

Figure 1.12*b* shows the same conductors when they carry currents i_I and i_{II} in the same direction. The positive charges move to the right and the negative charges move to the left, both at the same speed v as seen from the laboratory frame of reference. (Actual currents in metals consist of flows of negative electrons only, of course, but the electrically equivalent model here is easier to analyze and the results are the same.) Because the charges are moving, their spacing is smaller than before by the factor $\sqrt{1 - v^2/c^2}$. Since v is the same for both sets of charges, their spacings shrink by the same amounts, and both conductors remain neutral to an observer in the laboratory. However, the conductors now attract each other. Why?

Let us look at conductor II from the frame of reference of one of the negative charges in conductor I. Because the negative charges in II appear at rest in this frame, their spacing is not contracted, as in Fig. 1.12*c*. On the other hand, the positive charges in II now have the velocity $2v$, and their spacing is accordingly contracted to a greater extent than they are in the laboratory frame. Conductor II therefore appears to have a net positive charge, and an attractive force acts on the negative charge in I.

Next we look at conductor II from the frame of reference of one of the positive charges in conductor I. The positive charges in II are now at rest, and the negative charges there move to the left at the speed $2v$. Hence the negative charges are closer together than the positive ones, as in Fig. 1.12*d*, and the entire conductor appears negatively charged. An attractive force therefore acts on the positive charges in I.

Identical arguments show that the negative and positive charges in II are attracted to I. Thus all the charges in each conductor experience forces directed toward the other conductor. To each charge, the force on it is an "ordinary" electric force that arises because the charges of opposite sign in the other conductor are closer together than the charges of the same sign, so the other conductor appears to have a net charge. From the laboratory frame the situation is less straightforward. Both conductors are electrically neutral in this frame, and it is natural to explain their mutual attraction by attributing it to a special "magnetic" interaction between the currents.

A similar analysis explains the repulsive force between parallel conductors that carry currents in opposite directions. Although it is convenient to think of magnetic forces as being different from electric ones, they both result from a single electromagnetic interaction that occurs between charged particles.

Clearly a current-carrying conductor that is electrically neutral in one frame of reference might not be neutral in another frame. How can this observation be reconciled

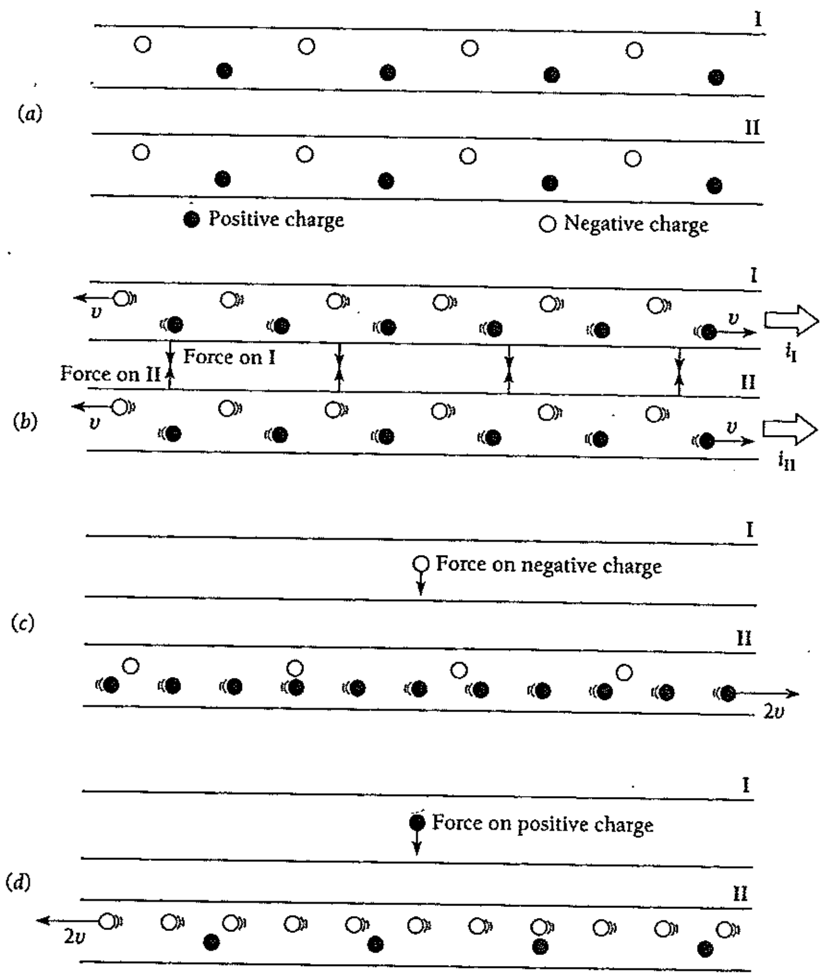


Figure 1.12 How the magnetic attraction between parallel currents arises. (*a*) Idealized parallel conductors that contain equal numbers of positive and negative charges. (*b*) When the conductors carry currents, the spacing of their moving charges undergoes a relativistic contraction as seen from the laboratory. The conductors attract each other when i_I and i_{II} are in the same direction. (*c*) As seen by a negative charge in I, the negative charges in II are at rest whereas the positive charges are in motion. The contracted spacing of the latter leads to a net positive charge in II that attracts the negative charge in I. (*d*) As seen by a positive charges in I, the positive charges in II are at rest whereas the negative charges are in motion. The contracted spacing of the latter leads to a net negative charge on II that attrats the positive charge in I. The contracted spacings in *b*, *c*, and *d* are greatly exaggerated.

with charge invariance? The answer is that we must consider the entire circuit of which the conductor is a part. Because the circuit must be closed for a current to occur in it, for every current element in one direction that a moving observer finds to have, say, a positive charge, there must be another current element in the opposite direction which the same observer finds to have a negative charge. Hence magnetic forces always act between different parts of the same circuit, even though the circuit as a whole appears electrically neutral to all observers.

The preceding discussion considered only a particular magnetic effect. All other magnetic phenomena can also be interpreted on the basis of Coulomb's law, charge invariance, and special relativity, although the analysis is usually more complicated.

1.7 RELATIVISTIC MOMENTUM

Redefining an important quantity

In classical mechanics linear momentum $\mathbf{p} = m\mathbf{v}$ is a useful quantity because it is conserved in a system of particles not acted upon by outside forces. When an event such as a collision or an explosion occurs inside an isolated system, the vector sum of the momenta of its particles before the event is equal to their vector sum afterward. We now have to ask whether $\mathbf{p} = m\mathbf{v}$ is valid as the definition of momentum in inertial frames in relative motion, and if not, what a relativistically correct definition is.

To start with, we require that $\mathbf{p}$ be conserved in a collision for all observers in relative motion at constant velocity. Also, we know that $\mathbf{p} = m\mathbf{v}$ holds in classical mechanics, that is, for $v \ll c$. Whatever the relativistically correct $\mathbf{p}$ is, then, it must reduce to $m\mathbf{v}$ for such velocities.

Let us consider an elastic collision (that is, a collision in which kinetic energy is conserved) between two particles A and B, as witnessed by observers in the reference frames S and S' which are in uniform relative motion. The properties of A and B are identical when determined in reference frames in which they are at rest. The frames S and S' are oriented as in Fig. 1.13, with S' moving in the $+x$ direction with respect to S at the velocity $\mathbf{v}$.

Before the collision, particle A had been at rest in frame S and particle B in frame S'. Then, at the same instant, A was thrown in the $+y$ direction at the speed V_A while B was thrown in the $-y'$ direction at the speed V'_B, where

$$V_A = V'_B \tag{1.10}$$

Hence the behavior of A as seen from S is exactly the same as the behavior of B as seen from S'.

When the two particles collide, A rebounds in the $-y$ direction at the speed V_A, while B rebounds in the $+y'$ direction at the speed V'_B. If the particles are thrown from positions Y apart, an observer in S finds that the collision occurs at $y = \frac{1}{2}Y$ and one in S' finds that it occurs at $y' = y = \frac{1}{2}Y$. The round-trip time T_0 for A as measured in frame S is therefore

$$T_0 = \frac{Y}{V_A} \tag{1.11}$$

and it is the same for B in S':

$$T_0 = \frac{Y}{V'_B}$$

In S the speed V_B is found from

$$V_B = \frac{Y}{T} \tag{1.12}$$

where T is the time required for B to make its round trip *as measured in S*. In S', however, B's trip requires the time T_0, where

$$T = \frac{T_0}{\sqrt{1 - v^2/c^2}} \tag{1.13}$$

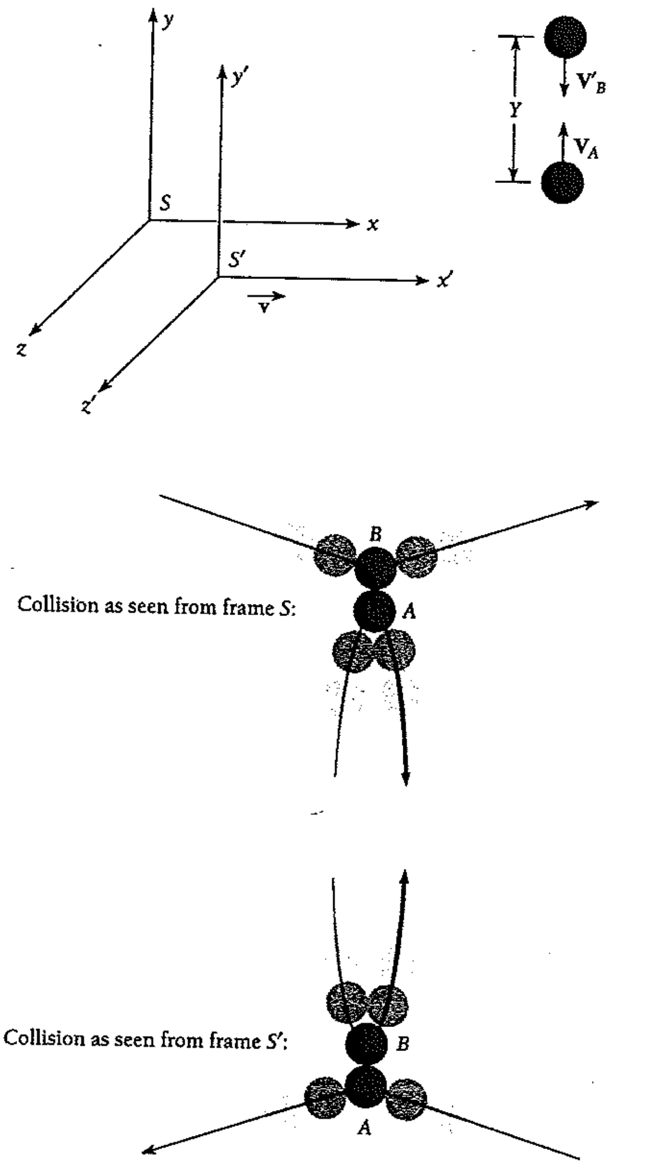


Figure 1.13 An elastic collision as observed in two different frames of reference. The balls are initially Y apart, which is the same distance in both frames since S' moves only in the x direction.

according to our previous results. Although observers in both frames see the same event, they disagree about the length of time the particle thrown from the other frame requires to make the collision and return.

Replacing T in Eq. (1.12) with its equivalent in terms of T_0, we have

$$V_B = \frac{Y\sqrt{1 - v^2/c^2}}{T_0}$$

From Eq. (1.11), $$V_A = \frac{Y}{T_0}$$

If we use the classical definition of momentum, $\mathbf{p} = m\mathbf{v}$, then in frame S

$$p_A = m_A V_A = m_A\left(\frac{Y}{T_0}\right)$$

$$p_B = m_B V_B = m_B \sqrt{1 - v^2/c^2}\left(\frac{Y}{T_0}\right)$$

This means that, in this frame, momentum will not be conserved if $m_A = m_B$, where m_A and m_B are the masses as measured in S. However, if

$$m_B = \frac{m_A}{\sqrt{1 - v^2/c^2}} \tag{1.14}$$

then momentum *will* be conserved.

In the collision of Fig. 1.13 both A and B are moving in both frames. Suppose now that V_A and V'_B are very small compared with v, the relative velocity of the two frames. In this case an observer in S will see B approach A with the velocity v, make a glancing collision (since $V'_B \ll v$), and then continue on. In the limit of $V_A = 0$, if m is the mass in S of A when A is at rest, then $m_A = m$. In the limit of $V'_B = 0$, if $m(v)$ is the mass in S of B, which is moving at the velocity v, then $m_B = m(v)$. Hence Eq. (1.14) becomes

$$m(v) = \frac{m}{\sqrt{1 - v^2/c^2}} \tag{1.15}$$

We can see that if linear momentum is defined as

Relativistic momentum

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - v^2/c^2}} \tag{1.16}$$

then conservation of momentum is valid in special relativity. When $v \ll c$, Eq. (1.16) becomes just $\mathbf{p} = m\mathbf{v}$, the classical momentum, as required. Equation (1.16) is often written as

Relativistic momentum

$$\mathbf{p} = \gamma m\mathbf{v} \tag{1.17}$$

where

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \tag{1.18}$$

In this definition, m is the **proper mass** (or **rest mass**) of an object, its mass when measured at rest relative to an observer. (The symbol γ is the Greek letter gamma.)

"Relativistic Mass"

We could alternatively regard the increase in an object's momentum over the classical value as being due to an increase in the object's mass. Then we would call $m_0 = m$ the rest mass of the object and $m = m(v)$ from Eq. (1.17) its relativistic mass, its mass when moving relative to an observer, so that $\mathbf{p} = m\mathbf{v}$. This is the view often taken in the past, at one time even by Einstein. However, as Einstein later wrote, the idea of relativistic mass is "not good" because "no clear definition can be given. It is better to introduce no other mass concept than the 'rest mass' m." In this book the term mass and the symbol m will always refer to proper (or rest) mass, which will be considered relativistically invariant.

Figure 1.14 shows how p varies with v/c for both γmv and mv. When v/c is small, mv and γmv are very nearly the same. (For $v = 0.01c$, the difference is only 0.005 percent; for $v = 0.1c$, it is 0.5 percent, still small). As v approaches c, however, the curve for γmv rises more and more steeply (for $v = 0.9c$, the difference is 229 percent). If $v = c$, $p = \gamma mv = \infty$, which is impossible. We conclude that no material object can travel as fast as light.

But what if a spacecraft moving at $v_1 = 0.5c$ relative to the earth fires a projectile at $v_2 = 0.5c$ in the same direction? We on earth might expect to observe the projectile's speed as $v_1 + v_2 = c$. Actually, as discussed in Appendix I to this chapter, velocity addition in relativity is not so simple a process, and we would find the projectile's speed to be only $0.8c$ in such a case.

Relativistic Second Law

In relativity Newton's second law of motion is given by

Relativistic second law

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(\gamma m\mathbf{v}) \qquad (1.19)$$

This is more complicated than the classical formula $\mathbf{F} = m\mathbf{a}$ because γ is a function of v. When $v \ll c$, γ is very nearly equal to 1, and $\mathbf{F}$ is very nearly equal to $m\mathbf{v}$, as it should be.

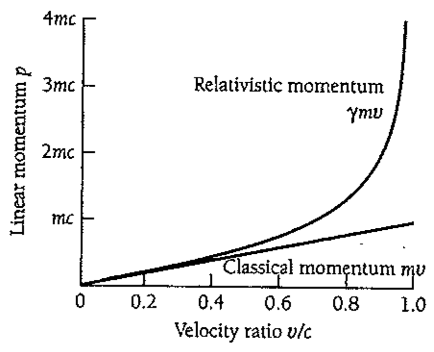


Figure 1.14 The momentum of an object moving at the velocity v relative to an observer. The mass m of the object is its value when it is at rest relative to the observer. The object's velocity can never reach c because its momentum would then be infinite, which is impossible. The relativistic momentum γmv is always correct; the classical momentum mv is valid for velocities much smaller than c.

Example 1.5

Find the acceleration of a particle of mass m and velocity $\mathbf{v}$ when it is acted upon by the constant force $\mathbf{F}$, where $\mathbf{F}$ is parallel to $\mathbf{v}$.

Solution

From Eq. (1.19), since $a = dv/dt$,

$$F = \frac{d}{dt}(\gamma m v) = m\frac{d}{dt}\left(\frac{v}{\sqrt{1 - v^2/c^2}}\right)$$

$$= m\left[\frac{1}{\sqrt{1 - v^2/c^2}} + \frac{v^2/c^2}{(1 - v^2/c^2)^{3/2}}\right]\frac{dv}{dt}$$

$$= \frac{ma}{(1 - v^2/c^2)^{3/2}}$$

We note that F is equal to $\gamma^3 ma$, *not* to γma. Merely replacing m by γm in classical formulas does not always give a relativistically correct result.

The acceleration of the particle is therefore

$$a = \frac{F}{m}(1 - v^2/c^2)^{3/2}$$

Even though the force is constant, the acceleration of the particle decreases as its velocity increases. As $v \to c$, $a \to 0$, so the particle can never reach the speed of light, a conclusion we expect.

1.8 MASS AND ENERGY

Where $E_0 = mc^2$ comes from

The most famous relationship Einstein obtained from the postulates of special relativity—how powerful they turn out to be!—concerns mass and energy. Let us see how this relationship can be derived from what we already know.

As we recall from elementary physics, the work W done on an object by a constant force of magnitude F that acts through the distance s, where $\mathbf{F}$ is in the same direction as $\mathbf{s}$, is given by $W = Fs$. If no other forces act on the object and the object starts from rest, all the work done on it becomes kinetic energy KE, so KE $= Fs$. In the general case where F need not be constant, the formula for kinetic energy is the integral

$$\text{KE} = \int_0^s F\,ds$$

In nonrelativistic physics, the kinetic energy of an object of mass m and speed v is KE $= \frac{1}{2}mv^2$. To find the correct relativistic formula for KE we start from the relativistic form of the second law of motion, Eq. (1.19), which gives

$$\text{KE} = \int_0^s \frac{d(\gamma m v)}{dt}\,ds = \int_0^{mv} v\,d(\gamma m v) = \int_0^v v\,d\left(\frac{mv}{\sqrt{1 - v^2/c^2}}\right)$$

Integrating by parts ($\int x\,dy = xy - \int y\,dx$),

$$\therefore\ \text{KE} = \frac{mv^2}{\sqrt{1 - v^2/c^2}} - m\int_0^v \frac{v\,dv}{\sqrt{1 - v^2/c^2}}$$

$$= \frac{mv^2}{\sqrt{1 - v^2/c^2}} + \left[mc^2\sqrt{1 - v^2/c^2}\right]_0^v$$

$$= \frac{mc^2}{\sqrt{1 - v^2/c^2}} - mc^2$$

Kinetic energy $$\text{KE} = \gamma mc^2 - mc^2 = (\gamma - 1)mc^2 \tag{1.20}$$

This result states that the kinetic energy of an object is equal to the difference between γmc^2 and mc^2. Equation (1.20) may be written

Total energy $$E = \gamma mc^2 = mc^2 + \text{KE} \tag{1.21}$$

If we interpret γmc^2 as the **total energy** E of the object, we see that when it is at rest and KE = 0, it nevertheless possesses the energy mc^2. Accordingly mc^2 is called the **rest energy** E_0 of something whose mass is m. We therefore have

$$E = E_0 + \text{KE}$$

where

Rest energy $$E_0 = mc^2 \tag{1.22}$$

If the object is moving, its total energy is

Total energy $$E = \gamma mc^2 = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \tag{1.23}$$

Example 1.6

A stationary body explodes into two fragments each of mass 1.0 kg that move apart at speeds of 0.6*c* relative to the original body. Find the mass of the original body.

Solution

The rest energy of the original body must equal the sum of the total energies of the fragments. Hence

$$E_0 = mc^2 = \gamma m_1c^2 + \gamma m_2c^2 = \frac{m_1c^2}{\sqrt{1 - v_1^2/c^2}} + \frac{m_2c^2}{\sqrt{1 - v_2^2/c^2}}$$

and

$$m = \frac{E_0}{c^2} = \frac{(2)(1.0\ \text{kg})}{\sqrt{1 - (0.60)^2}} = 2.5\ \text{kg}$$

Since mass and energy are not independent entities, their separate conservation principles are properly a single one—the principle of conservation of mass energy. Mass *can* be created or destroyed, but when this happens, an equivalent amount of energy simultaneously vanishes or comes into being, and vice versa. Mass and energy are different aspects of the same thing.

It is worth emphasizing the difference between a *conserved* quantity, such as total energy, and an *invariant* quantity, such as proper mass. Conservation of E means that, in a given reference frame, the total energy of some isolated system remains the same regardless of what events occur in the system. However, the total energy may be different as measured from another frame. On the other hand, the invariance of m means that m has the same value in all inertial frames.

The conversion factor between the unit of mass (the kilogram, kg) and the unit of energy (the joule, J) is c^2, so 1 kg of matter—the mass of this book is about that—has an energy content of $mc^2 = (1\text{ kg})(3 \times 10^8\text{ m/s})^2 = 9 \times 10^{16}$ J. This is enough to send a payload of a million tons to the moon. How is it possible for so much energy to be bottled up in even a modest amount of matter without anybody having been aware of it until Einstein's work?

In fact, processes in which rest energy is liberated are very familiar. It is simply that we do not usually think of them in such terms. In every chemical reaction that evolves energy, a certain amount of matter disappears, but the lost mass is so small a fraction of the total mass of the reacting substances that it is imperceptible. Hence the "law" of conservation of mass in chemistry. For instance, only about 6×10^{-11} kg of matter vanishes when 1 kg of dynamite explodes, which is impossible to measure directly, but the more than 5 million joules of energy that is released is hard to avoid noticing.

Example 1.7

Solar energy reaches the earth at the rate of about 1.4 kW per square meter of surface perpendicular to the direction of the sun (Fig. 1.15). By how much does the mass of the sun decrease per second owing to this energy loss? The mean radius of the earth's orbit is 1.5×10^{11} m.

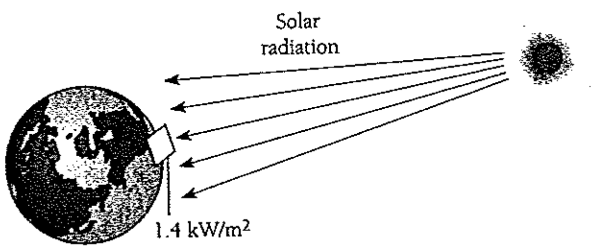


Figure 1.15

Solution

The surface area of a sphere of radius r is $A = 4\pi r^2$. The total power radiated by the sun, which is equal to the power received by a sphere whose radius is that of the earth's orbit, is therefore

$$P = \frac{P}{A}A = \frac{P}{A}(4\pi r^2) = (1.4 \times 10^3\text{ W/m}^2)(4\pi)(1.5 \times 10^{11}\text{ m})^2 = 4.0 \times 10^{26}\text{ W}$$

Thus the sun loses $E_0 = 4.0 \times 10^{26}$ J of rest energy per second, which means that the sun's rest mass decreases by

$$m = \frac{E_0}{c^2} = \frac{4.0 \times 10^{26}\text{ J}}{(3.0 \times 10^8\text{ m/s})^2} = 4.4 \times 10^9\text{ kg}$$

per second. Since the sun's mass is 2.0×10^{30} kg, it is in no immediate danger of running out of matter. The chief energy-producing process in the sun and most other stars is the conversion of hydrogen to helium in its interior. The formation of each helium nucleus is accompanied by the release of 4.0×10^{-11} J of energy, so 10^{37} helium nuclei are produced in the sun per second.

Kinetic Energy at Low Speeds

When the relative speed v is small compared with c, the formula for kinetic energy must reduce to the familiar $\frac{1}{2}mv^2$, which has been verified by experiment at such speeds. Let us see if this is true. The relativistic formula for kinetic energy is

Kinetic energy

$$\text{KE} = \gamma mc^2 - mc^2 = \frac{mc^2}{\sqrt{1 - v^2/c^2}} - mc^2 \tag{1.20}$$

Since $v^2/c^2 \ll 1$, we can use the binomial approximation $(1 + x)^n \approx 1 + nx$, valid for $|x| \ll 1$, to obtain

$$\frac{1}{\sqrt{1 - v^2/c^2}} \approx 1 + \frac{1}{2}\frac{v^2}{c^2} \qquad v \ll c$$

Thus we have the result

$$\text{KE} \approx \left(1 + \frac{1}{2}\frac{v^2}{c^2}\right)mc^2 - mc^2 \approx \frac{1}{2}mv^2 \qquad v \ll c$$

At low speeds the relativistic expression for the kinetic energy of a moving object does indeed reduce to the classical one. So far as is known, the correct formulation of mechanics has its basis in relativity, with classical mechanics representing an approximation that is valid only when $v \ll c$. Figure 1.16 shows how the kinetic energy of

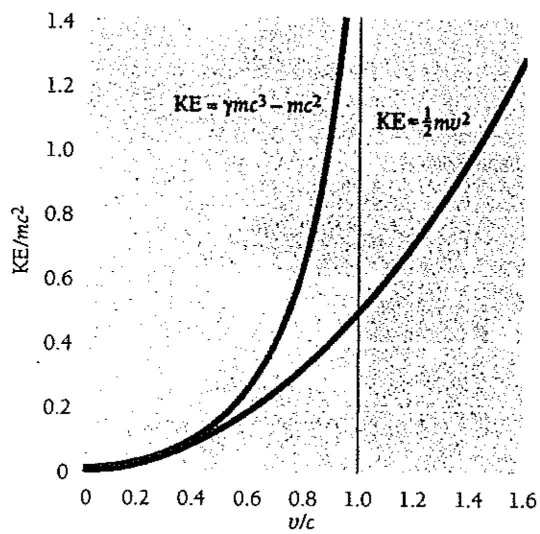


Figure 1.16 A comparison between the classical and relativistic formulas for the ratio between kinetic energy KE of a moving body and its rest energy mc^2. At low speeds the two formulas give the same result, but they diverge at speeds approaching that of light. According to relativistic mechanics, a body would need an infinite kinetic energy to travel with the speed of light, whereas in classical mechanics it would need only a kinetic energy of half its rest energy to have this speed.

a moving object varies with its speed according to both classical and relativistic mechanics.

The degree of accuracy required is what determines whether it is more appropriate to use the classical or to use the relativistic formulas for kinetic energy. For instance, when $v = 10^7$ m/s $(0.033c)$, the formula $\frac{1}{2}mv^2$ understates the true kinetic energy by only 0.08 percent; when $v = 3 \times 10^7$ m/s $(0.1c)$, it understates the true kinetic energy by 0.8 percent; but when $v = 1.5 \times 10^8$ m/s $(0.5c)$, the understatement is a significant 19 percent; and when $v = 0.999c$, the understatement is a whopping 4300 percent. Since 10^7 m/s is about 6310 mi/s, the nonrelativistic formula $\frac{1}{2}mv^2$ is entirely satisfactory for finding the kinetic energies of ordinary objects, and it fails only at the extremely high speeds reached by elementary particles under certain circumstances.

1.9 ENERGY AND MOMENTUM

How they fit together in relativity

Total energy and momentum are conserved in an isolated system, and the rest energy of a particle is invariant. Hence these quantities are in some sense more fundamental than velocity or kinetic energy, which are neither. Let us look into how the total energy, rest energy, and momentum of a particle are related.

We begin with Eq. (1.23) for total energy,

Total energy
$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \tag{1.23}$$

and square it to give

$$E^2 = \frac{m^2c^4}{1 - v^2/c^2}$$

From Eq. (1.17) for momentum,

Momentum
$$p = \frac{mv}{\sqrt{1 - v^2/c^2}} \tag{1.17}$$

we find that

$$p^2c^2 = \frac{m^2v^2c^2}{1 - v^2/c^2}$$

Now we subtract p^2c^2 from E^2:

$$E^2 - p^2c^2 = \frac{m^2c^4 - m^2v^2c^2}{1 - v^2/c^2} = \frac{m^2c^4(1 - v^2/c^2)}{1 - v^2/c^2}$$
$$= (mc^2)^2$$

Hence

Energy and momentum

$$E^2 = (mc^2)^2 + p^2c^2 \tag{1.24}$$

which is the formula we want. We note that, because mc^2 is invariant, so is $E^2 - p^2c^2$: this quantity for a particle has the same value in all frames of reference.

For a system of particles rather than a single particle, Eq. (1.24) holds provided that the rest energy mc^2—and hence mass m—is that of the entire system. If the particles in the system are moving with respect to one another, the sum of their individual rest energies may not equal the rest energy of the system. We saw this in Example 1.7 when a stationary body of mass 2.5 kg exploded into two smaller bodies, each of mass 1.0 kg, that then moved apart. If we were inside the system, we would interpret the difference of 0.5 kg of mass as representing its conversion into kinetic energy of the smaller bodies. But seen as a whole, the system is at rest both before and after the explosion, so the *system* did not gain kinetic energy. Therefore the rest energy of the system includes the kinetic energies of its internal motions and it corresponds to a mass of 2.5 kg both before and after the explosion.

In a given situation, the rest energy of an isolated system may be greater than, the same as, or less than the sum of the rest energies of its members. An important case in which the system rest energy is less than the rest energies of its members is that of a system of particles held together by attractive forces, such as the neutrons and protons in an atomic nucleus. The rest energy of a nucleus (except that of ordinary hydrogen, which is a single proton) is less than the total of the rest energies of its constituent particles. The difference is called the *binding energy* of the nucleus. To break a nucleus up completely calls for an amount of energy at least equal to its binding energy. This topic will be explored in detail in Sec. 11.4. For the moment it is interesting to note how large nuclear binding energies are—nearly 10^{12} kJ per kg of nuclear matter is typical. By comparison, the binding energy of water molecules in liquid water is only 2260 kJ/kg; this is the energy needed to turn 1 kg of water at 100°C to steam at the same temperature.

Massless Particles

Can a massless particle exist? To be more precise, can a particle exist which has no rest mass but which nevertheless exhibits such particlelike properties as energy and momentum? In classical mechanics, a particle must have rest mass in order to have energy and momentum, but in relativistic mechanics this requirement does not hold.

From Eqs. (1.17) and (1.23), when $m = 0$ and $v \ll c$, it is clear that $E = p = 0$. A massless particle with a speed less than that of light can have neither energy nor momentum. However, when $m = 0$ and $v = c$, $E = 0/0$ and $p = 0/0$, which are indeterminate: E and p can have any values. Thus Eqs. (1.17) and (1.23) are consistent with the existence of massless particles that possess energy and momentum *provided that they travel with the speed of light.*

Equation (1.24) gives us the relationship between E and p for a particle with $m = 0$:

Massless particle

$$E = pc \tag{1.25}$$

The conclusion is not that massless particles necessarily occur, only that the laws of physics do not exclude the possibility as long as $v = c$ and $E = pc$ for them. In fact, a massless particle—the photon—indeed exists and its behavior is as expected, as we shall find in Chap. 2.

Electronvolts

In atomic physics the usual unit of energy is the **electronvolt** (eV), where 1 eV is the energy gained by an electron accelerated through a potential difference of 1 volt. Since $W = QV$,

$$1 \text{ eV} = (1.602 \times 10^{-19} \text{ C})(1.000 \text{ V}) = 1.602 \times 10^{-19} \text{ J}$$

Two quantities normally expressed in electronvolts are the ionization energy of an atom (the work needed to remove one of its electrons) and the binding energy of a molecule (the energy needed to break it apart into separate atoms). Thus the ionization energy of nitrogen is 14.5 eV and the binding energy of the hydrogen molecule H_2 is 4.5 eV. Higher energies in the atomic realm are expressed in **kiloelectronvolts** (keV), where 1 keV $= 10^3$ eV.

In nuclear and elementary-particle physics even the keV is too small a unit in most cases, and the **megaelectronvolt** (MeV) and **gigaelectronvolt** (GeV) are more appropriate, where

$$1 \text{ MeV} = 10^6 \text{ eV} \qquad 1 \text{ GeV} = 10^9 \text{ eV}$$

An example of a quantity expressed in MeV is the energy liberated when the nucleus of a certain type of uranium atom splits into two parts. Each such fission event releases about 200 MeV; this is the process that powers nuclear reactors and weapons.

The rest energies of elementary particles are often expressed in MeV and GeV and the corresponding rest masses in MeV/c^2 and GeV/c^2. The advantage of the latter units is that the rest energy equivalent to a rest mass of, say, 0.938 GeV/c^2 (the rest mass of the proton) is just $E_0 = mc^2 = 0.938$ GeV. If the proton's kinetic energy is 5.000 GeV, finding its total energy is simple:

$$E = E_0 + \text{KE} = (0.938 + 5.000) \text{ GeV} = 5.938 \text{ GeV}$$

In a similar way the MeV/c and GeV/c are sometimes convenient units of linear momentum. Suppose we want to know the momentum of a proton whose speed is $0.800c$. From Eq. (1.17) we have

$$p = \frac{mv}{\sqrt{1 - v^2/c^2}} = \frac{(0.938 \text{ GeV}/c^2)(0.800c)}{\sqrt{1 - (0.800c)^2/c^2}}$$

$$= \frac{0.750 \text{ GeV}/c}{0.600} = 1.25 \text{ GeV}/c$$

Example 1.8

An electron ($m = 0.511$ MeV/c^2) and a photon ($m = 0$) both have momenta of 2.000 MeV/c. Find the total energy of each.

Solution

(*a*) From Eq. (1.24) the electron's total energy is

$$E = \sqrt{m^2c^4 + p^2c^2} = \sqrt{(0.511\ \text{MeV}/c^2)^2c^4 + (2.000\ \text{MeV}/c)^2c^2}$$
$$= \sqrt{(0.511\ \text{MeV})^2 + (2.000\ \text{MeV})^2} = 2.064\ \text{MeV}$$

(*b*) From Eq. (1.25) the photon's total energy is

$$E = pc = (2.000\ \text{MeV}/c)c = 2.000\ \text{MeV}$$

1.10 GENERAL RELATIVITY

Gravity is a warping of spacetime

Special relativity is concerned only with inertial frames of reference, that is, frames that are not accelerated. Einstein's 1916 **general theory of relativity** goes further by including the effects of accelerations on what we observe. Its essential conclusion is that the force of gravity arises from a warping of spacetime around a body of matter (Fig. 1.17). As a result, an object moving through such a region of space in general follows a curved path rather than a straight one, and may even be trapped there.

The **principle of equivalence** is central to general relativity:

> An observer in a closed laboratory cannot distinguish between the effects produced by a gravitational field and those produced by an acceleration of the laboratory.

This principle follows from the experimental observation (to better than 1 part in 10^{12}) that the inertial mass of an object, which governs the object's acceleration when a force acts on it, is always equal to its gravitational mass, which governs the gravitational force another object exerts on it. (The two masses are actually proportional; the constant of proportionality is set equal to 1 by an appropriate choice of the constant of gravitation G.)

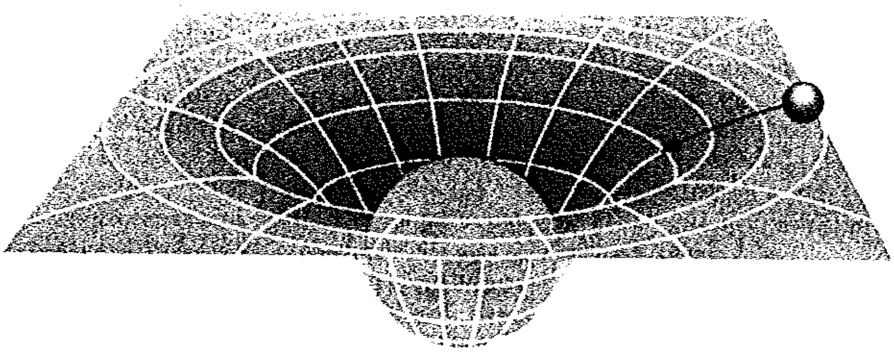


Figure 1.17 General relativity pictures gravity as a warping of spacetime due to the presence of a body of matter. An object nearby experiences an attractive force as a result of this distortion, much as a marble rolls toward the bottom of a depression in a rubber sheet. To paraphrase J. A. Wheeler, spacetime tells mass how to move, and mass tells spacetime how to curve.

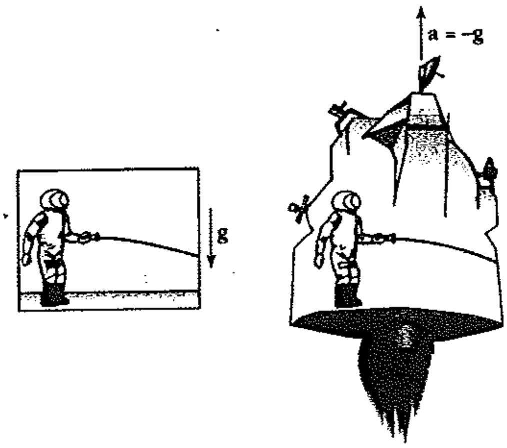


Figure 1.18 According to the principle of equivalence, events that take place in an accelerated laboratory cannot be distinguished from those which take place in a gravitational field. Hence the deflection of a light beam relative to an observer in an accelerated laboratory means that light must be similarly deflected in a gravitational field.

Gravity and Light

It follows from the principle of equivalence that light should be subject to gravity. If a light beam is directed across an accelerated laboratory, as in Fig. 1.18, its path relative to the laboratory will be curved. This means that, if the light beam is subject to the gravitational field to which the laboratory's acceleration is equivalent, the beam would follow the same curved path.

According to general relativity, light rays that graze the sun should have their paths bent toward it by 0.005°—the diameter of a dime seen from a mile away. This prediction was first confirmed in 1919 by photographs of stars that appeared in the sky near the sun during an eclipse, when they could be seen because the sun's disk was covered by the moon. The photographs were then compared with other photographs of the same part of the sky taken when the sun was in a distant part of the sky (Fig. 1.19). Einstein became a world celebrity as a result.

Because light is deflected in a gravitational field, a dense concentration of mass—such as a galaxy of stars—can act as a lens to produce multiple images of a distant light source located behind it (Fig. 1.20). A **quasar**, the nucleus of a young galaxy, is brighter than 100 billion stars but is no larger than the solar system. The first observation of gravitational lensing was the discovery in 1979 of what seemed to be a pair of nearby quasars but was actually a single one whose light was deviated by an intervening massive object. Since then a number of other gravitational lenses have been found; the effect occurs in radio waves from distant sources as well as in light waves.

The interaction between gravity and light also gives rise to the gravitational red shift and to black holes, topics that are considered in Chap. 2.

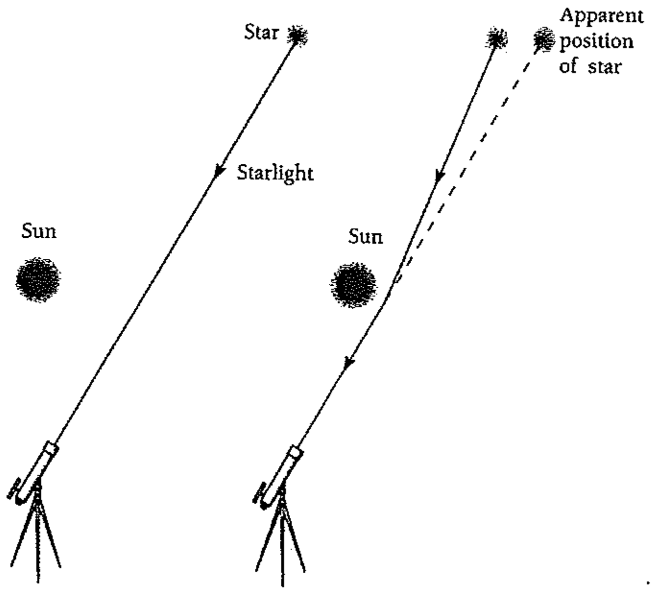


Figure 1.19 Starlight passing near the sun is deflected by its strong gravitational field. The deflection can be measured during a solar eclipse when the sun's disk is obscured by the moon.

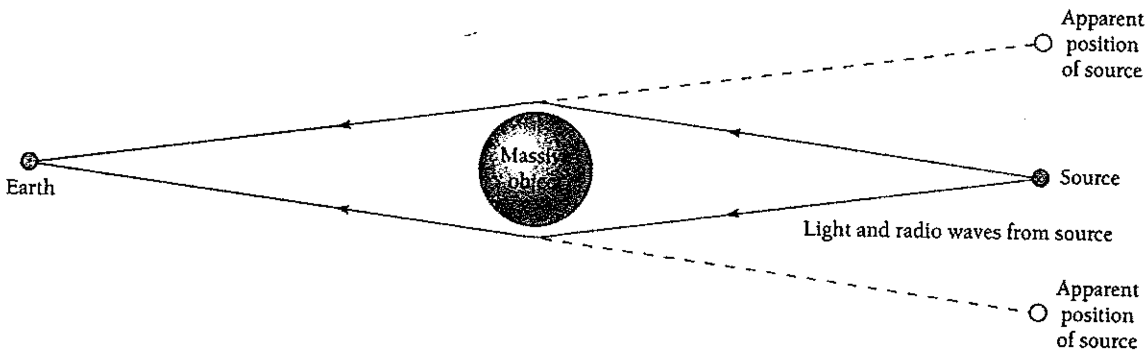


Figure 1.20 A gravitational lens. Light and radio waves from a source such as a quasar are deviated by a massive object such as a galaxy so that they seem to come from two or more identical sources. A number of such gravitational lenses have been identified.

Other Findings of General Relativity

A further success of general relativity was the clearing up of a long-standing puzzle in astronomy. The perihelion of a planetary orbit is the point in the orbit nearest the sun. Mercury's orbit has the peculiarity that its perihelion shifts (precesses) about 1.6° per century (Fig. 1.21). All but 43″ (1″ = 1 arc second = $\frac{1}{3600}$ of a degree) of this shift is due to the attractions of other planets, and for a while the discrepancy was used as evidence for an undiscovered planet called Vulcan whose orbit was supposed to lie

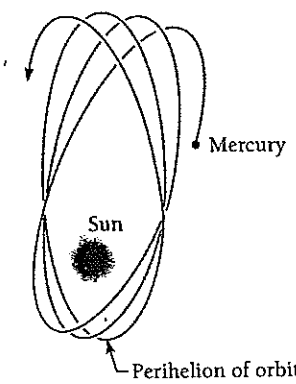


Figure 1.21 The precession of the perihelion of Mercury's orbit.

inside that of Mercury. When gravity is weak, general relativity gives very nearly the same results as Newton's formula $F = Gm_1m_2/r^2$. But Mercury is close to the sun and so moves in a strong gravitational field, and Einstein was able to show from general relativity that a precession of 43″ per century was to be expected for its orbit.

The existence of **gravitational waves** that travel with the speed of light was the prediction of general relativity that had to wait the longest to be verified. To visualize gravitational waves, we can think in terms of the model of Fig. 1.17 in which two-dimensional space is represented by a rubber sheet distorted by masses embedded in it. If one of the masses vibrates, waves will be sent out in the sheet that set other masses in vibration. A vibrating electric charge similarly sends out electromagnetic waves that excite vibrations in other charges.

A big difference between the two kinds of waves is that gravitational waves are extremely weak, so that despite much effort none have as yet been directly detected. However, in 1974 strong evidence for gravitational waves was found in the behavior of a system of two nearby stars, one a pulsar, that revolve around each other. A **pulsar** is a very small, dense star, composed mainly of neutrons, that spins rapidly and sends out flashes of light and radio waves at a regular rate, much as the rotating beam of a lighthouse does (see Sec. 9.11). The pulsar in this particular binary system emits pulses every 59 milliseconds (ms), and it and its companion (probably another neutron star) have an orbital period of about 8 h. According to general relativity, such a system should give off gravitational waves and lose energy as a result, which would reduce the orbital period as the stars spiral in toward each other. A change in orbital period means a change in the arrival times of the pulsar's flashes, and in the case of the observed binary system the orbital period was found to be decreasing at 75 ms per year. This is so close to the figure that general relativity predicts for the system that there seems to be no doubt that gravitational radiation is responsible. The 1993 Nobel Prize in physics was awarded to Joseph Taylor and Russell Hulse for this work.

Much more powerful sources of gravitational waves ought to be such events as two black holes colliding and supernova explosions in which the remnant star cores collapse into neutron stars (again, see Sec. 9.11). A gravitational wave that passes through a body of matter will cause distortions to ripple through it due to fluctuations in the gravitational field. Because gravitational forces are feeble—the electric attraction between a proton and an electron is over 10^{39} times greater than the gravitational attraction between them—such distortions at the earth induced by gravitational waves from a supernova in our galaxy (which occurs an average of once every 30 years or so) would amount to only about 1 part in 10^{18}, even less for a more distant supernova. This corresponds to a change in, say, the height of a person by well under the diameter of an atomic nucleus, yet it seems to be detectable—just—with current technology.

In one method, a large metal bar cooled to a low temperature to minimize the random thermal motions of its atoms is monitored by sensors for vibrations due to gravitational waves. In another method, an interferometer similar to the one shown in Fig. 1.2 with a laser as the light source is used to look for changes in the lengths of the arms to which the mirrors are attached. Instruments of both kinds are operating, thus far with no success.

A really ambitious scheme has been proposed that would use six spacecraft in orbit around the sun placed in pairs at the corners of a triangle whose sides are 5 million kilometers (km) long. Lasers, mirrors, and sensors in the spacecraft would detect changes in their spacings resulting from the passing of a gravitational wave. It may only be a matter of time before gravitational waves will be providing information about a variety of cosmic disturbances on the largest scale.

Appendix I to Chapter 1

The Lorentz Transformation

Suppose we are in an inertial frame of reference S and find the coordinates of some event that occurs at the time t are x, y, z. An observer located in a different inertial frame S' which is moving with respect to S at the constant velocity $\mathbf{v}$ will find that the same event occurs at the time t' and has the coordinates x', y', z'. (In order to simplify our work, we shall assume that $\mathbf{v}$ is in the $+x$ direction, as in Fig. 1.22.) How are the measurements x, y, z, t related to x', y', z', t'?

Galilean Transformation

Before special relativity, transforming measurements from one inertial system to another seemed obvious. If clocks in both systems are started when the origins of S and S' coincide, measurements in the x direction made is S will be greater than those made in S' by the amount vt, which is the distance S' has moved in the x direction. That is,

$$x' = x - vt \tag{1.26}$$

There is no relative motion in the y and z directions, and so

$$y' = y \tag{1.27}$$

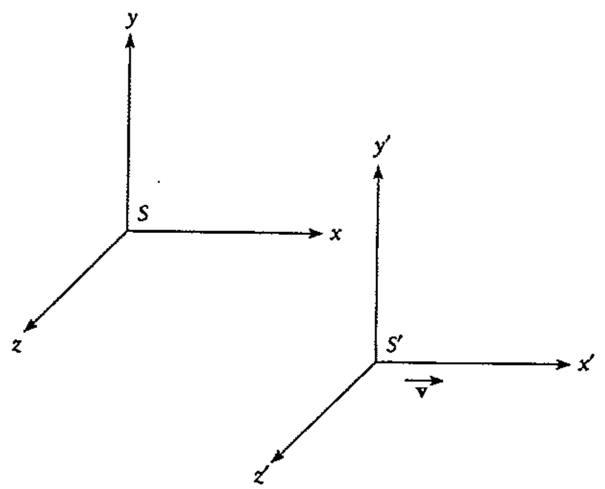


Figure 1.22 Frame S' moves in the $+x$ direction with the speed v relative to frame S. The Lorentz transformation must be used to convert measurements made in one of these frames to their equivalents in the other.

$$z' = z \tag{1.28}$$

In the absence of any indication to the contrary in our everyday experience, we further assume that

$$t' = t \tag{1.29}$$

The set of Eqs. (1.26) to (1.29) is known as the **Galilean transformation.**

To convert velocity components measured in the S frame to their equivalents in the S' frame according to the Galilean transformation, we simply differentiate x', y', and z' with respect to time:

$$v'_x = \frac{dx'}{dt'} = v_x - v \tag{1.30}$$

$$v'_y = \frac{dy'}{dt'} = v_y \tag{1.31}$$

$$v'_z = \frac{dz'}{dt'} = v_z \tag{1.32}$$

Although the Galilean transformation and the corresponding velocity transformation seem straightforward enough, they violate both of the postulates of special relativity. The first postulate calls for the same equations of physics in both the S and S' inertial frames, but the equations of electricity and magnetism become very different when the Galilean transformation is used to convert quantities measured in one frame into their equivalents in the other. The second postulate calls for the same value of the speed of light c whether determined in S or S'. If we measure the speed of light in the x direction in the S system to be c, however, in the S' system it will be

$$c' = c - v$$

according to Eq. (1.30). Clearly a different transformation is required if the postulates of special relativity are to be satisfied. We would expect both time dilation and length contraction to follow naturally from this new transformation.

Lorentz Transformation

A reasonable guess about the nature of the correct relationship between x and x' is

$$x' = k(x - vt) \tag{1.33}$$

Here k is a factor that does not depend upon either x or t but may be a function of v. The choice of Eq. (1.33) follows from several considerations:

1 It is linear in x and x', so that a single event in frame S corresponds to a single event in frame S', as it must.
2 It is simple, and a simple solution to a problem should always be explored first.
3 It has the possibility of reducing to Eq. (1.26), which we know to be correct in ordinary mechanics.

Because the equations of physics must have the same form in both S and S', we need only change the sign of v (in order to take into account the difference in the direction of relative motion) to write the corresponding equation for x in terms of x' and t':

$$x = k(x' + vt') \tag{1.34}$$

The factor k must be the same in both frames of reference since there is no difference between S and S' other than in the sign of v.

As in the case of the Galilean transformation, there is nothing to indicate that there might be differences between the corresponding coordinates y, y' and z, z' which are perpendicular to the direction of v. Hence we again take

$$y' = y \tag{1.35}$$

$$z' = z \tag{1.36}$$

The time coordinates t and t', however, are *not* equal. We can see this by substituting the value of x' given by Eq. (1.33) into Eq. (1.34). This gives

$$x = k^2(x - vt) + kvt'$$

from which we find that

$$t' = kt + \left(\frac{1 - k^2}{kv}\right)x \tag{1.37}$$

Equations (1.33) and (1.35) to (1.37) constitute a coordinate transformation that satisfies the first postulate of special relativity.

The second postulate of relativity gives us a way to evaluate k. At the instant $t = 0$, the origins of the two frames of reference S and S' are in the same place, according to our initial conditions, and $t' = 0$ then also. Suppose that a flare is set off at the common origin of S and S' at $t = t' = 0$, and the observers in each system measure the speed with which the flare's light spreads out. Both observers must find the same speed c (Fig. 1.23), which means that in the S frame

$$x = ct \tag{1.38}$$

and in the S' frame

$$x' = ct' \tag{1.39}$$

Substituting for x' and t' in Eq. (1.39) with the help of Eqs. (1.33) and (1.37) gives

$$k(x - vt) = ckt + \left(\frac{1 - k^2}{kv}\right)cx$$

and solving for x,

$$x = \frac{ckt + vkt}{k - \left(\dfrac{1 - k^2}{kv}\right)c} = ct\left[\frac{k + \dfrac{v}{c}k}{k - \left(\dfrac{1 - k^2}{kv}\right)c}\right] = ct\left[\frac{1 + \dfrac{v}{c}}{1 - \left(\dfrac{1}{k^2} - 1\right)\dfrac{c}{v}}\right]$$

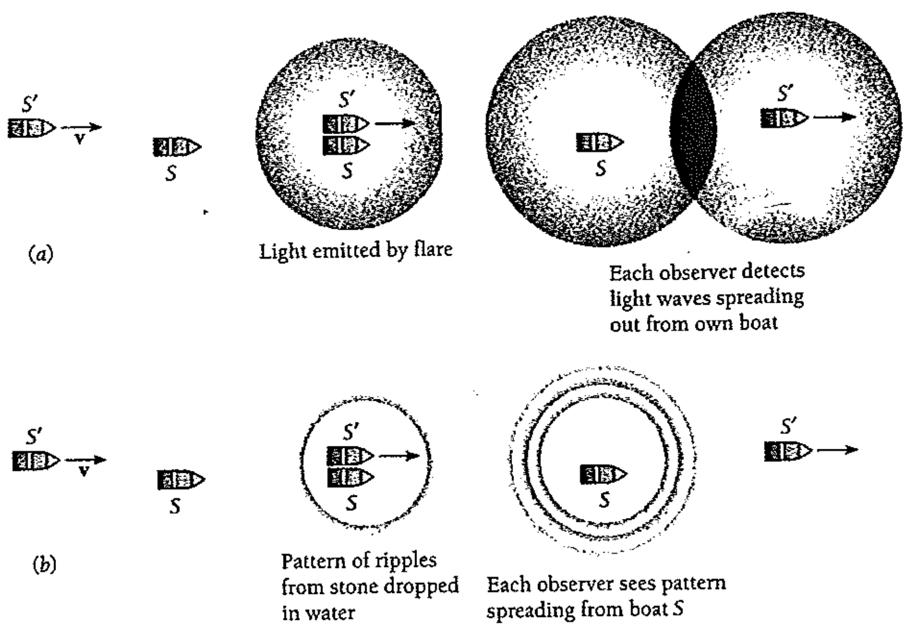


Figure 1.23 (*a*) Inertial frame S' is a boat moving at speed v in the $+x$ direction relative to another boat, which is the inertial frame S. When $t = t_0 = 0$, S' is next to S, and $x = x_0 = 0$. At this moment a flare is fired from one of the boats. An observer on boat S detects light waves spreading out at speed c from his boat. An observer on boat S' also detects light waves spreading out at speed c from her boat, even though S' is moving to the right relative to S. (*b*) If instead a stone were dropped in the water at $t = t_0 = 0$, the observers would find a pattern of ripples spreading out around S at different speeds relative to their boats. The difference between (*a*) and (*b*) is that water, in which the ripples move, is itself a frame of reference whereas space, in which light moves, is not.

This expression for x will be the same as that given by Eq. (1.38), namely, $x = ct$, provided that the quantity in the brackets equals 1. Therefore

$$y' = y \qquad (1.42)$$

$$z' = z \qquad (1.43)$$

$$t' = \frac{t - \dfrac{vx}{c^2}}{\sqrt{1 - v^2/c^2}} \qquad (1.44)$$

These equations comprise the **Lorentz transformation.** They were first obtained by the Dutch physicist H.A. Lorentz, who showed that the basic formulas of electromagnetism are the same in all inertial frames only when Eqs. (1.41) to (1.44) are used. It was not until several years later that Einstein discovered their full significance. It is obvious that the Lorentz transformation reduces to the Galilean transformation when the relative velocity v is small compared with the velocity of light c.

Example 1.9

Derive the relativistic length contraction using the Lorentz transformation.

Solution

Let us consider a rod lying along the x' axis in the moving frame S'. An observer in this frame determines the coordinates of its ends to be x'_1 and x'_2, and so the proper length of the rod is

$$L_0 = x'_2 - x'_1$$

Hendrik A. Lorentz (1853–1928) was born in Arnhem, Holland, and studied at the University of Leyden. At nineteen he returned to Arnhem and taught at the high school there while preparing a doctoral thesis that extended Maxwell's theory of electromagnetism to cover the details of the refraction and reflection of light. In 1878 he became professor of theoretical physics at Leyden, the first such post in Holland, where he remained for thirty-four years until he moved to Haarlem. Lorentz went on to reformulate and simplify Maxwell's theory and to introduce the idea that electromagnetic fields are created by electric charges on the atomic level. He proposed that the emission of light by atoms and various optical phenomena could be traced to the motions and interactions of atomic electrons. The discovery in 1896 by Pieter Zeeman, a student of his, that the spectral lines of atoms that radiate in a magnetic field are split into components of slightly different frequency confirmed Lorentz's work and led to a Nobel Prize for both of them in 1902.

The set of equations that enables electromagnetic quantities in one frame of reference to be transformed into their values in another frame of reference moving relative to the first were found by Lorentz in 1895, although their full significance was not realized until Einstein's theory of special relativity ten years afterward. Lorentz (and, independently, the Irish physicist G. F. Fitzgerald) suggested that the negative result of the Michelson-Morley experiment could be understood if lengths in the direction of motion relative to an observer were contracted. Subsequent experiments showed that although such contractions do occur, they are not the real reason for the Michelson-Morley result, which is that there is no "ether" to serve as a universal frame of reference.

In order to find $L = x_2 - x_1$, the length of the rod as measured in the stationary frame S at the time t, we make use of Eq. (1.41) to give

$$x'_1 = \frac{x_1 - vt}{\sqrt{1 - v^2/c^2}} \qquad x'_2 = \frac{x_2 - vt}{\sqrt{1 - v^2/c^2}}$$

Hence $$L = x_2 - x_1 = (x'_2 - x'_1)\sqrt{1 - v^2/c^2} = L_0\sqrt{1 - v^2/c^2}$$

This is the same as Eq. (1.9)

Inverse Lorentz Transformation

In Example 1.9 the coordinates of the ends of the moving rod were measured in the stationary frame S at the same time t, and it was easy to use Eq. (1.41) to find L in terms of L_0 and v. If we want to examine time dilation, though, Eq. (1.44) is not convenient, because t_1 and t_2, the start and finish of the chosen time interval, must be measured when the moving clock is at the respective *different* positions x_1 and x_2. In situations of this kind it is easier to use the **inverse Lorentz transformation,** which converts measurements made in the moving frame S' to their equivalents in S.

To obtain the inverse transformation, primed and unprimed quantities in Eqs. (1.41) to (1.44) are exchanged, and v is replaced by $-v$:

Inverse Lorentz transformation

$$x = \frac{x' + vt'}{\sqrt{1 - v^2/c^2}} \tag{1.45}$$

$$y = y' \tag{1.46}$$

$$z' = z' \tag{1.47}$$

$$t = \frac{t' + \dfrac{vx'}{c^2}}{\sqrt{1 - v^2/c^2}} \tag{1.48}$$

Example 1.10

Derive the formula for time dilation using the inverse Lorentz transformation.

Solution

Let us consider a clock at the point x' in the moving frame S'. When an observer in S' finds that the time is t'_1, an observer in S will find it to be t_1, where, from Eq. (1.48),

$$t_1 = \frac{t'_1 + \dfrac{vx'}{c^2}}{\sqrt{1 - v^2/c^2}}$$

After a time interval of t_0 (to him), the observer in the moving system finds that the time is now t'_2 according to his clock. That is,

$$t_0 = t'_2 - t'_1$$

The observer in S, however, measures the end of the same time interval to be

$$t_2 = \frac{t'_2 + \dfrac{vx'}{c^2}}{\sqrt{1 - v^2/c^2}}$$

so to her the duration of the interval t is

$$t = t_2 - t_1 = \frac{t'_2 - t'_1}{\sqrt{1 - v^2/c^2}} = \frac{t_0}{\sqrt{1 - v^2/c^2}}$$

This is what we found earlier with the help of a light-pulse clock.

Velocity Addition

Special relativity postulates that the speed of light c in free space has the same value for all observers, regardless of their relative motion. "Common sense" (which means here the Galilean transformation) tells us that if we throw a ball forward at 10 m/s from a car moving at 30 m/s, the ball's speed relative to the road will be 40 m/s, the sum of the two speeds. What if we switch on the car's headlights when its speed is v? The same reasoning suggests that their light, which is emitted from the reference frame S' (the car) in the direction of its motion relative to another frame S (the road), ought to have a speed of $c + v$ as measured in S. But this violates the above postulate, which has had ample experimental verification. Common sense is no more reliable as a guide in science than it is elsewhere, and we must turn to the Lorentz transformation equations for the correct scheme of velocity addition.

Suppose something is moving relative to both S and S'. An observer in S measures its three velocity components to be

$$V_x = \frac{dx}{dt} \qquad V_y = \frac{dy}{dt} \qquad V_z = \frac{dz}{dt}$$

while to an observer in S' they are

$$V'_x = \frac{dx'}{dt'} \qquad V'_y = \frac{dy'}{dt'} \qquad V'_z = \frac{dz'}{dt'}$$

By differentiating the inverse Lorentz transformation equations for x, y, z, and t, we obtain

$$dx = \frac{dx' + v\,dt'}{\sqrt{1 - v^2/c^2}} \qquad dy = dy' \qquad dz = dz' \qquad dt = \frac{dt' + \dfrac{v\,dz'}{c^2}}{\sqrt{1 - v^2/c^2}}$$

and so

$$V_x = \frac{dx}{dt} = \frac{dx' + v\,dt'}{dt' + \dfrac{v\,dx'}{c^2}} = \frac{\dfrac{dx'}{dt'} + v}{1 + \dfrac{v}{c^2}\dfrac{dx'}{dt'}}$$

Relativistic velocity transformation

$$V_x = \frac{V'_x + v}{1 + \frac{vV'_x}{c^2}} \tag{1.49}$$

Similarly,

$$V_y = \frac{V'_y\sqrt{1 - v^2/c^2}}{1 + \frac{vV'_x}{c^2}} \tag{1.50}$$

$$V_z = \frac{V'_z\sqrt{1 - v^2/c^2}}{1 + \frac{vV'_x}{c^2}} \tag{1.51}$$

If $V'_x = c$, that is, if light is emitted in the moving frame S' in its direction of motion relative to S, an observer in frame S will measure the speed

$$V_x = \frac{V'_x + v}{1 + \frac{vV'_x}{c^2}} = \frac{c + v}{1 + \frac{vc}{c^2}} = \frac{c(c + v)}{c + v} = c$$

Thus observers in the car and on the road both find the same value for the speed of light, as they must.

Example 1.11

Spacecraft Alpha is moving at 0.90c with respect to the earth. If spacecraft Beta is to pass Alpha at a relative speed of 0.50c in the same direction, what speed must Beta have with respect to the earth?

Solution

According to the Galilean transformation, Beta would need a speed relative to the earth of $0.90c + 0.50c = 1.40c$, which we know is impossible. According to Eq. (1.49), however, with $V'_x = 0.50c$ and $v = 0.90c$, the required speed is only

$$V_x = \frac{V'_x + v}{1 + \frac{vV'_x}{c^2}} = \frac{0.50c + 0.90c}{1 + \frac{(0.90c)(0.50c)}{c^2}} = 0.97c$$

which is less than c. It is necessary to go less than 10 percent faster than a spacecraft traveling at 0.90c in order to pass it at a relative speed of 0.50c.

Simultaneity

The relative character of time as well as space has many implications. Notably, events that seem to take place simultaneously to one observer may not be simultaneous to another observer in relative motion, and vice versa.

Let us examine two events—the setting off of a pair of flares, say—that occur at the same time t_0 to somebody on the earth but at the different locations x_1 and x_2. What does the pilot of a spacecraft in flight see? To her, the flare at x_1 and t_0 appears at the time

$$t_1' = \frac{t_0 - vx_1/c^2}{\sqrt{1 - v^2/c^2}}$$

according to Eq. (1.44), while the flare at x_2 and t_0 appears at the time

$$t_2' = \frac{t_0 - vx_2/c^2}{\sqrt{1 - v^2/c^2}}$$

Hence two events that occur simultaneously to one observer are separated by a time interval of

$$t_2' - t_1' = \frac{v(x_1 - x_2)/c^2}{\sqrt{1 - v^2/c^2}}$$

to an observer moving at the speed v relative to the other observer. Who is right? The question is, of course, meaningless: both observers are "right" since each simply measures what he or she sees.

Because simultaneity is a relative concept and not an absolute one, physical theories that require simultaneity in events at different locations cannot be valid. For instance, saying that total energy is conserved in an isolated system does not rule out a process in which an amount of energy ΔE vanishes at one place while an equal amount of energy ΔE comes into being somewhere else with no actual transport of energy from one place to the other. Because simultaneity is relative, some observers of the process will find energy not being conserved. To rescue conservation of energy in the light of special relativity, then, we have to say that, when energy disappears somewhere and appears elsewhere, it has actually flowed from the first location to the second. Thus energy is conserved *locally* everywhere, not merely when an isolated system is considered—a much stronger statement of this principle.

Appendix II to Chapter 1

Spacetime

As we have seen, the concepts of space and time are inextricably mixed in nature. A length that one observer can measure with only a meter stick may have to be measured with both a meter stick and a clock by another observer.

A convenient and elegant way to express the results of special relativity is to regard events as occurring in a four-dimensional **spacetime** in which the usual three coordinates x, y, z refer to space and a fourth coordinate ict refers to time, where $i = \sqrt{-1}$. Although we cannot visualize spacetime, it is no harder to deal with mathematically than three-dimensional space.

The reason that ict is chosen as the time coordinate instead of just t is that the quantity

$$s^2 = x^2 + y^2 + z^2 - (ct)^2 \tag{1.52}$$

is **invariant** under a Lorentz transformation. That is, if an event occurs at x, y, z, t in an inertial frame S and at x', y', z', t' in another inertial frame S', then

$$s^2 = x^2 + y^2 + z^2 - (ct)^2 = x'^2 + y'^2 + z'^2 - (ct')^2$$

Because s^2 is invariant, we can think of a Lorentz transformation merely as a rotation in spacetime of the coordinate axes x, y, z, ict (Fig. 1.24).

The four coordinates x, y, z, ict define a vector in spacetime, and this *four-vector* remains fixed in spacetime regardless of any rotation of the coordinate system—that is, regardless of any shift in point of view from one inertial frame S to another S'.

Another four-vector whose magnitude remains constant under Lorentz transformations has the components $p_x, p_y, p_z, iE/c$. Here p_x, p_y, p_z are the usual components of the linear momentum of a body whose total energy is E. Hence the value of

$$p_x^2 + p_y^2 + p_z^2 - \frac{E^2}{c}$$

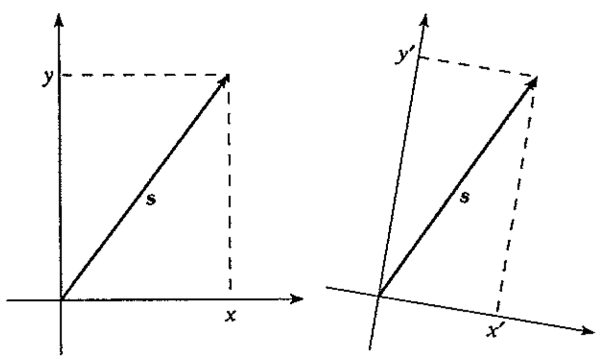


Figure 1.24 Rotating a two-dimensional coordinate system does not change the quantity $s^2 = x^2 + y^2 = x'^2 + y'^2$, where s is the length of the vector **s**. This result can be generalized to the four-dimensional spacetime coordinate system x, y, z, ict.

is the same in all inertial frames even though p_x, p_y, p_z and E separately may be different. This invariance was noted earlier in connection with Eq. (1.24); we note that $p^2 = p_x^2 + p_y^2 + p_z^2$.

A more mathematically elaborate formulation brings together the electric and magnetic fields **E** and **B** into an invariant quantity called a tensor. This approach to incorporating special relativity into physics has led both to a deeper understanding of natural laws and to the discovery of new phenomena and relationships.

Spacetime Intervals

The statements made at the end of Sec. 1.2 (P. 10) are easy to confirm using the idea of spacetime. Figure 1.25 shows two events plotted on the axes x and ct. Event 1 occurs at $x = 0$, $t = 0$ and event 2 occurs at $x = \Delta x$, $t = \Delta t$. The *spacetime interval* Δs between them is defined by

Spacetime interval between events

$$(\Delta s)^2 = (c\Delta t)^2 - (\Delta x)^2 \qquad (1.53)$$

The virtue of this definition is that $(\Delta s)^2$, like the s^2 of Eq. 1.52, is invariant under Lorentz transformations. If Δx and Δt are the differences in space and time between two events measured in the S frame and $\Delta x'$ and $\Delta t'$ are the same quantities measured in the S' frame,

$$(\Delta s)^2 = (c\Delta t)^2 - (\Delta x)^2 = (c\Delta t')^2 - (\Delta x')^2$$

Therefore whatever conclusions we arrive at in the S frame in which event 1 is at the origin hold equally well in any other frame in relative motion at constant velocity.

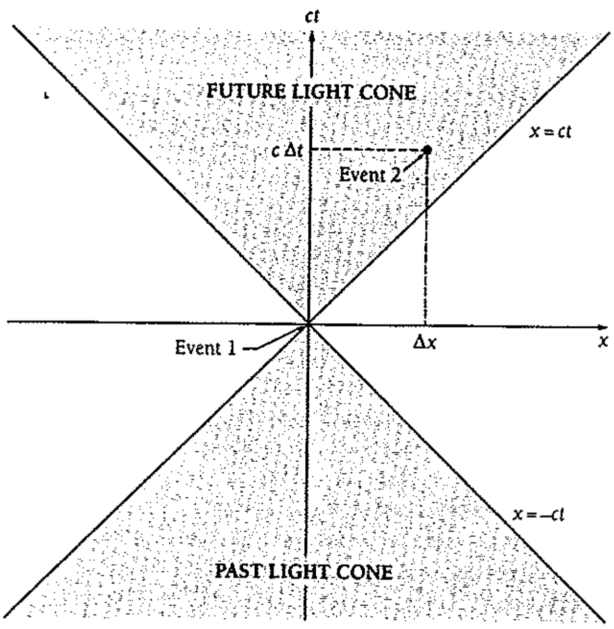


Figure 1.25 The past and future light cones in spacetime of event 1.

Now let us look into the possible relationships between events 1 and 2. Event 2 can be related causally in some way to event 1 provided that a signal traveling slower than the speed of light can connect these events, that is, provided that

$$c\Delta t > |\Delta x|$$

or

Timelike interval $$(\Delta s)^2 > 0 \tag{1.53}$$

An interval in which $(\Delta s)^2 > 0$ is said to be *timelike*. Every timelike interval that connects event 1 with another event lies within the *light cones* bounded by $x = \pm ct$ in Fig. 1.25. All events that could have affected event 1 lie in the past light cone; all events that event 1 is able to affect lie in the future light cone. (Events connected by timelike intervals need not *necessarily* be related, of course, but it is *possible* for them to be related.)

Conversely, the criterion for there being no causal relationship between events 1 and 2 is that

$$c\Delta t < |\Delta x|$$

or

Spacelike interval $$(\Delta s)^2 < 0 \tag{1.54}$$

An interval in which $(\Delta s)^2 < 0$ is said to be *spacelike*. Every event that is connected with event 1 by a spacelike interval lies outside the light cones of event 1 and neither has interacted with event 1 in the past nor is capable of interacting with it in the future; the two events must be entirely unrelated.

When events 1 and 2 can be connected with a light signal only,

$$c\Delta t = |\Delta x|$$

or

Lightlike interval $$\Delta s = 0 \tag{1.55}$$

An interval in which $\Delta s = 0$ is said to be *lightlike*. Events that can be connected with event 1 by lightlike intervals lie on the boundaries of the light cones.

These conclusions hold in terms of the light cones of event 2 because $(\Delta s)^2$ is invariant; for example, if event 2 is inside the past light cone of event 1, event 1 is inside the future light cone of event 2. In general, events that lie in the future of an event as seen in one frame of reference S lie in its future in every other frame S', and events that lie in the past of an event in S lie in its past in every other frame S'. Thus "future" and "past" have invariant meanings. However, "simultaneity" is an ambiguous concept, because all events that lie outside the past and future light cones of event 1 (that is, all events connected by spacelike intervals with event 1) can appear to occur simultaneously with event 1 in some particular frame of reference.

The path of a particle in spacetime is called its *world line* (Fig. 1.26). The world line of a particle must lie within its light cones.

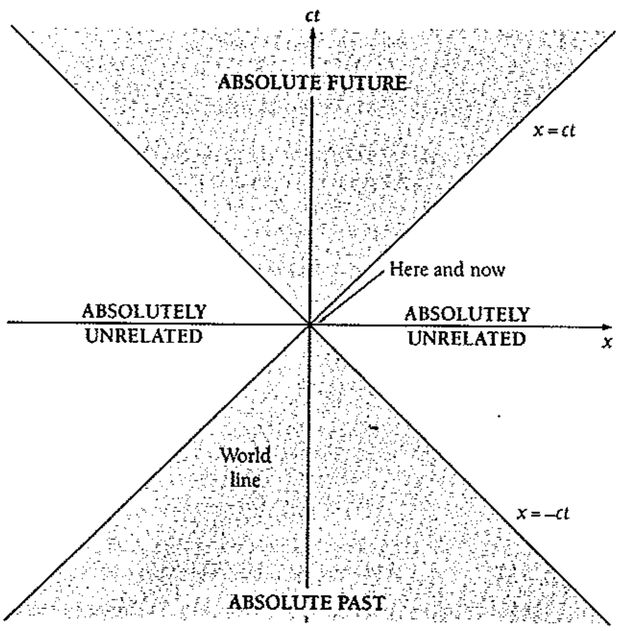


Figure 1.26 The world line of a particle in spacetime.

EXERCISES

But be ye doers of the word, and not hearers only, deceiving your own selves. —James I:22

1.1 Special Relativity

1. If the speed of light were smaller than it is, would relativistic phenomena be more or less conspicuous than they are now?
2. It is possible for the electron beam in a television picture tube to move across the screen at a speed faster than the speed of light. Why does this not contradict special relativity?

1.2 Time Dilation

3. An athlete has learned enough physics to know that if he measures from the earth a time interval on a moving spacecraft, what he finds will be greater than what somebody on the spacecraft would measure. He therefore proposes to set a world record for the 100-m dash by having his time taken by an observer on a moving spacecraft. Is this a good idea?
4. An observer on a spacecraft moving at $0.700c$ relative to the earth finds that a car takes 40.0 min to make a trip. How long does the trip take to the driver of the car?
5. Two observers, A on earth and B in a spacecraft whose speed is 2.00×10^8 m/s, both set their watches to the same time when the ship is abreast of the earth. (*a*) How much time must elapse by A's reckoning before the watches differ by 1.00 s? (*b*) To A, B's watch seems to run slow. To B, does A's watch seem to run fast, run slow, or keep the same time as his own watch?
6. An airplane is flying at 300 m/s (672 mi/h). How much time must elapse before a clock in the airplane and one on the ground differ by 1.00 s?
7. How fast must a spacecraft travel relative to the earth for each day on the spacecraft to correspond to 2 d on the earth?
8. The Apollo 11 spacecraft that landed on the moon in 1969 traveled there at a speed relative to the earth of 1.08×10^4 m/s. To an observer on the earth, how much longer than his own day was a day on the spacecraft?
9. A certain particle has a lifetime of 1.00×10^{-7} s when measured at rest. How far does it go before decaying if its speed is $0.99c$ when it is created?

1.3 Doppler Effect

10. A spacecraft receding from the earth at $0.97c$ transmits data at the rate of 1.00×10^4 pulses/s. At what rate are they received?
11. A galaxy in the constellation Ursa Major is receding from the earth at 15,000 km/s. If one of the characteristic wavelengths of the light the galaxy emits is 550 nm, what is the corresponding wavelength measured by astronomers on the earth?
12. The frequencies of the spectral lines in light from a distant galaxy are found to be two-thirds as great as those of the same lines in light from nearby stars. Find the recession speed of the distant galaxy.

13. A spacecraft receding from the earth emits radio waves at a constant frequency of 10^9 Hz. If the receiver on earth can measure frequencies to the nearest hertz, at what spacecraft speed can the difference between the relativistic and classical doppler effects be detected? For the classical effect, assume the earth is stationary.

14. A car moving at 150 km/h (93 mi/h) is approaching a stationary police car whose radar speed detector operates at a frequency of 15 GHz. What frequency change is found by the speed detector?

15. If the angle between the direction of motion of a light source of frequency ν_0 and the direction from it to an observer is θ, the frequency ν the observer finds is given by

$$\nu = \nu_0 \frac{\sqrt{1 - v^2/c^2}}{1 - (v/c) \cos \theta}$$

where v is the relative speed of the source. Show that this formula includes Eqs. (1.5) to (1.7) as special cases.

16. (*a*) Show that when $v \ll c$, the formulas for the doppler effect both in light and in sound for an observer approaching a source, and vice versa, all reduce to $\nu \approx \nu_0(1 + v/c)$, so that $\Delta\nu/\nu \approx v/c$. [*Hint:* For $x \ll 1$, $1/(1 + x) \approx 1 - x$.] (*b*) What do the formulas for an observer receding from a source, and vice versa, reduce to when $v \ll c$?

1.4 Length Contraction

17. An astronaut, whose height on the earth is exactly 6 ft is lying parallel to the axis of a spacecraft moving at 0.90*c* relative to the earth. What is his height as measured by an observer in the same spacecraft? By an observer on the earth?

18. An astronaut is standing in a spacecraft parallel to its direction of motion. An observer on the earth finds that the spacecraft speed is 0.60*c* and the astronaut is 1.3 m tall. What is the astronaut's height as measured in the spacecraft?

19. How much time does a meter stick moving at 0.100*c* relative to an observer take to pass the observer? The meter stick is parallel to its direction of motion.

20. A meter stick moving with respect to an observer appears only 500 mm long to her. What is its relative speed? How long does it take to pass her? The meter stick is parallel to its direction of motion.

21. A spacecraft antenna is at an angle of 10° relative to the axis of the spacecraft. If the spacecraft moves away from the earth at a speed of 0.70*c*, what is the angle of the antenna as seen from the earth?

1.5 Twin Paradox

22. Twin *A* makes a round trip at 0.6*c* to a star 12 light-years away, while twin *B* stays on the earth. Each twin sends the other a signal once a year by his own reckoning. (*a*) How many signals does *A* send during the trip? How many does *B* send? (*b*) How many signals does *A* receive? How many does *B* receive?

23. A woman leaves the earth in a spacecraft that makes a round trip to the nearest star, 4 light-years distant, at a speed of 0.9*c*. How much younger is she upon her return than her twin sister who remained behind?

1.7 Relativistic Momentum

24. (*a*) An electron's speed is doubled from 0.2*c* to 0.4*c*. By what ratio does its momentum increase? (*b*) What happens to the momentum ratio when the electron's speed is doubled again from 0.4*c* to 0.8*c*?

25. All definitions are arbitrary, but some are more useful than others. What is the objection to defining linear momentum as $\mathbf{p} = m\mathbf{v}$ instead of the more complicated $\mathbf{p} = \gamma m\mathbf{v}$?

26. Verify that

$$\frac{1}{\sqrt{1 - v^2/c^2}} = 1 + \frac{p^2}{m^2c^2}$$

1.8 Mass and Energy

27. Dynamite liberates about 5.4×10^6 J/kg when it explodes. What fraction of its total energy content is this?

28. A certain quantity of ice at 0°C melts into water at 0°C and in so doing gains 1.00 kg of mass. What was its initial mass?

29. At what speed does the kinetic energy of a particle equal its rest energy?

30. How many joules of energy per kilogram of rest mass are needed to bring a spacecraft from rest to a speed of 0.90*c*?

31. An electron has a kinetic energy of 0.100 MeV. Find its speed according to classical and relativistic mechanics.

32. Verify that, for $E \gg E_0$,

$$\frac{v}{c} \approx 1 - \frac{1}{2}\left(\frac{E_0}{E}\right)^2$$

33. A particle has a kinetic energy 20 times its rest energy. Find the speed of the particle in terms of *c*.

34. (*a*) The speed of a proton is increased from 0.20*c* to 0.40*c*. By what factor does its kinetic energy increase? (*b*) The proton speed is again doubled, this time to 0.80*c*. By what factor does its kinetic energy increase now?

35. How much work (in MeV) must be done to increase the speed of an electron from 1.2×10^8 m/s to 2.4×10^8 m/s?

36. (*a*) Derive a formula for the minimum kinetic energy needed by a particle of rest mass *m* to emit Cerenkov radiation in a medium of index of refraction *n*. [*Hint:* Start from Eqs. (1.21) and (1.23).] (*b*) Use this formula to find KE_{min} for an electron in a medium of $n = 1.5$.

37. Prove that $\frac{1}{2}\gamma mv^2$, does *not* equal the kinetic energy of a particle moving at relativistic speeds.

38. A moving electron collides with a stationary electron and an electron-positron pair comes into being as a result (a positron is a positively charged electron). When all four particles have the same velocity after the collision, the kinetic energy required for this process is a minimum. Use a relativistic calculation to show that $KE_{min} = 6mc^2$, where *m* is the rest mass of the electron.

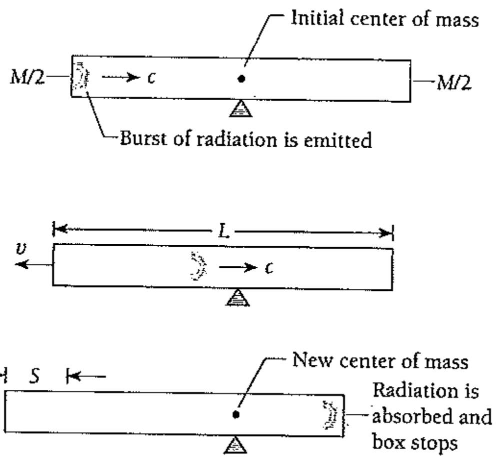


Figure 1.27 The box has moved the distance S to the left when it stops.

39. An alternative derivation of the mass-energy formula $E_0 = mc^2$, also given by Einstein, is based on the principle that the location of the center of mass (CM) of an isolated system cannot be changed by any process that occurs inside the system. Figure 1.27 shows a rigid box of length L that rests on a frictionless surface; the mass M of the box is equally divided between its two ends. A burst of electromagnetic radiation of energy E_0 is emitted by one end of the box. According to classical physics, the radiation has the momentum $p = E_0/c$, and when it is emitted, the box recoils with the speed $v \approx E_0/Mc$ so that the total momentum of the system remains zero. After a time $t \approx L/c$ the radiation reaches the other end of the box and is absorbed there, which brings the box to a stop after having moved the distance S. If the CM of the box is to remain in its original place, the radiation must have transferred mass from one end to the other. Show that this amount of mass is $m = E_0/c^2$.

1.9 Energy and Momentum

40. Find the SI equivalents of the mass unit MeV/c^2 and the momentum unit MeV/c.

41. In its own frame of reference, a proton takes 5 min to cross the Milky Way galaxy, which is about 10^5 light-years in diameter. (*a*) What is the approximate energy of the proton in electronvolts? (*b*) About how long would the proton take to cross the galaxy as measured by an observer in the galaxy's reference frame?

42. What is the energy of a photon whose momentum is the same as that of a proton whose kinetic energy is 10.0 MeV?

43. Find the momentum (in MeV/c) of an electron whose speed is $0.600c$.

44. Find the total energy and kinetic energy (in GeV) and the momentum (in GeV/c) of a proton whose speed is $0.900c$. The mass of the proton is 0.938 GeV/c^2.

45. Find the momentum of an electron whose kinetic energy equals its rest energy of 511 keV.

46. Verify that $v/c = pc/E$.

47. Find the speed and momentum (in GeV/c) of a proton whose total energy is 3.500 GeV.

48. Find the total energy of a neutron (m = 0.940 GeV/c^2) whose momentum is 1.200 GeV/c.

49. A particle has a kinetic energy of 62 MeV and a momentum of 335 MeV/c. Find its mass (in MeV/c^2) and speed (as a fraction of c).

50. (*a*) Find the mass (in GeV/c^2) of a particle whose total energy is 4.00 GeV and whose momentum is 1.45 GeV/c. (*b*) Find the total energy of this particle in a reference frame in which its momentum is 2.00 GeV/c.

Appendix I: The Lorentz Transformation

51. An observer detects two explosions, one that occurs near her at a certain time and another that occurs 2.00 ms later 100 km away. Another observer finds that the two explosions occur at the same place. What time interval separates the explosions to the second observer?

52. An observer detects two explosions that occur at the same time, one near her and the other 100 km away. Another observer finds that the two explosions occur 160 km apart. What time interval separates the explosions to the second observer?

53. A spacecraft moving in the $+x$ direction receives a light signal from a source in the xy plane. In the reference frame of the fixed stars, the speed of the spacecraft is v and the signal arrives at an angle θ to the axis of the spacecraft. (*a*) With the help of the Lorentz transformation find the angle θ' at which the signal arrives in the reference frame of the spacecraft. (*b*) What would you conclude from this result about the view of the stars from a porthole on the side of the spacecraft?

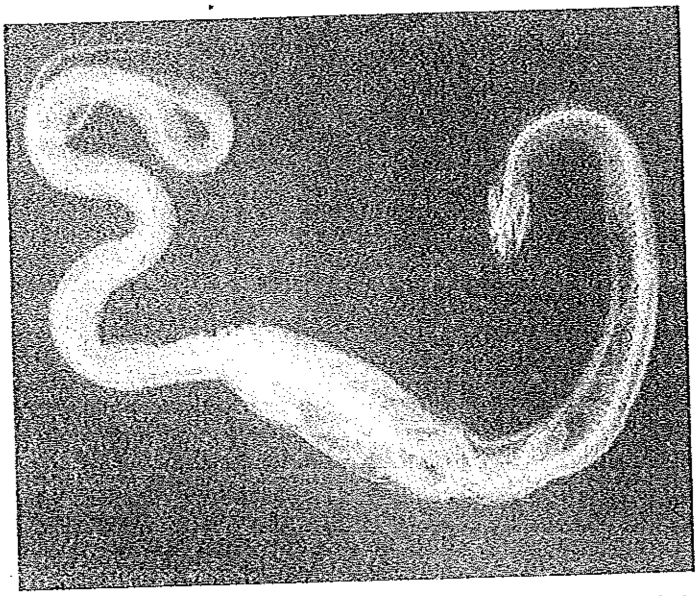
54. A body moving at $0.500c$ with respect to an observer disintegrates into two fragments that move in opposite directions relative to their center of mass along the same line of motion as the original body. One fragment has a velocity of $0.600c$ in the backward direction relative to the center of mass and the other has a velocity of $0.500c$ in the forward direction. What velocities will the observer find?

55. A man on the moon sees two spacecraft, A and B, coming toward him from opposite directions at the respective speeds of $0.800c$ and $0.900c$. (*a*) What does a man on A measure for the speed with which he is approaching the moon? For the speed with which he is approaching B? (*b*) What does a man on B measure for the speed with which he is approaching the moon? For the speed with which he is approaching A?

56. An electron whose speed relative to an observer in a laboratory is $0.800c$ is also being studied by an observer moving in the same direction as the electron at a speed of $0.500c$ relative to the laboratory. What is the kinetic energy (in MeV) of the electron to each observer?

CHAPTER 2

Particle Properties of Waves



The penetrating ability of x-rays enabled them to reveal the frog which this snake had swallowed. The snake's jaws are very loosely joined and so can open widely.

2.1 ELECTROMAGNETIC WAVES
Coupled electric and magnetic oscillations that move with the speed of light and exhibit typical wave behavior

2.2 BLACKBODY RADIATION
Only the quantum theory of light can explain its origin

2.3 PHOTOELECTRIC EFFECT
The energies of electrons liberated by light depend on the frequency of the light

2.4 WHAT IS LIGHT?
Both wave and particle

2.5 X-RAYS
They consist of high-energy photons

2.6 X-RAY DIFFRACTION
How x-ray wavelengths can be determined

2.7 COMPTON EFFECT
Further confirmation of the photon model

2.8 PAIR PRODUCTION
Energy into matter

2.9 PHOTONS AND GRAVITY
Although they lack rest mass, photons behave though they have gravitational mass

In our everyday experience there is nothing mysterious or ambiguous about the concepts of **particle** and **wave.** A stone dropped into a lake and the ripples that spread out from its point of impact apparently have in common only the ability to carry energy and momentum from one place to another. Classical physics, which mirrors the "physical reality" of our sense impressions, treats particles and waves as separate components of that reality. The mechanics of particles and the optics of waves are traditionally independent disciplines, each with its own chain of experiments and principles based on their results.

The physical reality we perceive has its roots in the microscopic world of atoms and molecules, electrons and nuclei, but in this world there are neither particles nor waves in our sense of these terms. We regard electrons as particles because they possess charge and mass and behave according to the laws of particle mechanics in such familiar devices as television picture tubes. We shall see, however, that it is just as correct to interpret a moving electron as a wave manifestation as it is to interpret it as a particle manifestation. We regard electromagnetic waves as waves because under suitable circumstances they exhibit diffraction, interference, and polarization. Similarly, we shall see that under other circumstances electromagnetic waves behave as though they consist of streams of particles. Together with special relativity, the wave-particle duality is central to an understanding of modern physics, and in this book there are few arguments that do not draw upon either or both of these fundamental ideas.

2.1 ELECTROMAGNETIC WAVES

Coupled electric and magnetic oscillations that move with the speed of light and exhibit typical wave behavior

In 1864 the British physicist James Clerk Maxwell made the remarkable suggestion that accelerated electric charges generate linked electric and magnetic disturbances that can travel indefinitely through space. If the charges oscillate periodically, the disturbances are waves whose electric and magnetic components are perpendicular to each other and to the direction of propagation, as in Fig. 2.1.

From the earlier work of Faraday, Maxwell knew that a changing magnetic field can induce a current in a wire loop. Thus a changing magnetic field is equivalent in its effects to an electric field. Maxwell proposed the converse: a changing electric field has a magnetic field associated with it. The electric fields produced by electromagnetic induction are easy to demonstrate because metals offer little resistance to the flow of charge. Even a weak field can lead to a measurable current in a metal. Weak magnetic fields are much harder to detect, however, and Maxwell's hypothesis was based on a symmetry argument rather than on experimental findings.

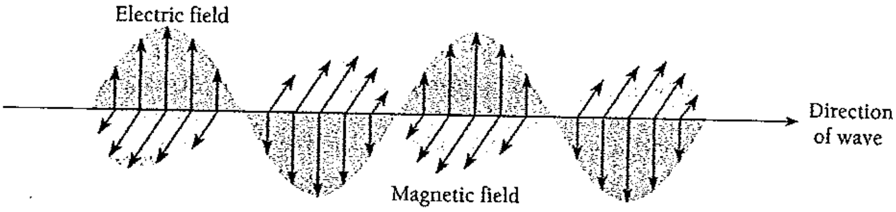


Figure 2.1 The electric and magnetic fields in an electromagnetic wave vary together. The fields are perpendicular to each other and to the direction of propagation of the wave.

James Clerk Maxwell (1831–1879) was born in Scotland shortly before Michael Faraday discovered electromagnetic induction. At nineteen he entered Cambridge University to study physics and mathematics. While still a student, he investigated the physics of color vision and later used his ideas to make the first color photograph. Maxwell became known to the scientific world at twenty-four when he showed that the rings of Saturn could not be solid or liquid but must consist of separate small bodies. At about this time Maxwell became interested in electricity and magnetism and grew convinced that the wealth of phenomena Faraday and others had discovered were not isolated effects but had an underlying unity of some kind. Maxwell's initial step in establishing that unity came in 1856 with the paper "On Faraday's Lines of Force," in which he developed a mathematical description of electric and magnetic fields.

Maxwell left Cambridge in 1856 to teach at a college in Scotland and later at King's College in London. In this period he expanded his ideas on electricity and magnetism to create a single comprehensive theory of electromagnetism. The fundamental equations he arrived at remain the foundations of the subject today. From these equations Maxwell predicted that electromagnetic waves should exist that travel with the speed of light, described the properties the waves should have, and surmised that light consisted of electromagnetic waves. Sadly, he did not live to see his work confirmed in the experiments of the German physicist Heinrich Hertz.

Maxwell's contributions to kinetic theory and statistical mechanics were on the same profound level as his contributions to electromagnetic theory. His calculations showed that the viscosity of a gas ought to be independent of its pressure, a surprising result that Maxwell, with the help of his wife, confirmed in the laboratory. They also found that the viscosity was proportional to the absolute temperature of the gas. Maxwell's explanation for this proportionality gave him a way to estimate the size and mass of molecules, which until then could only be guessed at. Maxwell shares with Boltzmann credit for the equation that gives the distribution of molecular energies in a gas.

In 1865 Maxwell returned to his family's home in Scotland. There he continued his research and also composed a treatise on electromagnetism that was to be the standard text on the subject for many decades. It was still in print a century later. In 1871 Maxwell went back to Cambridge to establish and direct the Cavendish Laboratory, named in honor of the pioneering physicist Henry Cavendish. Maxwell died of cancer at the age of forty-eight in 1879, the year in which Albert Einstein was born. Maxwell had been the greatest theoretical physicist of the nineteenth century; Einstein was to be the greatest theoretical physicist of the twentieth century. (By a similar coincidence, Newton was born in the year of Galileo's death.)

If Maxwell was right, electromagnetic (em) waves must occur in which constantly varying electric and magnetic fields are coupled together by both electromagnetic induction and the converse mechanism he proposed. Maxwell was able to show that the speed c of electromagnetic waves in free space is given by

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} = 2.998 \times 10^8 \text{ m/s}$$

where ϵ_0 is the electric permittivity of free space and μ_0 is its magnetic permeability. This is the same as the speed of light waves. The correspondence was too great to be accidental, and Maxwell concluded that light consists of electromagnetic waves.

During Maxwell's lifetime the notion of em waves remained without direct experimental support. Finally, in 1888, the German physicist Heinrich Hertz showed that em waves indeed exist and behave exactly as Maxwell had predicted. Hertz generated the waves by applying an alternating current to an air gap between two metal balls. The width of the gap was such that a spark occurred each time the current reached a peak. A wire loop with a small gap was the detector; em waves set up oscillations in the loop that produced sparks in the gap. Hertz determined the wavelength and speed of the waves he generated, showed that they have both electric and magnetic components, and found that they could be reflected, refracted, and diffracted.

Light is not the only example of an em wave. Although all such waves have the same fundamental nature, many features of their interaction with matter depend upon

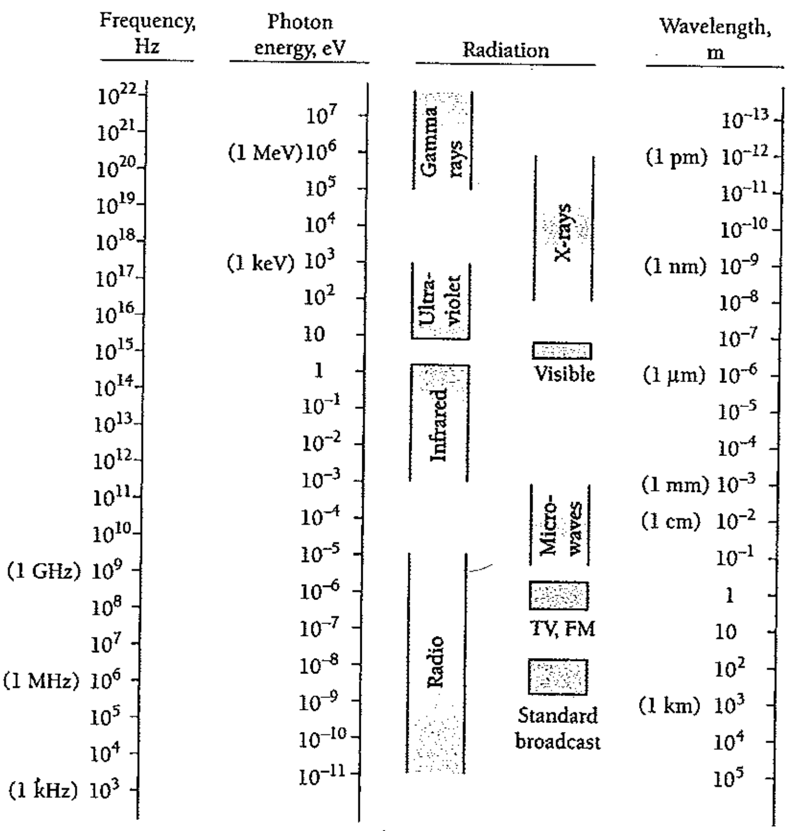


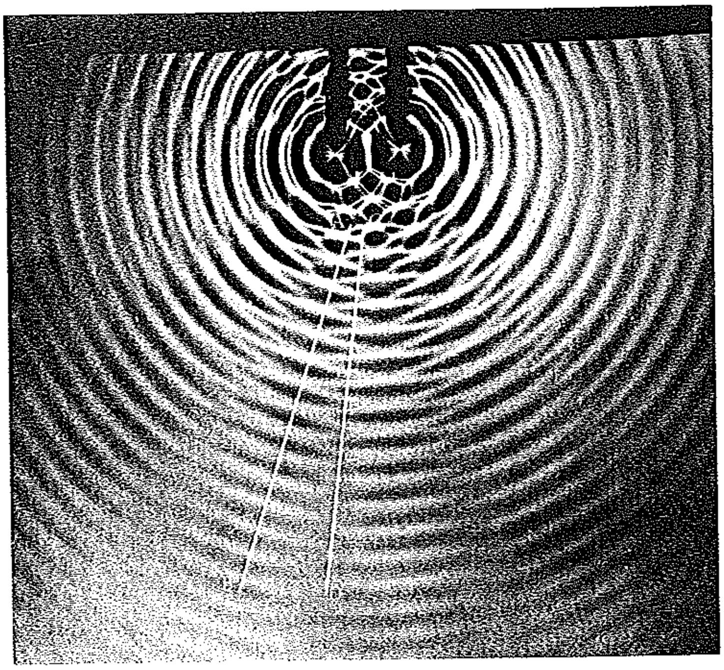
Figure 2.2 The spectrum of electromagnetic radiation.

their frequencies. Light waves, which are em waves the eye responds to, span only a brief frequency interval, from about 4.3×10^{14} Hz for red light to about 7.5×10^{14} Hz for violet light. Figure 2.2 shows the em wave spectrum from the low frequencies used in radio communication to the high frequencies found in x-rays and gamma rays.

A characteristic property of all waves is that they obey the **principle of superposition:**

> When two or more waves of the same nature travel past a point at the same time, the instantaneous amplitude there is the sum of the instantaneous amplitudes of the individual waves.

Instantaneous amplitude refers to the value at a certain place and time of the quantity whose variations constitute the wave. ("Amplitude" without qualification refers to the maximum value of the wave variable.) Thus the instantaneous amplitude of a wave in a stretched string is the displacement of the string from its normal position; that of a water wave is the height of the water surface relative to its normal level; that of a sound wave is the change in pressure relative to the normal pressure. Since the electric and magnetic fields in a light wave are related by $E = cB$, its instantaneous amplitude can be taken as either E or B. Usually E is used, since it is the electric fields of light waves whose interactions with matter give rise to nearly all common optical effects.



The interference of water waves. Constructive interference occurs along the line *AB* and destructive interference occurs along the line *CD*.

When two or more trains of light waves meet in a region, they **interfere** to produce a new wave there whose instantaneous amplitude is the sum of those of the original waves. Constructive interference refers to the reinforcement of waves with the same phase to produce a greater amplitude, and destructive interference refers to the partial or complete cancellation of waves whose phases differ (Fig. 2.3). If the original waves have different frequencies, the result will be a mixture of constructive and destructive interference, as in Fig. 3.4.

The interference of light waves was first demonstrated in 1801 by Thomas Young, who used a pair of slits illuminated by monochromatic light from a single source (Fig. 2.4). From each slit secondary waves spread out as though originating at the slit; this is an example of **diffraction**, which, like interference, is a characteristic wave phenomenon. Owing to interference, the screen is not evenly lit but shows a pattern of alternate bright and dark lines. At those places on the screen where the path lengths from the two slits differ by an odd number of half wavelengths ($\lambda/2$, $3\lambda/2$, $5\lambda/2$, . . .), destructive interference occurs and a dark line is the result. At those places where the path lengths are

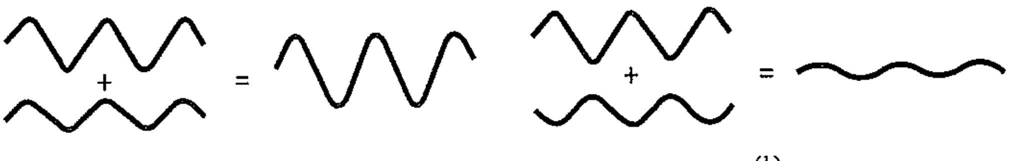


Figure 2.3 (*a*) In constructive interference, superposed waves in phase reinforce each other. (*b*) In destructive interference, waves out of phase partially or completely cancel each other.

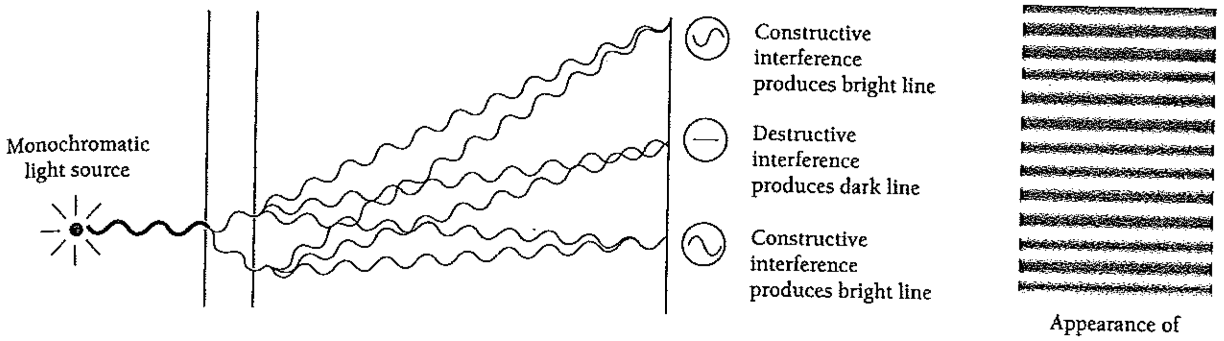


Figure 2.4 Origin of the interference pattern in Young's experiment. Constructive interference occurs where the difference in path lengths from the slits to the screen is θ, λ, 2λ, ... Destructive interference occurs where the path difference is $\lambda/2$, $3\lambda/2$, $5\lambda/2$, ...

equal or differ by a whole number of wavelengths (λ, 2λ, 3λ, . . .), constructive interference occurs and a bright line is the result. At intermediate places the interference is only partial, so the light intensity on the screen varies gradually between the bright and dark lines.

Interference and diffraction are found only in waves—the particles we are familiar with do not behave in those ways. If light consisted of a stream of classical particles, the entire screen would be dark. Thus Young's experiment is proof that light consists of waves. Maxwell's theory further tells us what kind of waves they are: electromagnetic. Until the end of the nineteenth century the nature of light seemed settled forever.

2.2 BLACKBODY RADIATION

Only the quantum theory of light can explain its origin

Following Hertz's experiments, the question of the fundamental nature of light seemed clear: light consisted of em waves that obeyed Maxwell's theory. This certainty lasted only a dozen years. The first sign that something was seriously amiss came from attempts to understand the origin of the radiation emitted by bodies of matter.

We are all familiar with the glow of a hot piece of metal, which gives off visible light whose color varies with the temperature of the metal, going from red to yellow to white as it becomes hotter and hotter. In fact, other frequencies to which our eyes do not respond are present as well. An object need not be so hot that it is luminous for it to be radiating em energy; *all* objects radiate such energy continuously whatever their temperatures, though which frequencies predominate depends on the temperature. At room temperature most of the radiation is in the infrared part of the spectrum and hence is invisible.

The ability of a body to radiate is closely related to its ability to absorb radiation. This is to be expected, since a body at a constant temperature is in thermal equilibrium with its surroundings and must absorb energy from them at the same rate as it emits energy. It is convenient to consider as an ideal body one that absorbs *all* radiation incident upon it, regardless of frequency. Such a body is called a **blackbody.**

The point of introducing the idealized blackbody in a discussion of thermal radiation is that we can now disregard the precise nature of whatever is radiating, since

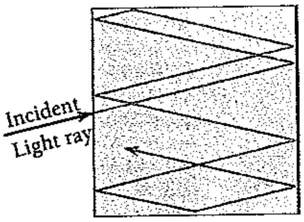


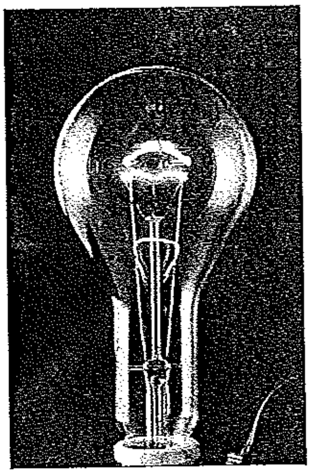
Figure 2.5 A hole in the wall of a hollow object is an excellent approximation of a blackbody.

all blackbodies behave identically. In the laboratory a blackbody can be approximated by a hollow object with a very small hole leading to its interior (Fig. 2.5). Any radiation striking the hole enters the cavity, where it is trapped by reflection back and forth until it is absorbed. The cavity walls are constantly emitting and absorbing radiation, and it is in the properties of this radiation **(blackbody radiation)** that we are interested.

Experimentally we can sample blackbody radiation simply by inspecting what emerges from the hole in the cavity. The results agree with everyday experience. A blackbody radiates more when it is hot than when it is cold, and the spectrum of a hot blackbody has its peak at a higher frequency than the peak in the spectrum of a cooler one. We recall the behavior of an iron bar as it is heated to progressively higher temperatures: at first it glows dull red, then bright orange-red, and eventually it becomes "white hot." The spectrum of blackbody radiation is shown in Fig. 2.6 for two temperatures.

The Ultraviolet Catastrophe

Why does the blackbody spectrum have the shape shown in Fig. 2.6? This problem was examined at the end of the nineteenth century by Lord Rayleigh and James Jeans. The details of their calculation are given in Chap. 9. They started by considering the radiation inside a cavity of absolute temperature T whose walls are perfect reflectors to be a series of standing em waves (Fig. 2.7). This is a three-dimensional generalization of standing waves in a stretched string. The condition



The color and brightness of an object heated until it glows, such as the filament of this light bulb, depends upon its temperature, which here is about 3000 K. An object that glows white is hotter than it is when it glows red, and it gives off more light as well.

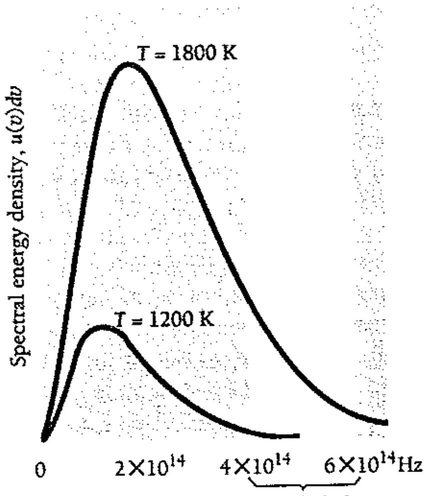


Figure 2.6 Blackbody spectra. The spectral distribution of energy in the radiation depends only on the temperature of the body. The higher the temperature, the greater the amount of radiation and the higher the frequency at which the maximum emission occurs. The dependence of the latter frequency on temperature follows a formula called Wien's displacement law, which is discussed in Sec. 9.6.

for standing waves in such a cavity is that the path length from wall to wall, whatever the direction, must be a whole number of half-wavelengths, so that a node occurs at each reflecting surface. The number of independent standing waves $G(\nu)d\nu$ in the frequency interval between ν and $d\nu$ per unit volume in the cavity turned out to be

Density of standing waves in cavity

$$G(\nu)d\nu = \frac{8\pi\nu^2 d\nu}{c^3} \tag{2.1}$$

This formula is independent of the shape of the cavity. As we would expect, the higher the frequency ν, the shorter the wavelength and the greater the number of possible standing waves.

The next step is to find the average energy per standing wave. According to the **theorem of equipartition of energy**, a mainstay of classical physics, the average energy per degree of freedom of an entity (such as a molecule of an ideal gas) that is a member of a system of such entities in thermal equilibrium at the temperature T is $\frac{1}{2}kT$. Here k is **Boltzmann's constant:**

Boltzmann's constant

$$k = 1.381 \times 10^{-23} \text{ J/K}$$

A degree of freedom is a mode of energy possession. Thus a monatomic ideal gas molecule has three degrees of freedom, corresponding to kinetic energy of motion in three independent directions, for an average total energy of $\frac{3}{2}kT$.

A one-dimensional harmonic oscillator has two degrees of freedom, one that corresponds to its kinetic energy and one that corresponds to its potential energy. Because each standing wave in a cavity originates in an oscillating electric charge in the cavity wall, two degrees of freedom are associated with the wave and it should have an average energy of $2(\frac{1}{2})kT$:

Classical average energy per standing wave

$$\bar{\epsilon} = kT \tag{2.2}$$

The total energy $u(\nu)\, d\nu$ per unit volume in the cavity in the frequency interval from ν to $\nu + d\nu$ is therefore

Rayleigh-Jeans formula

$$u(\nu)\, d\nu = \bar{\epsilon}G(\nu)\, d\nu = \frac{8\pi kT}{c^3}\nu^2\, d\nu \tag{2.3}$$

This radiation rate is proportional to this energy density for frequencies between ν and $\nu + d\nu$. Equation (2.3), the **Rayleigh-Jeans formula,** contains everything that classical physics can say about the spectrum of blackbody radiation.

Even a glance at Eq. (2.3) shows that it cannot possibly be correct. As the frequency ν increases toward the ultraviolet end of the spectrum, this formula predicts that the energy density should increase as ν^2. In the limit of infinitely high frequencies, $u(\nu)\, d\nu$ therefore should also go to infinity. In reality, of course, the energy density (and radiation rate) falls to 0 as $\nu \rightarrow \infty$ (Fig. 2.8). This discrepancy became known as the **ultraviolet catastrophe** of classical physics. Where did Rayleigh and Jeans go wrong?

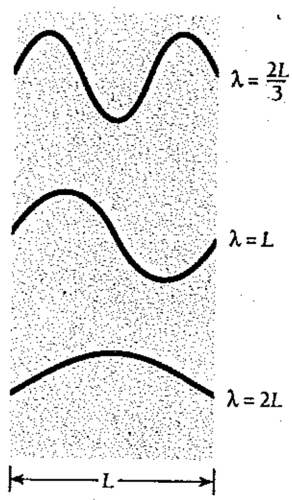


Figure 2.7 Em radiation in a cavity whose walls are perfect reflectors consists of standing waves that have nodes at the walls, which restricts their possible wavelengths. Shown are three possible wavelengths when the distance between opposite walls is L.

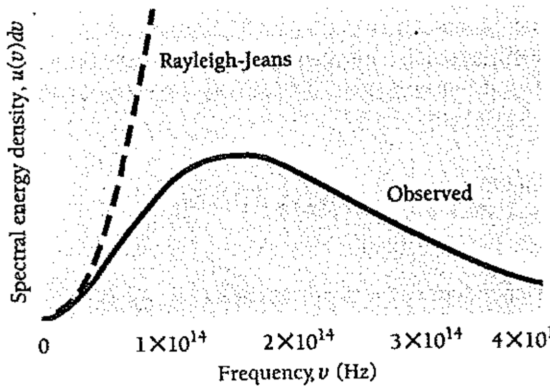


Figure 2.8 Comparison of the Rayleigh-Jeans formula for the spectrum of the radiation from a blackbody at 1500 K with the observed spectrum. The discrepancy is known as the ultraviolet catastrophe because it increases with increasing frequency. This failure of classical physics led Planck to the discovery that radiation is emitted in quanta whose energy is $h\nu$.

Planck Radiation Formula

In 1900 the German physicist Max Planck used "lucky guesswork" (as he later called it) to come up with a formula for the spectral energy density of blackbody radiation:

Planck radiation formula

$$u(\nu)\, d\nu = \frac{8\pi h}{c^3} \frac{\nu^3\, d\nu}{e^{h\nu/kT} - 1} \tag{2.4}$$

Here h is a constant whose value is

Planck's constant

$$h = 6.626 \times 10^{-34}\ \text{J} \cdot \text{s}$$

Max Planck (1858–1947) was born in Kiel and educated in Munich and Berlin. At the University of Berlin he studied under Kirchhoff and Helmholtz, as Hertz had done earlier. Planck realized that blackbody radiation was important because it was a fundamental effect independent of atomic structure, which was still a mystery in the late nineteenth century, and worked at understanding it for six years before finding the formula the radiation obeyed. He "strived from the day of its discovery to give it a real physical interpretation." The result was the discovery that radiation is emitted in energy steps of $h\nu$. Although this discovery, for which he received the Nobel Prize in 1918, is now considered to mark the start of modern physics, Planck himself remained skeptical for a long time of the physical reality of quanta. As he later wrote, "My vain attempts to somehow reconcile the elementary quantum with classical theory continued for many years and cost me great effort. . . . Now I know for certain that the quantum of action has a much more fundamental significance than I originally suspected."

Like many physicists, Planck was a competent musician (he sometimes played with Einstein) and in addition enjoyed mountain climbing. Although Planck remained in Germany during the Hitler era, he protested the Nazi treatment of Jewish scientists and lost his presidency of the Kaiser Wilhelm Institute as a result. In 1945 one of his sons was implicated in a plot to kill Hitler and was executed. After World War II the Institute was renamed after Planck and he was again its head until his death.

At high frequencies, $h\nu \gg kT$ and $e^{h\nu/kT} \rightarrow \infty$, which means that $u(\nu)\,d\nu \rightarrow 0$ as observed. No more ultraviolet catastrophe. At low frequencies, where the Rayleigh-Jeans formula is a good approximation to the data (see Fig. 2.8), $h\nu \ll kT$ and $h\nu/kT \ll 1$. In general,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

If x is small, $e^x \approx 1 + x$, and so for $h\nu/kT \ll 1$ we have

$$\frac{1}{e^{h\nu/kT} - 1} \approx \frac{1}{1 + \frac{h\nu}{kT} - 1} \approx \frac{kT}{h\nu} \qquad h\nu \ll kT$$

Thus at low frequencies Planck's formula becomes

$$u(\nu)\,d\nu \approx \frac{8\pi h}{c^3}\nu^3\left(\frac{kT}{h\nu}\right)d\nu \approx \frac{8\pi kT}{c^3}\nu^2\,d\nu$$

which is the Rayleigh-Jeans formula. Planck's formula is clearly at least on the right track; in fact, it has turned out to be completely correct.

Next Planck had the problem of justifying Eq. (2.4) in terms of physical principles. A new principle seemed needed to explain his formula, but what was it? After several weeks of "the most strenuous work of my life," Planck found the answer: The oscillators in the cavity walls could not have a continuous distribution of possible energies ϵ but must have only the specific energies

Oscillator energies $\qquad \epsilon_n = nh\nu \qquad n = 0, 1, 2, \ldots \qquad$ (2.5)

An oscillator emits radiation of frequency ν when it drops from one energy state to the next lower one, and it jumps to the next higher state when it absorbs radiation of frequency ν. Each discrete bundle of energy $h\nu$ is called a **quantum** (plural **quanta**) from the Latin for "how much."

With oscillator energies limited to $nh\nu$, the average energy per oscillator in the cavity walls—and so per standing wave—turned out to be not $\bar{\epsilon} = kT$ as for a continuous distribution of oscillator energies, but instead

Actual average energy per standing wave $\qquad \bar{\epsilon} = \frac{h\nu}{e^{h\nu/kT} - 1} \qquad$ (2.6)

This average energy leads to Eq. (2.4). Blackbody radiation is further discussed in Chap. 9.

Example 2.1

Assume that a certain 660-Hz tuning fork can be considered as a harmonic oscillator whose vibrational energy is 0.04 J. Compare the energy quanta of this tuning fork with those of an atomic oscillator that emits and absorbs orange light whose frequency is 5.00×10^{14} Hz.

Solution

(*a*) For the tuning fork,

$$h\nu_1 = (6.63 \times 10^{-34}\ \text{J} \cdot \text{s})\ (660\ \text{s}^{-1}) = 4.38 \times 10^{-31}\ \text{J}$$

The total energy of the vibrating tines of the fork is therefore about 10^{29} times the quantum energy $h\nu$. The quantization of energy in the tuning fork is obviously far too small to be observed, and we are justified in regarding the fork as obeying classical physics.

(*b*) For the atomic oscillator,

$$h\nu_2 = (6.63 \times 10^{-34}\ \text{J} \cdot \text{s})\ (5.00 \times 10^{14}\ \text{s}^{-1}) = 3.32 \times 10^{-19}\ \text{J}$$

In electronvolts, the usual energy unit in atomic physics,

$$h\nu_2 = \frac{3.32 \times 10^{-19}\ \text{J}}{1.60 \times 10^{-19}\ \text{J/eV}} = 2.08\ \text{eV}$$

This is a significant amount of energy on an atomic scale, and it is not surprising that classical physics fails to account for phenomena on this scale.

The concept that the oscillators in the cavity walls can interchange energy with standing waves in the cavity only in quanta of $h\nu$ is, from the point of view of classical physics, impossible to understand. Planck regarded his quantum hypothesis as an "act of desperation" and, along with other physicists of his time, was unsure of how seriously to regard it as an element of physical reality. For many years he held that, although the energy transfers between electric oscillators and em waves apparently are quantized, em waves themselves behave in an entirely classical way with a continuous range of possible energies.

2.3 PHOTOELECTRIC EFFECT

The energies of electrons liberated by light depend on the frequency of the light

During his experiments on em waves, Hertz noticed that sparks occurred more readily in the air gap of his transmitter when ultraviolet light was directed at one of the metal balls. He did not follow up this observation, but others did. They soon discovered that the cause was electrons emitted when the frequency of the light was sufficiently high. This phenomenon is known as the **photoelectric effect** and the emitted electrons are called **photoelectrons.** It is one of the ironies of history that the same work to demonstrate that light consists of em waves also gave the first hint that this was not the whole story.

Figure 2.9 shows how the photoelectric effect was studied. An evacuated tube contains two electrodes connected to a source of variable voltage, with the metal plate whose surface is irradiated as the anode. Some of the photoelectrons that emerge from this surface have enough energy to reach the cathode despite its negative polarity, and they constitute the measured current. The slower photoelectrons are repelled before they get to the cathode. When the voltage is increased to a certain value V_0, of the order of several volts, no more photoelectrons arrive, as indicated by the current dropping to zero. This extinction voltage corresponds to the maximum photoelectron kinetic energy.

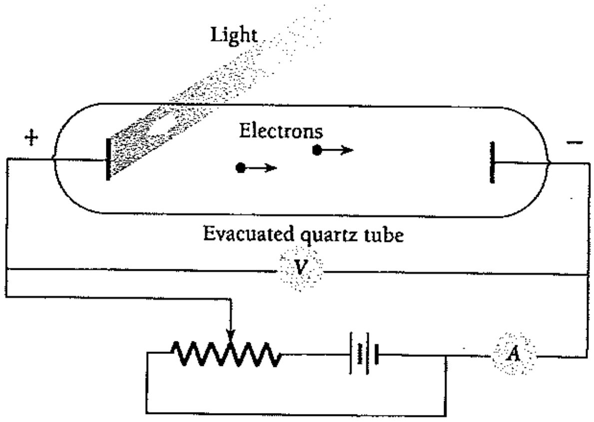


Figure 2.9 Experimental observation of the photoelectric effect.

The existence of the photoelectric effect is not surprising. After all, light waves carry energy, and some of the energy absorbed by the metal may somehow concentrate on individual electrons and reappear as their kinetic energy. The situation should be like water waves dislodging pebbles from a beach. But three experimental findings show that no such simple explanation is possible.

1 Within the limits of experimental accuracy (about 10^{-9} s), there is no time interval between the arrival of light at a metal surface and the emission of photoelectrons. However, because the energy in an em wave is supposed to be spread across the wavefronts, a period of time should elapse before an individual electron accumulates enough energy (several eV) to leave the metal. A detectable photoelectron current results when 10^{-6} W/m^2 of em energy is absorbed by a sodium surface. A layer of sodium 1 atom thick and 1 m^2 in area contains about 10^{19} atoms, so if the incident light is absorbed in the uppermost atomic layer, each atom receives energy at an average rate of 10^{-25} W. At this rate over a month would be needed for an atom to accumulate energy of the magnitude that photoelectrons from a sodium surface are observed to have.
2 A bright light yields more photoelectrons than a dim one of the same frequency, but the electron energies remain the same (Fig. 2.10). The em theory of light, on the contrary, predicts that the more intense the light, the greater the energies of the electrons.
3 The higher the frequency of the light, the more energy the photoelectrons have (Fig. 2.11). Blue light results in faster electrons than red light. At frequencies below a certain critical frequency ν_0, which is characteristic of each particular metal, no electrons are emitted. Above ν_0 the photoelectrons range in energy from 0 to a maximum value that increases linearly with increasing frequency (Fig. 2.12). This observation, also, cannot be explained by the em theory of light.

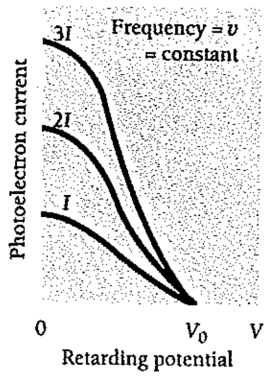


Figure 2.10 Photoelectron current is proportional to light intensity I for all retarding voltages. The stopping potential V_0, which corresponds to the maximum photoelectron energy, is the same for all intensities of light of the same frequency ν.

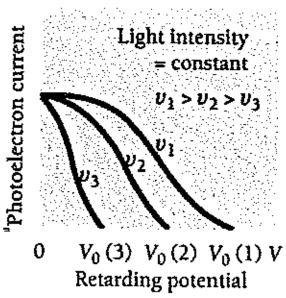


Figure 2.11 The stopping potential V_0, and hence the maximum photoelectron energy, depends on the frequency of the light. When the retarding potential is $V = 0$, the photoelectron current is the same for light of a given intensity regardless of its frequency.

Quantum Theory of Light

When Planck's derivation of his formula appeared, Einstein was one of the first—perhaps the first—to understand just how radical the postulate of energy quantization

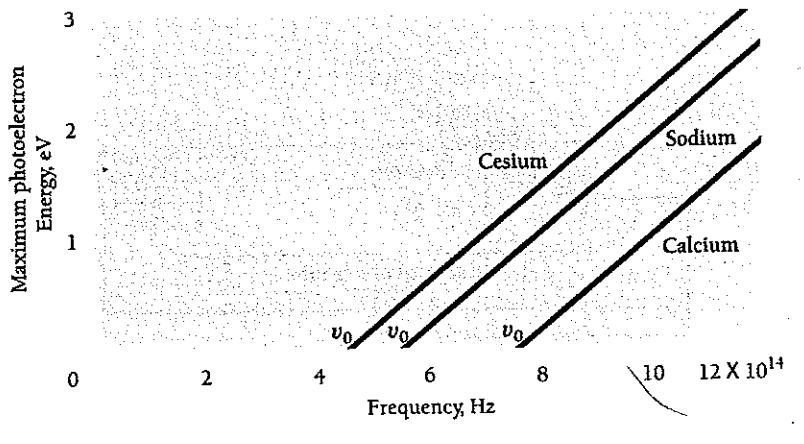


Figure 2.12 Maximum photoelectron kinetic energy KE_{max} versus frequency of incident light for three metal surfaces.

of oscillators was: "It was as if the ground was pulled from under one." A few years later, in 1905, Einstein realized that the photoelectric effect could be understood if the energy in light is not spread out over wavefronts but is concentrated in small packets, or **photons**. (The term photon was coined by the chemist Gilbert Lewis in 1926.) Each photon of light of frequency ν has the energy $h\nu$, the same as Planck's quantum energy. Planck had thought that, although energy from an electric oscillator apparently had to be given to em waves in separate quanta of $h\nu$ each, the waves themselves behaved exactly as in conventional wave theory. Einstein's break with classical physics was more drastic: Energy was not only given to em waves in separate quanta but was also carried by the waves in separate quanta.

The three experimental observations listed above follow directly from Einstein's hypothesis. (1) Because em wave energy is concentrated in photons and not spread out, there should be no delay in the emission of photoelectrons. (2) All photons of frequency ν have the same energy, so changing the intensity of a monochromatic light beam will change the number of photoelectrons but not their energies. (3) The higher the frequency ν, the greater the photon energy $h\nu$ and so the more energy the photoelectrons have.

What is the meaning of the critical frequency ν_0 below which no photoelectrons are emitted? There must be a minimum energy ϕ for an electron to escape from a particular metal surface or else electrons would pour out all the time. This energy is called the **work function** of the metal, and is related to ν_0 by the formula

Work function $$\phi = h\nu_0 \tag{2.7}$$

The greater the work function of a metal, the more energy is needed for an electron to leave its surface, and the higher the critical frequency for photoelectric emission to occur.

Some examples of photoelectric work functions are given in Table 2.1. To pull an electron from a metal surface generally takes about half as much energy as that needed

Table 2.1 Photoelectric Work Functions

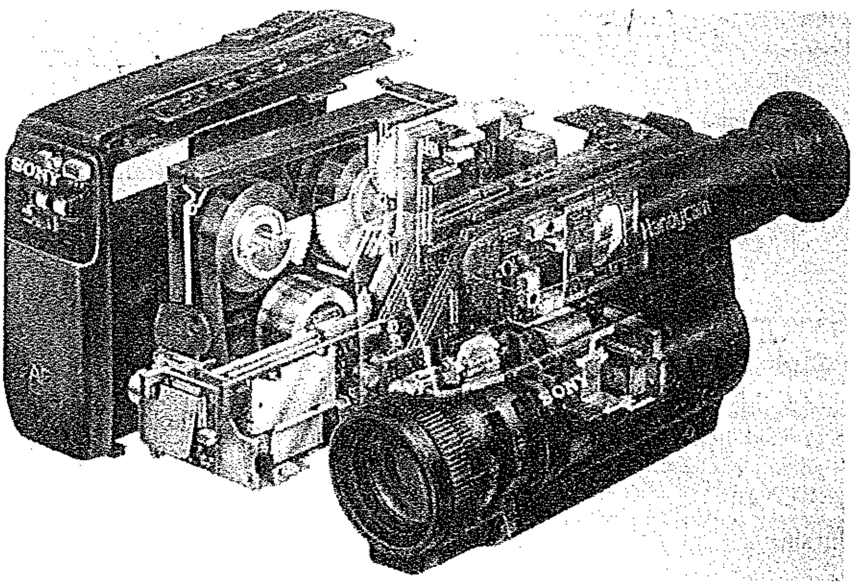
Metal	Symbol	Work Function, eV
Cesium	Cs	1.9
Potassium	K	2.2
Sodium	Na	2.3
Lithium	Li	2.5
Calcium	Ca	3.2
Copper	Cu	4.7
Silver	Ag	4.7
Platinum	Pt	6.4

to pull an electron from a free atom of that metal (see Fig. 7.10); for instance, the ionization energy of cesium is 3.9 eV compared with its work function of 1.9 eV. Since the visible spectrum extends from about 4.3 to about 7.5×10^{14} Hz, which corresponds to quantum energies of 1.7 to 3.3 eV, it is clear from Table 2.1 that the photoelectric effect is a phenomenon of the visible and ultraviolet regions.

According to Einstein, the photoelectric effect in a given metal should obey the equation

Photoelectric effect
$$h\nu = \text{KE}_{\text{max}} + \phi \tag{2.8}$$

where $h\nu$ is the photon energy, KE_{max} is the maximum photoelectron energy (which is proportional to the stopping potential), and ϕ is the minimum energy needed for an



All light-sensitive detectors, including the eye and the one used in this video camera, are based on the absorption of energy from photons of light by electrons in the atoms the light falls on.

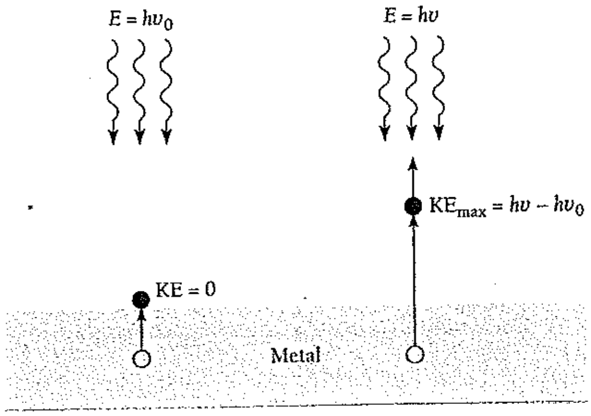


Figure 2.13 If the energy $h\nu_0$ (the work function of the surface) is needed to remove an electron from a metal surface, the maximum electron kinetic energy will be $h\nu - h\nu_0$ when light of frequency ν is directed at the surface.

electron to leave the metal. Because $\phi = h\nu_0$, Eq. (2.8) can be rewritten (Fig. 2.13)

$$h\nu = \text{KE}_{\text{max}} + h\nu_0$$

$$\text{KE}_{\text{max}} = h\nu - h\nu_0 = h(\nu - \nu_0) \tag{2.9}$$

This formula accounts for the relationships between KE_{max} and ν plotted in Fig. 2.12 from experimental data. If Einstein was right, the slopes of the lines should all be equal to Planck's constant h, and this is indeed the case.

In terms of electronvolts, the formula $E = h\nu$ for photon energy becomes

Photon energy

$$E = \left(\frac{6.626 \times 10^{-34}\ \text{J}\cdot\text{s}}{1.602 \times 10^{-19}\ \text{J/eV}}\right)\nu = (4.136 \times 10^{-15})\nu\ \text{eV}\cdot\text{s} \tag{2.10}$$

If we are given instead the wavelength λ of the light, then since $\nu = c/\lambda$ we have

Photon energy

$$E = \frac{(4.136 \times 10^{-15}\ \text{eV}\cdot\text{s})(2.998 \times 10^{8}\ \text{m/s})}{\lambda} = \frac{1.240 \times 10^{-6}\ \text{eV}\cdot\text{m}}{\lambda} \tag{2.11}$$

Example 2.2

Ultraviolet light of wavelength 350 nm and intensity 1.00 W/m^2 is directed at a potassium surface. (*a*) Find the maximum KE of the photoelectrons. (*b*) If 0.50 percent of the incident photons produce photoelectrons, how many are emitted per second if the potassium surface has an area of 1.00 cm^2?

Solution

(*a*) From Eq. (2.11) the energy of the photons is, since 1 nm = 1 nanometer = 10^{-9} m,

$$E_p = \frac{1.24 \times 10^{-6}\ \text{eV}\cdot\text{m}}{(350\ \text{nm})(10^{-9}\ \text{m/nm})} = 3.5\ \text{eV}$$

Table 2.1 gives the work function of potassium as 2.2 eV, so

$$KE_{max} = h\nu - \phi = 3.5 \text{ eV} - 2.2 \text{ eV} = 1.3 \text{ eV}$$

(*b*) The photon energy in joules is 5.68×10^{-19} J. Hence the number of photons that reach the surface per second is

$$n_p = \frac{E/t}{E_p} = \frac{(P/A)(A)}{E_p} = \frac{(1.00 \text{ W/m}^2)(1.00 \times 10^{-4} \text{ m}^2)}{5.68 \times 10^{-19} \text{ J/photon}} = 1.76 \times 10^{14} \text{ photons/s}$$

The rate at which photoelectrons are emitted is therefore

$$n_e = (0.0050)n_p = 8.8 \times 10^{11} \text{ photoelectrons/s}$$

Thermionic Emission

Einstein's interpretation of the photoelectric effect is supported by studies of thermionic emission. Long ago it was discovered that the presence of a very hot object increases the electric conductivity of the surrounding air. Eventually the reason for this effect was found to be the emission of electrons from such an object. Thermionic emission makes possible the operation of such devices as television picture tubes, in which metal filaments or specially coated cathodes at high temperature supply dense streams of electrons.

The emitted electrons evidently obtain their energy from the thermal agitation of the particles of the metal, and we would expect the electrons to need a certain minimum energy to escape. This minimum energy can be determined for many surfaces, and it is always close to the photoelectric work function for the same surfaces. In photoelectric emission, photons of light provide the energy required by an electron to escape, while in thermionic emission heat does so.

2.4 WHAT IS LIGHT?

Both wave and particle

The concept that light travels as a series of little packets is directly opposed to the wave theory of light (Fig. 2.14). Both views have strong experimental support, as we have seen. According to the wave theory, light waves leave a source with their energy spread out continuously through the wave pattern. According to the quantum theory, light consists of individual photons, each small enough to be absorbed by a single electron. Yet, despite the particle picture of light it presents, the quantum theory needs the frequency of the light to describe the photon energy.

Which theory are we to believe? A great many scientific ideas have had to be revised or discarded when they were found to disagree with new data. Here, for the first time, two different theories are needed to explain a single phenomenon. This situation is not the same as it is, say, in the case of relativistic versus newtonian mechanics, where one turns out to be an approximation of the other. The connection between the wave and quantum theories of light is something else entirely.

To appreciate this connection, let us consider the formation of a double-slit interference pattern on a screen. In the wave model, the light intensity at a place on the screen depends on $\overline{E^2}$, the average over a complete cycle of the square of the instantaneous magnitude E of the em wave's electric field. In the particle model, this

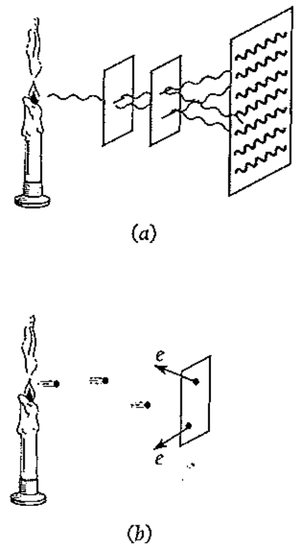


Figure 2.14 (*a*) The wave theory of light explains diffraction and interference, which the quantum theory cannot account for. (*b*) The quantum theory explains the photoelectric effect, which the wave theory cannot account for.

intensity depends instead on $Nh\nu$, where N is the number of photons per second per unit area that reach the same place on the screen. Both descriptions must give the same value for the intensity, so N is proportional to $\overline{E^2}$. If N is large enough, somebody looking at the screen would see the usual double-slit interference pattern and would have no reason to doubt the wave model. If N is small—perhaps so small that only one photon at a time reaches the screen—the observer would find a series of apparently random flashes and would assume that he or she is watching quantum behavior.

If the observer keeps track of the flashes for long enough, though, the pattern they form will be the same as when N is large. Thus the observer is entitled to conclude that the *probability* of finding a photon at a certain place and time depends on the value of $\overline{E^2}$ there. If we regard each photon as somehow having a wave associated with it, the intensity of this wave at a given place on the screen determines the likelihood that a photon will arrive there. When it passes through the slits, light is behaving as a wave does. When it strikes the screen, light is behaving as a particle does. Apparently light travels as a wave but absorbs and gives off energy as a series of particles.

We can think of light as having a dual character. **The wave theory and the quantum theory complement each other.** Either theory by itself is only part of the story and can explain only certain effects. A reader who finds it hard to understand how light can be both a wave and a stream of particles is in good company: shortly before his death, Einstein remarked that "All these fifty years of conscious brooding have brought me no nearer to the answer to the question, 'What are light quanta?'" The "true nature" of light includes both wave and particle characters, even though there is nothing in everyday life to help us visualize that.

2.5 X-RAYS

They consist of high-energy photons

The photoelectric effect provides convincing evidence that photons of light can transfer energy to electrons. Is the inverse process also possible? That is, can part or all of the kinetic energy of a moving electron be converted into a photon? As it happens, the inverse photoelectric effect not only does occur but had been discovered (though not understood) before the work of Planck and Einstein.

In 1895 Wilhelm Roentgen found that a highly penetrating radiation of unknown nature is produced when fast electrons impinge on matter. These **x-rays** were soon found to travel in straight lines, to be unaffected by electric and magnetic fields, to pass readily through opaque materials, to cause phosphorescent substances to glow, and to expose photographic plates. The faster the original electrons, the more penetrating the resulting x-rays, and the greater the number of electrons, the greater the intensity of the x-ray beam.

Not long after this discovery it became clear that x-rays are em waves. Electromagnetic theory predicts that an accelerated electric charge will radiate em waves, and a rapidly moving electron suddenly brought to rest is certainly accelerated. Radiation produced under these circumstances is given the German name **bremsstrahlung** ("braking radiation"). Energy loss due to bremsstrahlung is more important for electrons than for heavier particles because electrons are more violently accelerated when passing near nuclei in their paths. The greater the energy of an electron and the greater the atomic number of the nuclei it encounters, the more energetic the bremsstrahlung.

Wilhelm Konrad Roentgen (1845–1923) was born in Lennep, Germany, and studied in Holland and Switzerland. After periods at several German universities, Roentgen became professor of physics at Würzburg where, on November 8, 1895, he noticed that a sheet of paper coated with barium platinocyanide glowed when he switched on a nearby cathode-ray tube that was entirely covered with black cardboard. In a cathode-ray tube electrons are accelerated in a vacuum by an electric field, and it was the impact of these electrons on the glass end of the tube that produced the penetrating "x" (since their nature was then unknown) rays that caused the salt to glow. Roentgen said of his discovery that, when people heard of it, they would say, "Roentgen has probably gone crazy." In fact, x-rays were an immediate sensation, and only two months later were being used in medicine. They also stimulated research in new directions; Becquerel's discovery of radioactivity followed within a year. Roentgen received the first Nobel Prize in physics in 1902. He refused to benefit financially from his work and died in poverty in the German inflation that followed the end of World War I.

In 1912 a method was devised for measuring the wavelengths of x-rays. A diffraction experiment had been recognized as ideal, but as we recall from physical optics, the spacing between adjacent lines on a diffraction grating must be of the same order of magnitude as the wavelength of the light for satisfactory results, and gratings cannot be ruled with the minute spacing required by x-rays. Max von Laue realized that the wavelengths suggested for x-rays were comparable to the spacing between adjacent atoms in crystals. He therefore proposed that crystals be used to diffract x-rays, with their regular lattices acting as a kind of three-dimensional grating. In experiments carried out the following year, wavelengths from 0.013 to 0.048 nm were found, 10^{-4} of those in visible light and hence having quanta 10^4 times as energetic.

Electromagnetic radiation with wavelengths from about 0.01 to about 10 nm falls into the category of x-rays. The boundaries of this category are not sharp: the shorter-wavelength end overlaps gamma rays and the longer-wavelength end overlaps ultraviolet light (see Fig. 2.2).

Figure 2.15 is a diagram of an x-ray tube. A cathode, heated by a filament through which an electric current is passed, supplies electrons by thermionic emission. The high potential difference V maintained between the cathode and a metallic target accelerates the electrons toward the latter. The face of the target is at an angle relative to the electron beam, and the x-rays that leave the target pass through the

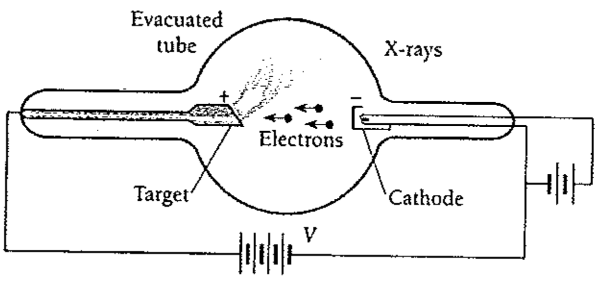
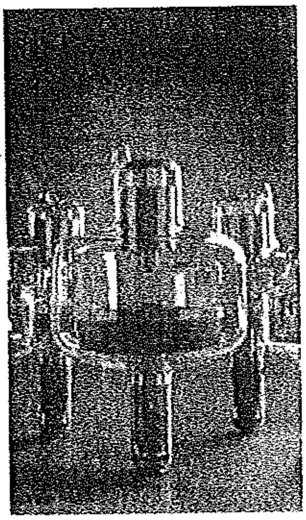


Figure 2.15 An x-ray tube. The higher the accelerating voltage V, the faster the electrons and the shorter the wavelengths of the x-rays.



In modern x-ray tubes like these, circulating oil carries heat away from the target and releases it to the outside air through a heat exchanger. The use of x-rays as a diagnostic tool in medicine is based upon the different extents to which different tissues absorb them. Because of its calcium content, bone is much more opaque to x-rays than muscle, which in turn is more opaque than fat. To enhance contrast, "meals" that contain barium are given to patients to better display their digestive systems, and other compounds may be injected into the bloodstream to enable the condition of blood vessels to be studied.

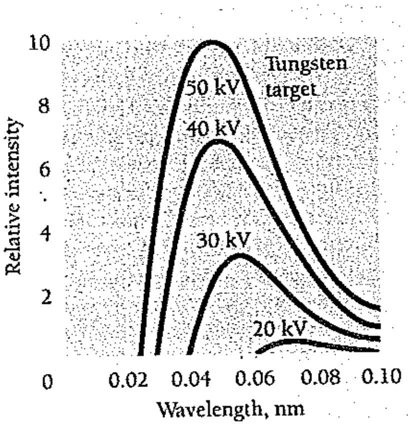


Figure 2.16 X-ray spectra of tungsten at various accelerating potentials.

side of the tube. The tube is evacuated to permit the electrons to get to the target unimpeded.

As mentioned earlier, classical electromagnetic theory predicts bremsstrahlung when electrons are accelerated, which accounts in general for the x-rays produced by an x-ray tube. However, the agreement between theory and experiment is not satisfactory in certain important respects. Figures 2.16 and 2.17 show the x-ray spectra that result when tungsten and molybdenum targets are bombarded by electrons at several different accelerating potentials. The curves exhibit two features electromagnetic theory cannot explain:

1 In the case of molybdenum, intensity peaks occur that indicate the enhanced production of x-rays at certain wavelengths. These peaks occur at specific wavelengths for each target material and originate in rearrangements of the electron structures of the

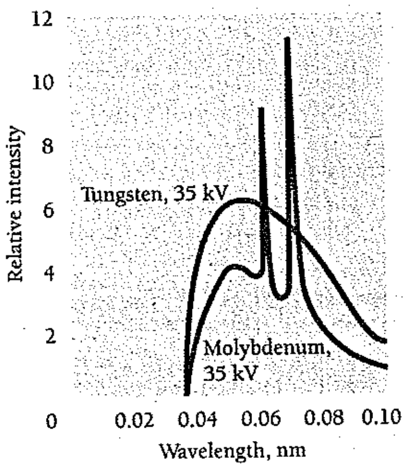
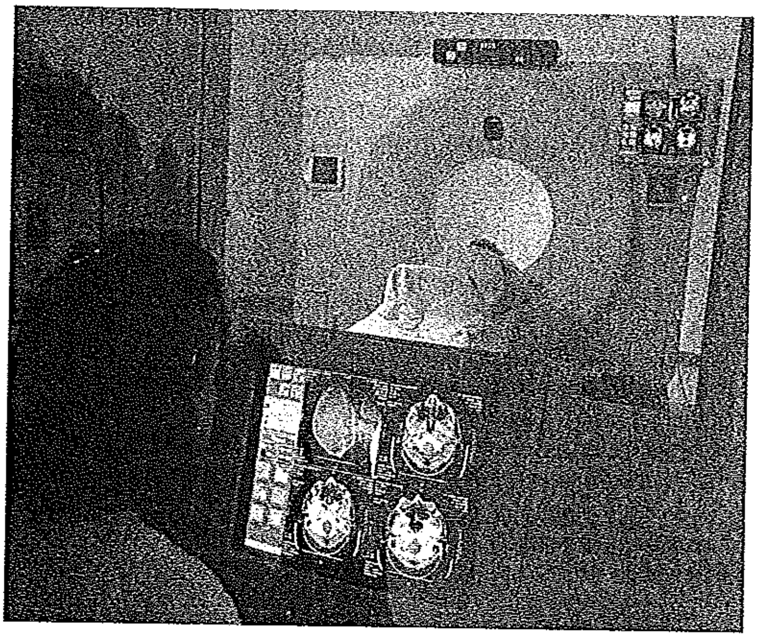


Figure 2.17 X-ray spectra of tungsten and molybdenum at 35 kV accelerating potential.



In a CT (computerized tomography) scanner, a series of x-ray exposures of a patient taken from different directions are combined by a computer to give cross-sectional images of the parts of the body being examined. In effect, the tissue is sliced up by the computer on the basis of the x-ray exposures, and any desired slice can be displayed. This technique enables an abnormality to be detected and its exact location established, which might be impossible to do from an ordinary x-ray picture. (The word tomography comes from tomos, Greek for "cut.")

target atoms after having been disturbed by the bombarding electrons. This phenomenon will be discussed in Sec. 7.9; the important thing to note at this point is the presence of x-rays of specific wavelengths, a decidedly nonclassical effect, in addition to a continuous x-ray spectrum.

2 The x-rays produced at a given accelerating potential V vary in wavelength, but none has a wavelength shorter than a certain value λ_{min}. Increasing V decreases λ_{min}. At a particular V, λ_{min} is the *same* for both the tungsten and molybdenum targets. Duane and Hunt found experimentally that λ_{min} is inversely proportional to V; their precise relationship is

X-ray production

$$\lambda_{min} = \frac{1.24 \times 10^{-6}}{V} \text{ V} \cdot \text{m} \tag{2.12}$$

The second observation fits in with the quantum theory of radiation. Most of the electrons that strike the target undergo numerous glancing collisions, with their energy going simply into heat. (This is why the targets in x-ray tubes are made from high-melting-point metals such as tungsten, and a means of cooling the target is usually employed.) A few electrons, though, lose most or all of their energy in single collisions with target atoms. This is the energy that becomes x-rays.

X-rays production, then, except for the peaks mentioned in observation 1 above, represents an inverse photoelectric effect. Instead of photon energy being transformed into electron KE, electron KE is being transformed into photon energy. A short wavelength means a high frequency, and a high frequency means a high photon energy $h\nu$.

Since work functions are only a few electronvolts whereas the accelerating potentials in x-ray tubes are typically tens or hundreds of thousands of volts, we can ignore the work function and interpret the short wavelength limit of Eq. (2.12) as corresponding to the case where the entire kinetic energy KE = Ve of a bombarding electron is given up to a single photon of energy $h\nu_{max}$. Hence

$$Ve = h\nu_{max} = \frac{hc}{\lambda_{min}}$$

$$\lambda_{min} = \frac{hc}{Ve} = \frac{1.240 \times 10^{-6}}{V} \text{ V} \cdot \text{m}$$

which is the Duane-Hunt formula of Eq. (2.12)—and, indeed, the same as Eq. (2.11) except for different units. It is therefore appropriate to regard x-ray production as the inverse of the photoelectric effect.

Example 2.3

Find the shortest wavelength present in the radiation from an x-ray machine whose accelerating potential is 50,000 V.

Solution

From Eq. (2.12) we have

$$\lambda_{min} = \frac{1.24 \times 10^{-6} \text{ V} \cdot \text{m}}{5.00 \times 10^{4} \text{ V}} = 2.48 \times 10^{-11} \text{ m} = 0.0248 \text{ nm}$$

This wavelength corresponds to the frequency

$$\nu_{max} = \frac{c}{\lambda_{min}} = \frac{3.00 \times 10^{8} \text{ m/s}}{2.48 \times 10^{-11} \text{ m}} = 1.21 \times 10^{19} \text{ Hz}$$

2.6 X-RAY DIFFRACTION

How x-ray wavelengths can be determined

A crystal consists of a regular array of atoms, each of which can scatter em waves. The mechanism of scattering is straightforward. An atom in a constant electric field becomes polarized since its negatively charged electrons and positively charged nucleus experience forces in opposite directions. These forces are small compared with the forces holding the atom together, and so the result is a distorted charge distribution equivalent to an electric dipole. In the presence of the alternating electric field of an em wave of frequency ν, the polarization changes back and forth with the same frequency ν. An oscillating electric dipole is thus created at the expense of some of the energy of the incoming wave. The oscillating dipole in turn radiates em waves of frequency ν, and these secondary waves go out in all directions except along the dipole axis. (In an assembly of atoms exposed to unpolarized radiation, the latter restriction does not apply since the contributions of the individual atoms are random.)

In wave terminology, the secondary waves have spherical wave fronts in place of the plane wave fronts of the incoming waves (Fig. 2.18). The scattering process, then,

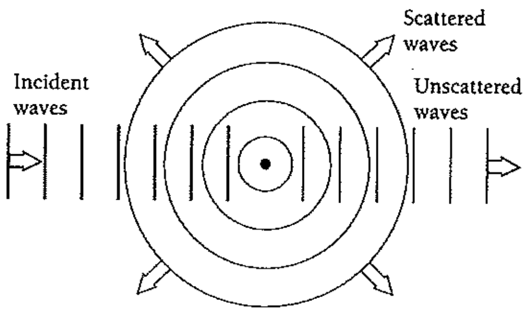


Figure 2.18 The scattering of electromagnetic radiation by a group of atoms. Incident plane waves are reemitted as spherical waves.

involves atoms that absorb incident plane waves and reemit spherical waves of the same frequency.

A monochromatic beam of x-rays that falls upon a crystal will be scattered in all directions inside it. However, owing to the regular arrangement of the atoms, in certain directions the scattered waves will constructively interfere with one another while in others they will destructively interfere. The atoms in a crystal may be thought of as defining families of parallel planes, as in Fig. 2.19, with each family having a characteristic separation between its component planes. This analysis was suggested in 1913 by W. L Bragg, in honor of whom the above planes are called **Bragg planes.**

The conditions that must be fulfilled for radiation scattered by crystal atoms to undergo constructive interference may be obtained from a diagram like that in Fig. 2.20. A beam containing x-rays of wavelength λ is incident upon a crystal at an angle θ with a family of Bragg planes whose spacing is d. The beam goes past atom A in the first plane and atom B in the next, and each of them scatters part of the beam in random directions. Constructive interference takes place only between those scattered rays that are parallel and whose paths differ by exactly λ, 2λ, 3λ, and so on. That is, the path difference must be $n\lambda$, where n is an integer. The only rays scattered by A and B for which this is true are those labeled I and II in Fig. 2.20.

The first condition on I and II is that their common scattering angle be equal to the angle of incidence θ of the original beam. (This condition, which is independent

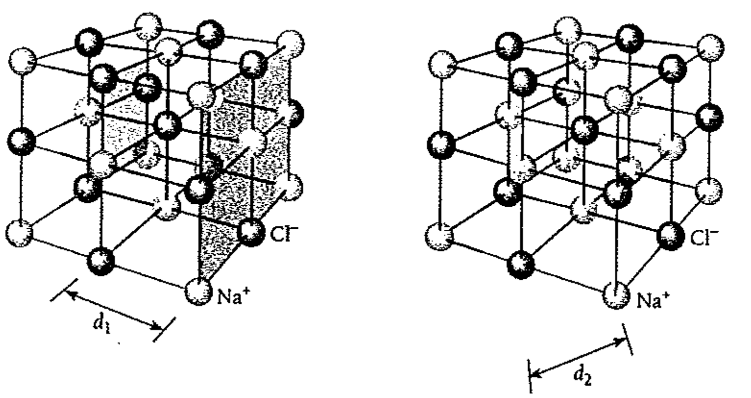
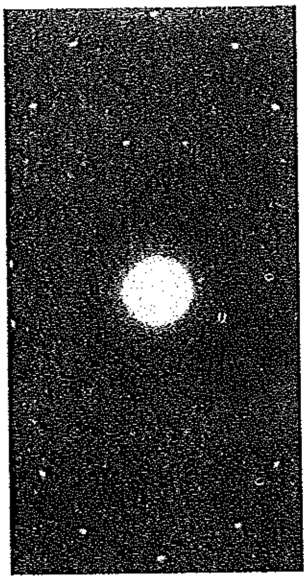


Figure 2.19 Two sets of Bragg planes in a NaCl crystal.



The interference pattern produced by the scattering of x-rays from ions in a crystal of NaCl. The bright spots correspond to the directions where x-rays scattered from various layers in the crystal interfere constructively. The cubic pattern of the NaCl lattice is suggested by he fourfold symmetry of the pattern. The large central spot is due to the unscattered x-ray beam.

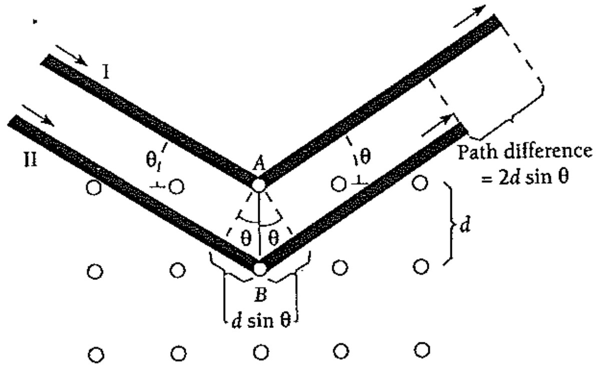


Figure 2.20 X-ray scattering from a cubic crystal.

of wavelength, is the same as that for ordinary specular reflection in optics: angle of incidence = angle of reflection.) The second condition is that

$$2d \sin \theta = n\lambda \qquad n = 1, 2, 3, \ldots \tag{2.13}$$

since ray II must travel the distance $2d \sin \theta$ farther than ray I. The integer n is the **order** of the scattered beam.

The schematic design of an x-ray spectrometer based upon Bragg's analysis is shown in Fig. 2.21. A narrow beam of x-rays falls upon a crystal at an angle θ, and a detector is placed so that it records those rays whose scattering angle is also θ. Any x-rays reaching the detector therefore obey the first Bragg condition. As θ is varied, the detector

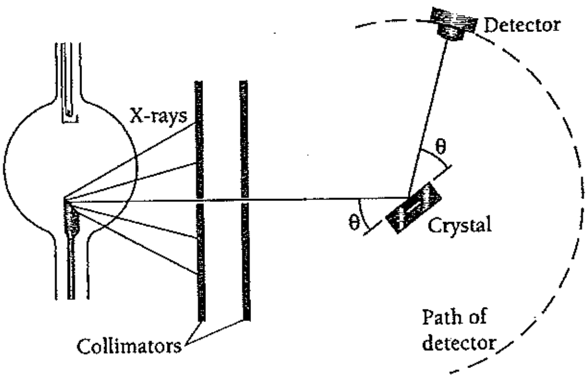


Figure 2.21 X-ray spectrometer.

will record intensity peaks corresponding to the orders predicted by Eq. (2.13). If the spacing d between adjacent Bragg planes in the crystal is known, the x-ray wavelength λ may be calculated.

2.7 COMPTON EFFECT

Further confirmation of the photon model

According to the quantum theory of light, photons behave like particles except for their lack of rest mass. How far can this analogy be carried? For instance, can we consider a collision between a photon and an electron as if both were billiard balls?

Figure 2.22 shows such a collision: an x-ray photon strikes an electron (assumed to be initially at rest in the laboratory coordinate system) and is scattered away from its original direction of motion while the electron receives an impulse and begins to move. We can think of the photon as losing an amount of energy in the collision that is the same as the kinetic energy KE gained by the electron, although actually separate photons are involved. If the initial photon has the frequency ν associated with it, the scattered photon has the lower frequency ν', where

$$\text{Loss in photon energy} = \text{gain in electron energy}$$
$$h\nu - h\nu' = \text{KE} \tag{2.14}$$

From Chap. 1 we recall that the momentum of a massless particle is related to its energy by the formula

$$E = pc \tag{1.25}$$

Since the energy of a photon is $h\nu$, its momentum is

Photon momentum
$$p = \frac{E}{c} = \frac{h\nu}{c} \tag{2.15}$$

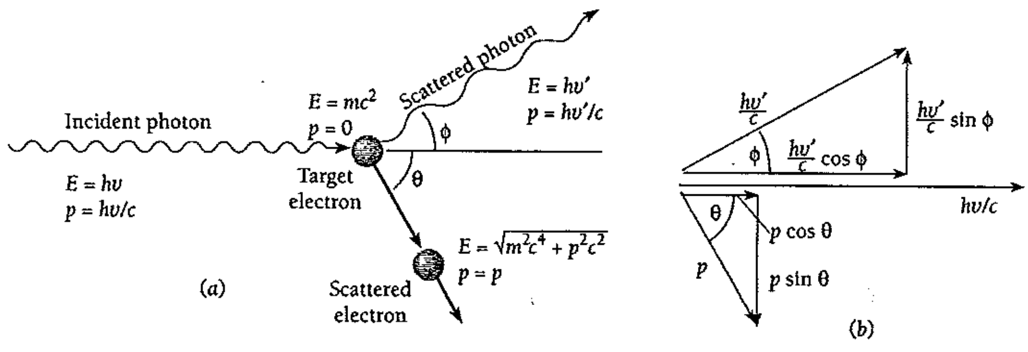


Figure 2.22 (*a*) The scattering of a photon by an electron is called the Compton effect. Energy and momentum are conserved in such an event, and as a result the scattered photon has less energy (longer wavelength) than the incident photon. (*b*) Vector diagram of the momenta and their components of the incident and scattered photons and the scattered electron.

Momentum, unlike energy, is a vector quantity that incorporates direction as well as magnitude, and in the collision momentum must be conserved in each of two mutually perpendicular directions. (When more than two bodies participate in a collision, momentum must be conserved in each of three mutually perpendicular directions.) The directions we choose here are that of the original photon and one perpendicular to it in the plane containing the electron and the scattered photon (Fig. 2.22).

The initial photon momentum is $h\nu/c$, the scattered photon momentum is $h\nu'/c$, and the initial and final electron momenta are respectively 0 and p. In the original photon direction

$$\text{Initial momentum} = \text{final momentum}$$

$$\frac{h\nu}{c} + 0 = \frac{h\nu'}{c}\cos\phi + p\cos\theta \tag{2.16}$$

and perpendicular to this direction

$$\text{Initial momentum} = \text{final momentum}$$

$$0 = \frac{h\nu'}{c}\sin\phi - p\sin\theta \tag{2.17}$$

The angle ϕ is that between the directions of the initial and scattered photons, and θ is that between the directions of the initial photon and the recoil electron. From Eqs. (2.14), (2.16), and (2.17) we can find a formula that relates the wavelength difference between initial and scattered photons with the angle ϕ between their directions, both of which are readily measurable quantities (unlike the energy and momentum of the recoil electron).

The first step is to multiply Eqs. (2.16) and (2.17) by c and rewrite them as

$$pc\cos\theta = h\nu - h\nu'\cos\phi$$

$$pc\sin\theta = h\nu'\sin\phi$$

By squaring each of these equations and adding the new ones together, the angle θ is eliminated, leaving

$$p^2c^2 = (h\nu)^2 - 2(h\nu)(h\nu')\cos\phi + (h\nu')^2 \tag{2.18}$$

Next we equate the two expressions for the total energy of a particle

$$E = \text{KE} + mc^2 \tag{1.20}$$

$$E = \sqrt{m^2c^4 + p^2c^2} \tag{1.24}$$

from Chap. 1 to give

$$(\text{KE} + mc^2)^2 = m^2c^4 + p^2c^2$$

$$p^2c^2 = \text{KE}^2 + 2mc^2\,\text{KE}$$

Since

$$KE = h\nu - h\nu'$$

we have

$$p^2c^2 = (h\nu)^2 - 2(h\nu)(h\nu') + (h\nu')^2 + 2mc^2(h\nu - h\nu') \qquad (2.19)$$

Substituting this value of p^2c^2 in Eq. (2.18), we finally obtain

$$2mc^2(h\nu - h\nu') = 2(h\nu)(h\nu')(1 - \cos\phi) \qquad (2.20)$$

This relationship is simpler when expressed in terms of wavelength λ. Dividing Eq. (2.20) by $2h^2 c^2$,

$$\frac{mc}{h}\left(\frac{\nu}{c} - \frac{\nu'}{c}\right) = \frac{\nu}{c}\frac{\nu'}{c}(1 - \cos\phi)$$

and so, since $\nu/c = 1/\lambda$ and $\nu'/c = 1/\lambda'$,

$$\frac{mc}{h}\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right) = \frac{1 - \cos\phi}{\lambda\lambda'}$$

Compton effect

$$\lambda' - \lambda = \frac{h}{mc}(1 - \cos\phi) \qquad (2.21)$$

Equation (2.21) was derived by Arthur H. Compton in the early 1920s, and the phenomenon it describes, which he was the first to observe, is known as the **Compton effect.** It constitutes very strong evidence in support of the quantum theory of radiation.

Equation (2.21) gives the change in wavelength expected for a photon that is scattered through the angle ϕ by a particle of rest mass m. This change is independent of the wavelength λ of the incident photon. The quantity

Compton wavelength

$$\lambda_C = \frac{h}{mc} \qquad (2.22)$$

is called the **Compton wavelength** of the scattering particle. For an electron $\lambda_C = 2.426 \times 10^{-12}$ m, which is 2.426 pm (1 pm = 1 picometer = 10^{-12} m). In terms of λ_C, Eq. (2.21) becomes

Compton effect

$$\lambda' - \lambda = \lambda_C(1 - \cos\phi) \qquad (2.23)$$

The Compton wavelength gives the scale of the wavelength change of the incident photon. From Eq. (2.23) we note that the greatest wavelength change possible corresponds to $\phi = 180°$, when the wavelength change will be twice the Compton wavelength λ_C. Because $\lambda_C = 2.426$ pm for an electron, and even less for other particles owing to their larger rest masses, the maximum wavelength change in the Compton effect is 4.852 pm. Changes of this magnitude or less are readily observable only in x-rays: the shift in wavelength for visible light is less than 0.01 percent of the initial wavelength, whereas for x-rays of $\lambda = 0.1$ nm it is several percent. The Compton effect is the chief means by which x-rays lose energy when they pass through matter.

Arthur Holly Compton (1892–1962), a native of Ohio, was educated at College of Wooster and Princeton. While at Washington University in St. Louis he found that x-rays increase in wavelength when scattered, which he explained in 1923 on the basis of the quantum theory of light. This work convinced remaining doubters of the reality of photons.

After receiving the Nobel Prize in 1927, Compton, now at the University of Chicago, studied cosmic rays and helped establish that they are fast charged particles (today known to be atomic nuclei, largely protons) that circulate in space and are not high-energy gamma rays as many had thought. He did this by showing that cosmic-ray intensity varies with latitude, which makes sense only if they are ions whose paths are influenced by the earth's magnetic field. During World War II Compton was one of the leaders in the development of the atomic bomb.

Example 2.4

X-rays of wavelength 10.0 pm are scattered from a target. (*a*) Find the wavelength of the x-rays scattered through 45°. (*b*) Find the maximum wavelength present in the scattered x-rays. (*c*) Find the maximum kinetic energy of the recoil electrons.

Solution

(*a*) From Eq. (2.23), $\lambda' - \lambda = \lambda_C(1 - \cos\phi)$, and so

$$\begin{aligned}\lambda' &= \lambda + \lambda_C(1 - \cos 45°)\\ &= 10.0 \text{ pm} + 0.293\lambda_C\\ &= 10.7 \text{ pm}\end{aligned}$$

(*b*) $\lambda' - \lambda$ is a maximum when $(1 - \cos\phi) = 2$, in which case

$$\lambda' = \lambda + 2\lambda_C = 10.0 \text{ pm} + 4.9 \text{ pm} = 14.9 \text{ pm}$$

(*c*) The maximum recoil kinetic energy is equal to the difference between the energies of the incident and scattered photons, so

$$\text{KE}_{max} = h(\nu - \nu') = hc\left(\frac{1}{\lambda} - \frac{1}{\lambda'}\right)$$

where λ' is given in (*b*). Hence

$$\begin{aligned}\text{KE}_{max} &= \frac{(6.626\times10^{-34}\text{ J}\cdot\text{s})(3.00\times10^{8}\text{ m/s})}{10^{-12}\text{ m/pm}}\left(\frac{1}{10.0\text{ pm}} - \frac{1}{14.9\text{ pm}}\right)\\ &= 6.54\times10^{-15}\text{ J}\end{aligned}$$

which is equal to 40.8 keV.

The experimental demonstration of the Compton effect is straightforward. As in Fig. 2.23, a beam of x-rays of a single, known wavelength is directed at a target, and the wavelengths of the scattered x-rays are determined at various angles ϕ. The results, shown in Fig. 2.24, exhibit the wavelength shift predicted by Eq. (2.21), but at each angle the scattered x-rays also include many that have the initial wavelength. This is not hard to understand. In deriving Eq. (2.21) it was assumed that the scattering particle is able to move freely, which is reasonable since many of the electrons in matter

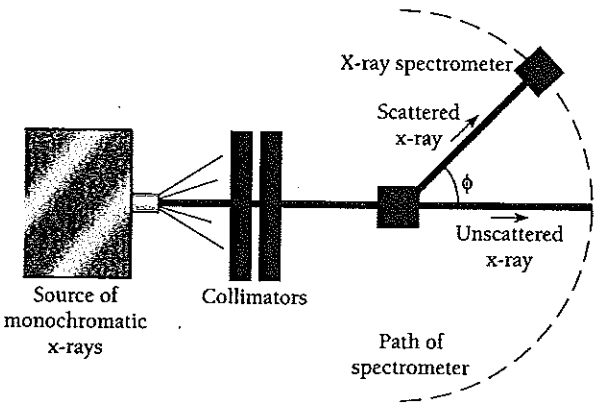


Figure 2.23 Experimental demonstration of the Compton effect.

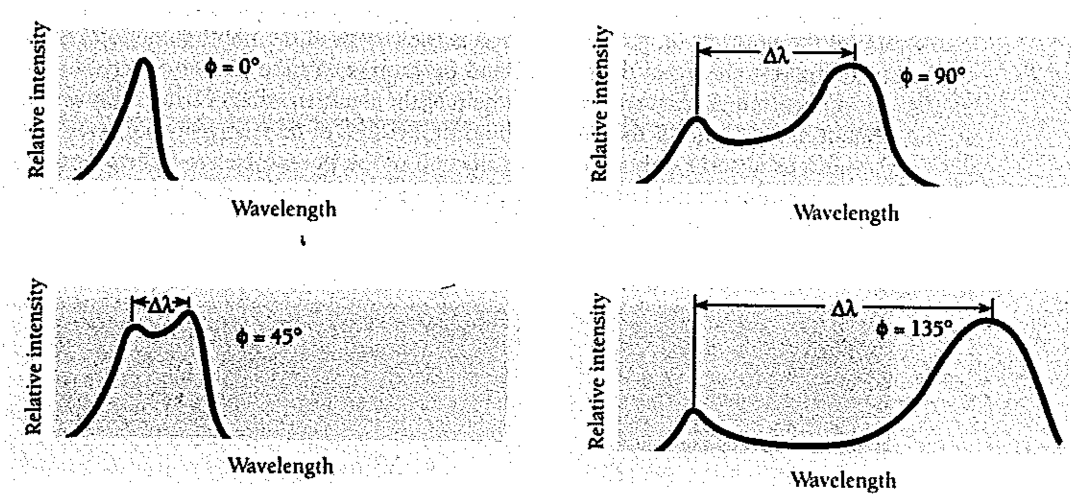


Figure 2.24 Experimental confirmation of Compton scattering. The greater the scattering angle, the greater the wavelength change, in accord with Eq. (2.21).

are only loosely bound to their parent atoms. Other electrons, however, are very tightly bound and when struck by a photon, the entire atom recoils instead of the single electron. In this event the value of m to use in Eq. (2.21) is that of the entire atom, which is tens of thousands of times greater than that of an electron, and the resulting Compton shift is accordingly so small as to be undetectable.

2.8 PAIR PRODUCTION

Energy into matter

As we have seen, in a collision a photon can give an electron all of its energy (the photoelectric effect) or only part (the Compton effect). It is also possible for a photon to materialize into an electron and a positron, which is a positively charged electron. In this process, called **pair production**, electromagnetic energy is converted into matter.

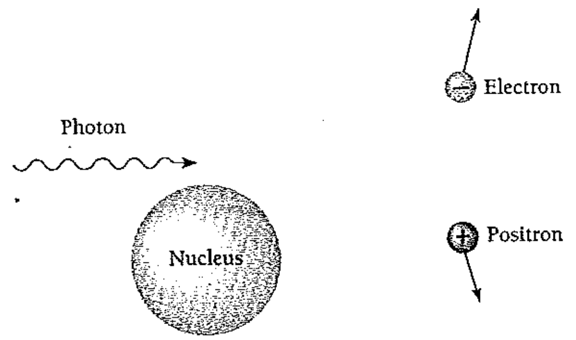
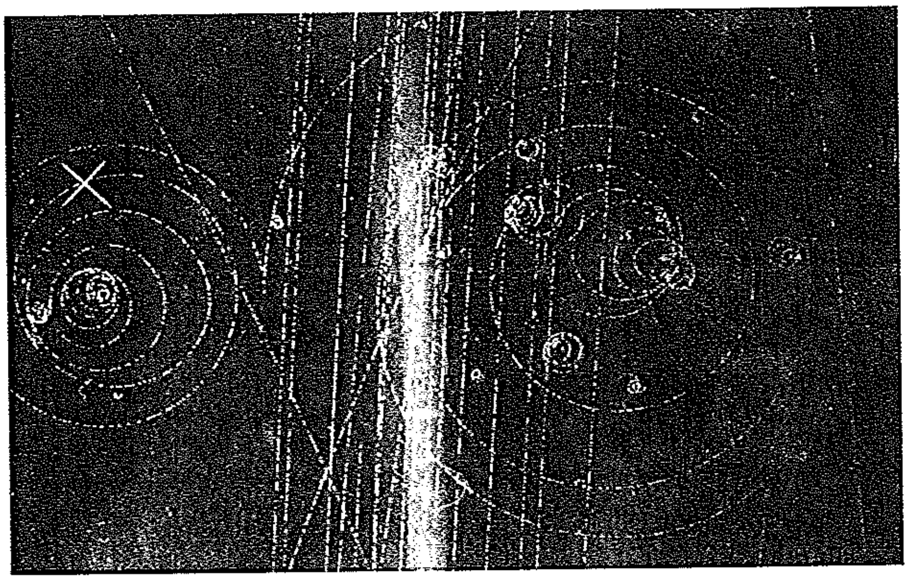


Figure 2.25 In the process of pair production, a photon of sufficient energy materializes into an electron and a positron.

No conservation principles are violated when an electron-positron pair is created near an atomic nucleus (Fig. 2.25). The sum of the charges of the electron ($q = -e$) and of the positron ($q = +e$) is zero, as is the charge of the photon; the total energy, including rest energy, of the electron and positron equals the photon energy; and linear momentum is conserved with the help of the nucleus, which carries away enough photon momentum for the process to occur. Because of its relatively enormous mass, the nucleus absorbs only a negligible fraction of the photon energy. (Energy and linear momentum could not both be conserved if pair production were to occur in empty space, so it does not occur there.)



Bubble-chamber photograph of electron-positron pair formation. A magnetic field perpendicular to the page caused the electron and positron to move in opposite curved paths, which are spirals because the particles lost energy as they moved through the chamber. In a bubble chamber, a liquid (here, hydrogen) is heated above its normal boiling point under a pressure great enough to keep it liquid. The pressure is then released, and bubbles form around any ions present in the resulting unstable superheated liquid. A charged particle moving through the liquid at this time leaves a track of bubbles that can be photographed.

The rest energy mc^2 of an electron or positron is 0.51 MeV, hence pair production requires a photon energy of at least 1.02 MeV. Any additional photon energy becomes kinetic energy of the electron and positron. The corresponding maximum photon wavelength is 1.2 pm. Electromagnetic waves with such wavelengths are called **gamma rays,** symbol γ, and are found in nature as one of the emissions from radioactive nuclei and in cosmic rays.

The inverse of pair production occurs when a positron is near an electron and the two come together under the influence of their opposite electric charges. Both particles vanish simultaneously, with the lost mass becoming energy in the form of two gamma-ray photons:

$$e^+ + e^- \rightarrow \gamma + \gamma$$

The total mass of the positron and electron is equivalent to 1.02 MeV, and each photon has an energy $h\nu$ of 0.51 MeV plus half the kinetic energy of the particles relative to their center of mass. The directions of the photons are such as to conserve both energy and linear momentum, and no nucleus or other particle is needed for this **pair annihilation** to take place.

Example 2.5

Show that pair production cannot occur in empty space.

Solution

From conservation of energy,

$$h\nu = 2\gamma mc^2$$

where $h\nu$ is the photon energy and γmc^2 is the total energy of each member of the electron-position pair. Figure 2.26 is a vector diagram of the linear momenta of the photon, electron, and positron. The angles θ are equal in order that momentum be conserved in the transverse direction. In the direction of motion of the photon, for momentum to be conserved it must be true that

$$\frac{h\nu}{c} = 2p \cos \theta$$

$$h\nu = 2pc \cos \theta$$

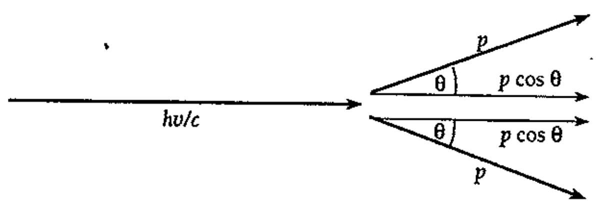


Figure 2.26 Vector diagram of the momenta involved if a photon were to materialize into an electron-positron pair in empty space. Because such an event cannot conserve both energy and momentum, it does not occur. Pair production always involves an atomic nucleus that carries away part of the initial photon momentum.

Since $p = \gamma mv$ for the electron and positron,

$$h\nu = 2\gamma mc^2\left(\frac{v}{c}\right)\cos\theta$$

Because $v/c < 1$ and $\cos\theta \leqslant 1$,

$$h\nu < 2\gamma mc^2$$

But conservation of energy requires that $h\nu = 2\gamma mc^2$. Hence it is impossible for pair production to conserve both energy and momentum unless some other object is involved in the process to carry away part of the initial photon momentum.

Example 2.6

An electron and a positron are moving side by side in the $+x$ direction at $0.500c$ when they annihilate each other. Two photons are produced that move along the x axis. (*a*) Do both photons move in the $+x$ direction? (*b*) What is the energy of each photon?

Solution

(*a*) In the center-of-mass (CM) system (which is the system moving with the original particles), the photons move off in opposite directions to conserve momentum. They must also do so in the lab system because the speed of the CM system is less than the speed c of the photons.
(*b*) Let p_1 be the momentum of the photon moving in the $+x$ direction and p_2 be the momentum of the photon moving in the $-x$ direction. Then conservation of momentum (in the lab system) gives

$$p_1 - p_2 = 2\gamma mv = \frac{2(mc^2)(v/c^2)}{\sqrt{1 - v/c^2}}$$

$$= \frac{2(0.511\ \text{MeV}/c^2)(c^2)(0.500c)/c^2}{\sqrt{1-(0.500)^2}} = 0.590\ \text{MeV}/c$$

Conservation of energy gives

$$p_1c + p_2c = 2\gamma mc^2 = \frac{2mc^2}{\sqrt{1 - v^2/c^2}} = \frac{2(0.511\ \text{MeV})}{\sqrt{1-(0.500)^2}} = 1.180\ \text{MeV}$$

and so

$$p_1 + p_2 = 1.180\ \text{MeV}/c$$

Now we add the two results and solve for p_1 and p_2:

$$(p_1 - p_2) + (p_1 + p_2) = 2p_1 = (0.590 + 1.180)\ \text{MeV}/c$$

$$p_1 = 0.885\ \text{MeV}/c$$

$$p_2 = (p_1 + p_2) - p_1 = 0.295\ \text{MeV}/c$$

The photon energies are accordingly

$$E_1 = p_1c = 0.885\ \text{MeV} \qquad E_2 = p_2c = 0.295\ \text{MeV}$$

Photon Absorption

The three chief ways in which photons of light, x-rays, and gamma rays interact with matter are summarized in Fig. 2.27. In all cases photon energy is transferred to electrons which in turn lose energy to atoms in the absorbing material.

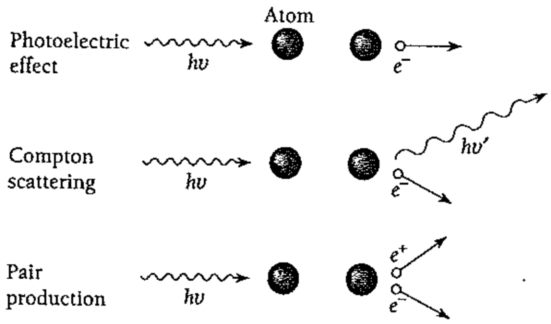


Figure 2.27 X- and gamma rays interact with matter chiefly through the photoelectric effect, Compton scattering, and pair production. Pair production requires a photon energy of at least 1.02 MeV.

At low photon energies the photoelectric effect is the chief mechanism of energy loss. The importance of the photoelectric effect decreases with increasing energy, to be succeeded by Compton scattering. The greater the atomic number of the absorber, the higher the energy at which the photoelectric effect remains significant. In the lighter elements, Compton scattering becomes dominant at photon energies of a few tens of keV, whereas in the heavier ones this does not happen until photon energies of nearly 1 MeV are reached (Fig. 2.28).

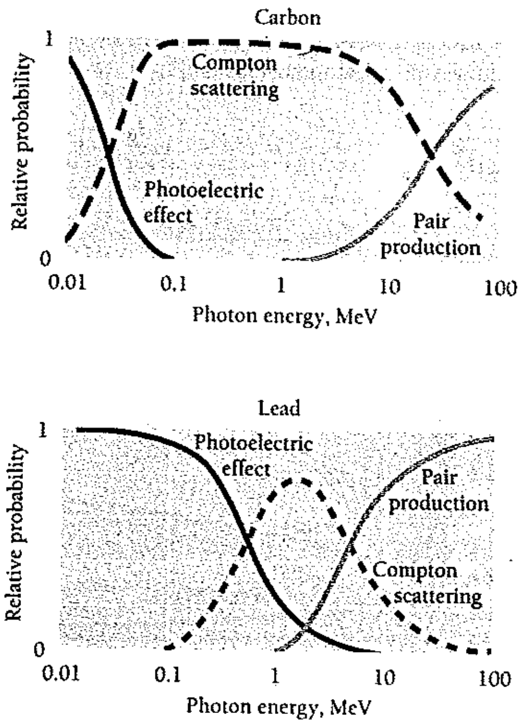


Figure 2.28 The relative probabilities of the photoelectric effect, Compton scattering, and pair production as functions of energy in carbon (a light element) and lead (a heavy element).

Pair production becomes increasingly likely the more the photon energy exceeds the threshold of 1.02 MeV. The greater the atomic number of the absorber, the lower the energy at which pair production takes over as the principal mechanism of energy loss by gamma rays. In the heaviest elements, the crossover energy is about 4 MeV, but it is over 10 MeV for the lighter ones. Thus gamma rays in the energy range typical of radioactive decay interact with matter largely through Compton scattering.

The intensity I of an x- or gamma-ray beam is equal to the rate at which it transports energy per unit cross-sectional area of the beam. The fractional energy $-dI/I$ lost by the beam in passing through a thickness dx of a certain absorber is found to be proportional to dx:

$$-\frac{dI}{I} = \mu\, dx \qquad (2.24)$$

The proportionality constant μ is called the **linear attenuation coefficient** and its value depends on the energy of the photons and on the nature of the absorbing material. Integrating Eq. (2.24) gives

Radiation intensity $$I = I_0 e^{-\mu x} \qquad (2.25)$$

The intensity of the radiation decreases exponentially with absorber thickness x. Figure 2.29 is a graph of the linear attenuation coefficient for photons in lead as a function of photon energy. The contribution to μ of the photoelectric effect, Compton scattering, and pair production are shown.

We can use Eq. (2.25) to relate the thickness x of absorber needed to reduce the intensity of an x- or gamma-ray beam by a given amount to the attenuation coefficient μ. If the ratio of the final and initial intensities is I/I_0,

$$\frac{I}{I_0} = e^{-\mu x} \qquad \frac{I_0}{I} = e^{\mu x} \qquad \ln\frac{I_0}{I} = \mu x$$

Absorber thickness $$x = \frac{\ln (I_0/I)}{\mu} \qquad (2.26)$$

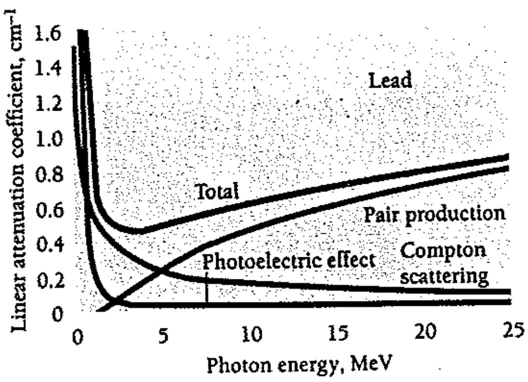


Figure 2.29 Linear attentuation coefficients for photons in lead.

Example 2.7

The linear attenuation coefficient for 2.0-MeV gamma rays in water is 4.9 m^{-1}. (*a*) Find the relative intensity of a beam of 2.0-MeV gamma rays after it has passed through 10 cm of water. (*b*) How far must such a beam travel in water before its intensity is reduced to 1 percent of its original value?

Solution

(*a*) Here $\mu x = (4.9\ m^{-1})(0.10\ m) = 0.49$ and so, from Eq. (2.25)

$$\frac{I}{I_0} = e^{-\mu x} = e^{-0.49} = 0.61$$

The intensity of the beam is reduced to 61 percent of its original value after passing through 10 cm of water.
(*b*) Since $I_0/I = 100$, Eq. (2.26) yields

$$x = \frac{\ln(I_0/I)}{\mu} = \frac{\ln 100}{4.9\ m^{-1}} = 0.94\ m$$

2.9 PHOTONS AND GRAVITY

Although they lack rest mass, photons behave as though they have gravitational mass

In Sec. 1.10 we learned that light is affected by gravity by virtue of the curvature of spacetime around a mass. Another way to approach the gravitational behavior of light follows from the observation that, although a photon has no rest mass, it nevertheless interacts with electrons as though it has the inertial mass

Photon "mass"
$$m = \frac{p}{v} = \frac{h\nu}{c^2} \tag{2.27}$$

(We recall that, for a photon, $p = h\nu/c$ and $v = c$.) According to the principle of equivalence, gravitational mass is always equal to inertial mass, so a photon of frequency ν ought to act gravitationally like a particle of mass $h\nu/c^2$.

The gravitational behavior of light can be demonstrated in the laboratory. When we drop a stone of mass m from a height H near the earth's surface, the gravitational pull of the earth accelerates it as it falls and the stone gains the energy mgH on the way to the ground. The stone's final kinetic energy $\frac{1}{2}mv^2$ is equal to mgH, so its final speed is $\sqrt{2gH}$.

All photons travel with the speed of light and so cannot go any faster. However, a photon that falls through a height H can manifest the increase of mgH in its energy by an increase in frequency from ν to ν' (Fig. 2.30). Because the frequency change is extremely small in a laboratory-scale experiment, we can neglect the corresponding change in the photon's "mass" $h\nu/c^2$.

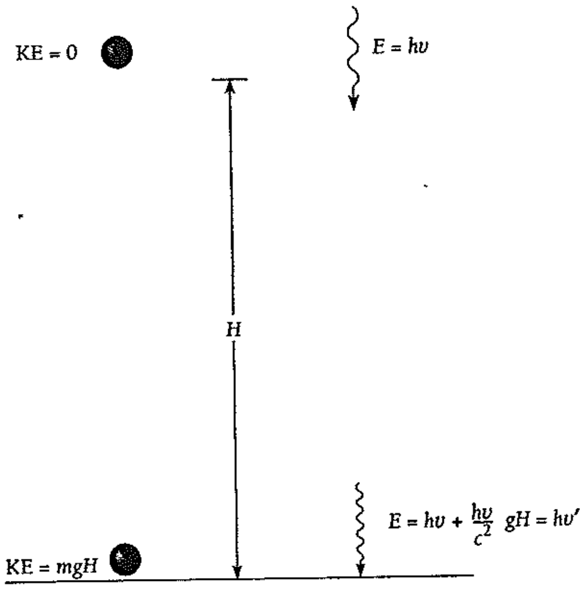


Figure 2.30 A photon that falls in a gravitational field gains energy, just as a stone does. This gain in energy is manifested as an increase in frequency from ν to ν'.

Hence,

$$\text{final photon energy} = \text{initial photon energy} + \text{increase in energy}$$

$$h\nu' = h\nu + mgH$$

and so

$$h\nu' = h\nu + \left(\frac{h\nu}{c^2}\right)gH$$

Photon energy after falling through height H

$$h\nu' = h\nu\left(1 + \frac{gH}{c^2}\right) \tag{2.28}$$

Example 2.8

The increase in energy of a fallen photon was first observed in 1960 by Pound and Rebka at Harvard. In their work H was 22.5 m. Find the change in frequency of a photon of red light whose original frequency is 7.3×10^{14} Hz when it falls through 22.5 m.

Solution

From Eq. (2.28) the change in frequency is

$$\nu' - \nu = \left(\frac{gH}{c^2}\right)\nu$$

$$= \frac{(9.8\ \text{m/s}^2)(22.5\ \text{m})(7.3 \times 10^{14}\ \text{Hz})}{(3.0 \times 10^8\ \text{m/s})^2} = 1.8\ \text{Hz}$$

Pound and Rebka actually used gamma rays of much higher frequency, as described in Exercise 53.

Gravitational Red Shift

An interesting astronomical effect is suggested by the gravitational behavior of light. If the frequency associated with a photon moving toward the earth increases, then the frequency of a photon moving away from it should decrease.

The earth's gravitational field is not particularly strong, but the fields of many stars are. Suppose a photon of initial frequency ν is emitted by a star of mass M and radius R, as in Fig. 2.31. The potential energy of a mass m on the star's surface is

$$\mathrm{PE} = -\frac{GMm}{R}$$

where the minus sign is required because the force between M and m is attractive. The potential energy of a photon of "mass" $h\nu/c^2$ on the star's surface is therefore

$$\mathrm{PE} = -\frac{GMh\nu}{c^2R}$$

and its total energy E, the sum of PE and its quantum energy $h\nu$, is

$$\mathrm{E} = h\nu - \frac{GMh\nu}{c^2R} = h\nu\left(1 - \frac{GM}{c^2R}\right)$$

At a larger distance from the star, for instance at the earth, the photon is beyond the star's gravitational field but its total energy remains the same. The photon's energy is now entirely electromagnetic, and

$$E = h\nu'$$

where ν' is the frequency of the arriving photon. (The potential energy of the photon in the earth's gravitational field is negligible compared with that in the star's field.) Hence

$$h\nu' = h\nu\left(1 - \frac{GM}{c^2R}\right)$$

$$\frac{\nu'}{\nu} = 1 - \frac{GM}{c^2R}$$

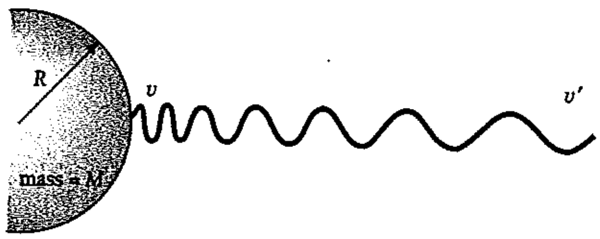


Figure 2.31 The frequency of a photon emitted from the surface of a star decreases as it moves away from the star.

and the relative frequency change is

Gravitational red shift
$$\frac{\Delta\nu}{\nu} = \frac{\nu - \nu'}{\nu} = 1 - \frac{\nu'}{\nu} = \frac{GM}{c^2R} \tag{2.29}$$

The photon has a *lower* frequency at the earth, corresponding to its loss in energy as it leaves the field of the star.

A photon in the visible region of the spectrum is thus shifted toward the red end, and this phenomenon is accordingly known as the **gravitational red shift.** It is different from the doppler red shift observed in the spectra of distant galaxies due to their apparent recession from the earth, a recession that seems to be due to a general expansion of the universe.

As we shall learn in Chap. 4, when suitably excited the atoms of every element emit photons of certain specific frequencies only. The validity of Eq. (2.29) can therefore be checked by comparing the frequencies found in stellar spectra with those in spectra obtained in the laboratory. For most stars, including the sun, the ratio M/R is too small for a gravitational red shift to be apparent. However, for a class of stars known as **white dwarfs**, it is just on the limit of measurement—and has been observed. A white dwarf is an old star whose interior consists of atoms whose electron structures have collapsed and so it is very small: a typical white dwarf is about the size of the earth but has the mass of the sun.

Black Holes

An interesting question is, what happens if a star is so dense that $GM/c^2R \geq 1$? If this is the case, then from Eq. (2.29) we see that no photon can ever leave the star, since to do so requires more energy than its initial energy $h\nu$. The red shift would, in effect, have then stretched the photon wavelength to infinity. A star of this kind cannot radiate and so would be invisible—a **black hole** in space.

In a situation in which gravitational energy is comparable with total energy, as for a photon in a black hole, general relativity must be applied in detail. The correct criterion for a star to be a black hole turns out to be $GM/c^2R \geq \frac{1}{2}$. The **Schwarzschild radius** R_S of a body of mass M is defined as

Quasars and Galaxies

In even the most powerful telescope, a **quasar** appears as a sharp point of light, just as a star does. Unlike stars, quasars are powerful sources of radio waves; hence their name, a contraction of *quasi-stellar radio sources.* Hundreds of quasars have been discovered, and there seem to be many more. Though a typical quasar is smaller than the solar system, its energy output may be thousands of times the output of our entire Milky Way galaxy.

Most astronomers believe that at the heart of every quasar is a black hole whose mass is at least that of 100 million suns. As nearby stars are pulled toward the black hole, their matter is squeezed and heated to produce the observed radiation. While being swallowed, a star may liberate 10 times as much energy as it would have given off had it lived out a normal life. A diet of a few stars a year seems enough to keep a quasar going at the observed rates. It is possible that quasars are the cores of newly formed galaxies. Did all galaxies once undergo a quasar phase? Nobody can say as yet, but there is evidence that all galaxies, including the Milky Way, contain massive black holes at their centers.

Schwarzschild radius $$R_S = \frac{2GM}{c^2} \tag{2.30}$$

The body is a black hole if all its mass is inside a sphere with this radius. The boundary of a black hole is called its **event horizon.** The escape speed from a black hole is equal to the speed of the light c at the Schwarzschild radius, hence nothing at all can ever leave a black hole. For a star with the sun's mass, R_S is 3 km, a quarter of a million times smaller than the sun's present radius. Anything passing near a black hole will be sucked into it, never to return to the outside world.

Since it is invisible, how can a black hole be detected? A black hole that is a member of a double-star system (double stars are quite common) will reveal its presence by its gravitational pull on the other star; the two stars circle each other. In addition, the intense gravitational field of the black hole will attract matter from the other star, which will be compressed and heated to such high temperatures that x-rays will be emitted profusely. One of a number of invisible objects that astronomers believe on this basis to be black holes is known as Cygnus X-1. Its mass is perhaps 8 times that of the sun, and its radius may be only about 10 km. The region around a black hole that emits x-rays should extend outward for several hundred kilometers.

Only very heavy stars end up as black holes. Lighter stars evolve into white dwarfs and neutron stars, which as their name suggests consist largely of neutrons (see Sec. 9.11). But as time goes on, the strong gravitational fields of both white dwarfs and neutron stars attract more and more cosmic dust and gas. When they have gathered up enough mass, they too will become black holes. If the universe lasts long enough, then everything in it may be in the form of black holes.

Black holes are also believed to be at the cores of galaxies. Again, the clues come from the motions of nearby bodies and from the amount and type of radiation emitted. Stars close to a galactic center are observed to move so rapidly that only the gravitational pull of an immense mass could keep them in their orbits instead of flying off. How immense? As much as a billion times the sun's mass. And, as in the case of black holes that were once stars, radiation pours out of galactic centers so copiously that only black holes could be responsible.

EXERCISES

"Why," said the Dodo, "the best way to explain it is to do it." —Lewis Carroll, *Alice's Adventures in Wonderland*

2.2 Blackbody Radiation

1. If Planck's constant were smaller than it is, would quantum phenomena be more or less conspicuous than they are now?
2. Express the Planck radiation formula in terms of wavelength.

2.3 Photoelectric Effect

3. Is it correct to say that the maximum photoelectron energy KE_{max} is proportional to the frequency ν of the incident light? If not, what would a correct statement of the relationship between KE_{max} and ν be?
4. Compare the properties of particles with those of waves. Why do you think the wave aspect of light was discovered earlier than its particle aspect?
5. Find the energy of a 700-nm photon.
6. Find the wavelength and frequency of a 100-MeV photon.
7. A 1.00-kW radio transmitter operates at a frequency of 880 kHz. How many photons per second does it emit?
8. Under favorable circumstances the human eye can detect 1.0×10^{-18} J of electromagnetic energy. How many 600-nm photons does this represent?

9. Light from the sun arrives at the earth, an average of 1.5 $\times 10^{11}$ m away, at the rate of 1.4 $\times 10^3$ W/m^2 of area perpendicular to the direction of the light. Assume that sunlight is monochromatic with a frequency of 5.0 $\times 10^{14}$ Hz. (*a*) How many photons fall per second on each square meter of the earth's surface directly facing the sun? (*b*) What is the power output of the sun, and how many photons per second does it emit? (*c*) How many photons per cubic meter are there near the earth?

10. A detached retina is being "welded" back in place using 20-ms pulses from a 0.50-W laser operating at a wavelength of 632 nm. How many photons are in each pulse?

11. The maximum wavelength for photoelectric emission in tungsten is 230 nm. What wavelength of light must be used in order for electrons with a maximum energy of 1.5 eV to be ejected?

12. The minimum frequency for photoelectric emission in copper is 1.1 $\times 10^{15}$ Hz. Find the maximum energy of the photoelectrons (in electronvolts) when light of frequency 1.5 $\times 10^{15}$ Hz is directed on a copper surface.

13. What is the maximum wavelength of light that will cause photoelectrons to be emitted from sodium? What will the maximum kinetic energy of the photoelectrons be if 200-nm light falls on a sodium surface?

14. A silver ball is suspended by a string in a vacuum chamber and ultraviolet light of wavelength 200 nm is directed at it. What electrical potential will the ball acquire as a result?

15. 1.5 mW of 400-nm light is directed at a photoelectric cell. If 0.10 percent of the incident photons produce photoelectrons, find the current in the cell.

16. Light of wavelength 400 nm is shone on a metal surface in an apparatus like that of Fig. 2.9. The work function of the metal is 2.50 eV. (*a*) Find the extinction voltage, that is, the retarding voltage at which the photoelectron current disappears. (*b*) Find the speed of the fastest photoelectrons.

17. A metal surface illuminated by 8.5 $\times 10^{14}$ Hz light emits electrons whose maximum energy is 0.52 eV. The same surface illuminated by 12.0 $\times 10^{14}$ Hz hight emits electrons whose maximum energy is 1.97 eV. From these data find Planck's constant and the work function of the surface.

18. The work function of a tungsten surface is 5.4 eV. When the surface is illuminated by light of wavelength 175 nm, the maximum photoelectron energy is 1.7 eV. Find Planck's constant from these data.

19. Show that it is impossible for a photon to give up all its energy and momentum to a free electron. This is the reason why the photoelectric effect can take place only when photons strike bound electrons.

2.5 X-Rays

20. What voltage must be applied to an x-ray tube for it to emit x-rays with a minimum wavelength of 30 pm?

21. Electrons are accelerated in television tubes through potential differences of about 10 kV. Find the highest frequency of the electromagnetic waves emitted when these electrons strike the screen of the tube. What kind of waves are these?

2.6 X-Ray Diffraction

22. The smallest angle of Bragg scattering in potassium chloride (KCl) is 28.4° for 0.30-nm x-rays. Find the distance between atomic planes in potassium chloride.

23. The distance between adjacent atomic planes in calcite ($CaCO_3$) is 0.300 nm. Find the smallest angle of Bragg scattering for 0.030-nm x-rays.

24. Find the atomic spacing in a crystal of rock salt (NaCl), whose structure is shown in Fig. 2.19. The density of rock salt is 2.16 $\times 10^3$ kg/m^3 and the average masses of the Na and Cl atoms are respectively 3.82 $\times 10^{-26}$ kg and 5.89 $\times 10^{-26}$ kg.

2.7 Compton Effect

25. What is the frequency of an x-ray photon whose momentum is 1.1 $\times 10^{-23}$ kg · m/s?

26. How much energy must a photon have if it is to have the momentum of a 10-MeV proton?

27. In Sec. 2.7 the x-rays scattered by a crystal were assumed to undergo no change in wavelength. Show that this assumption is reasonable by calculating the Compton wavelength of a Na atom and comparing it with the typical x-ray wavelength of 0.1 nm.

28. A monochromatic x-ray beam whose wavelength is 55.8 pm is scattered through 46°. Find the wavelength of the scattered beam.

29. A beam of x-rays is scattered by a target. At 45° from the beam direction the scattered x-rays have a wavelength of 2.2 pm. What is the wavelength of the x-rays in the direct beam?

30. An x-ray photon whose initial frequency was 1.5 $\times 10^{19}$ Hz emerges from a collision with an electron with a frequency of 1.2 $\times 10^{19}$ Hz. How much kinetic energy was imparted to the electron?

31. An x-ray photon of initial frequency 3.0 $\times 10^{19}$ Hz collides with an electron and is scattered through 90°. Find its new frequency.

32. Find the energy of an x-ray photon which can impart a maximum energy of 50 keV to an electron.

33. At what scattering angle will incident 100-keV x-rays leave a target with an energy of 90 keV?

34. (*a*) Find the change in wavelength of 80-pm x-rays that are scattered 120° by a target. (*b*) Find the angle between the directions of the recoil electron and the incident photon. (*c*) Find the energy of the recoil electron.

35. A photon of frequency ν is scattered by an electron initially at rest. Verify that the maximum kinetic energy of the recoil electron is $KE_{max} = (2h^2\nu^2/mc^2)/(1 + 2h\nu/mc^2)$.

36. In a Compton-effect experiment in which the incident x-rays have a wavelength of 10.0 pm, the scattered x-rays at a certain angle have a wavelength of 10.5 pm. Find the momentum (magnitude and direction) of the corresponding recoil electrons.

37. A photon whose energy equals the rest energy of the electron undergoes a Compton collision with an electron. If the electron moves off at an angle of 40° with the original photon direction, what is the energy of the scattered photon?

38. A photon of energy E is scattered by a particle of rest energy E_0. Find the maximum kinetic energy of the recoiling particle in terms of E and E_0.

2.8 Pair Production

39. A positron collides head on with an electron and both are annihilated. Each particle had a kinetic energy of 1.00 MeV. Find the wavelength of the resulting photons.

40. A positron with a kinetic energy of 2.000 MeV collides with an electron at rest and the two particles are annihilated. Two photons are produced; one moves in the same direction as the incoming positron and the other moves in the opposite direction. Find the energies of the photons.

41. Show that, regardless of its initial energy, a photon cannot undergo Compton scattering through an angle of more than 60° and still be able to produce an electron-positron pair. (*Hint:* Start by expressing the Compton wavelength of the electron in terms of the maximum photon wavelength needed for pair production.)

42. (*a*) Verify that the minimum energy a photon must have to create an electron-positron pair in the presence of a stationary nucleus of mass M is $2mc^2(1 + m/M)$, where m is the electron rest mass. (*b*) Find the minimum energy needed for pair production in the presence of a proton.

43. (*a*) Show that the thickness $x_{1/2}$ of an absorber required to reduce the intensity of a beam of radiation by a factor of 2 is given by $x_{1/2} = 0.693/\mu$. (*b*) Find the absorber thickness needed to produce an intensity reduction of a factor of 10.

44. (*a*) Show that the intensity of the radiation absorbed in a thickness x of an absorber is given by $I_0\mu x$ when $\mu x \lll 1$. (*b*) If $\mu x = 0.100$, what is the percentage error in using this formula instead of Eq. (2.25)?

45. The linear absorption coefficient for 1-MeV gamma rays in lead is 78 m^{-1}. Find the thickness of lead required to reduce by half the intensity of a beam of such gamma rays.

46. The linear absorption coefficient for 50-keV x-rays in sea-level air is 5.0×10^{-3} m^{-1}. By how much is the intensity of a beam of such x-rays reduced when it passes through 0.50 m of air? Through 5.0 m of air?

47. The linear absorption coefficients for 2.0-MeV gamma rays are 4.9 m^{-1} in water and 52 m^{-1} in lead. What thickness of water would give the same shielding for such gamma rays as 10 mm of lead?

48. The linear absorption coefficient of copper for 80-keV x-rays is 4.7×10^4 m^{-1}. Find the relative intensity of a beam of 80-keV x-rays after it has passed through a 0.10-mm copper foil.

49. What thickness of copper is needed to reduce the intensity of the beam in Exercise 48 by half?

50. The linear absorption coefficients for 0.05-nm x-rays in lead and in iron are, respectively, 5.8×10^4 m^{-1} and 1.1×10^4 m^{-1}. How thick should an iron shield be in order to provide the same protection from these x-rays as 10 mm of lead?

2.9 Photons and Gravity

51. The sun's mass is 2.0×10^{30} kg and its radius is 7.0×10^8 m. Find the approximate gravitational red shift in light of wavelength 500 nm emitted by the sun.

52. Find the approximate gravitational red shift in 500-nm light emitted by a white dwarf star whose mass is that of the sun but whose radius is that of the earth, 6.4×10^6 m.

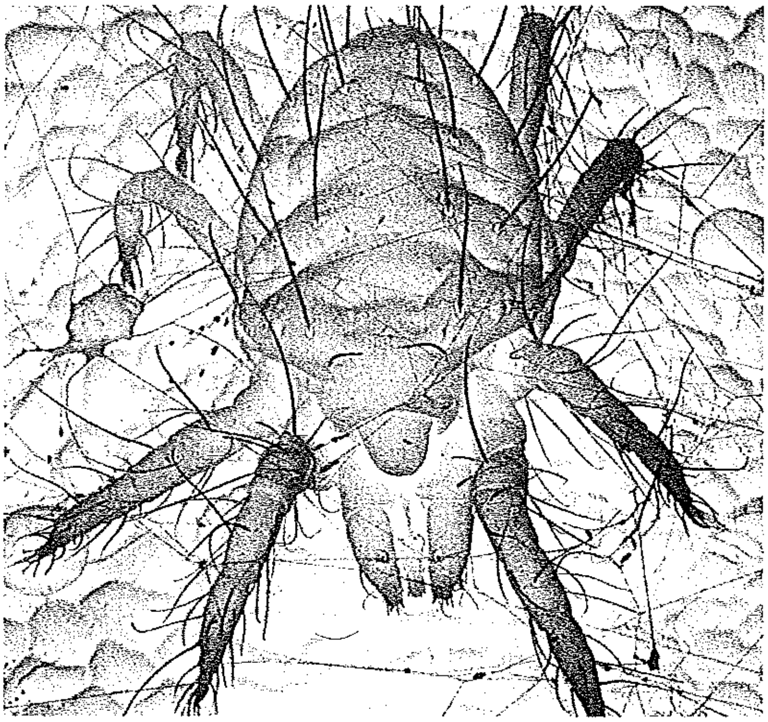
53. As discussed in Chap. 12, certain atomic nuclei emit photons in undergoing transitions from "excited" energy states to their "ground" or normal states. These photons constitute gamma rays. When a nucleus emits a photon, it recoils in the opposite direction. (*a*) The $^{57}_{27}Co$ nucleus decays by K capture to $^{57}_{26}Fe$, which then emits a photon in losing 14.4 keV to reach its ground state. The mass of a $^{57}_{26}Fe$ atom is 9.5×10^{-26} kg. By how much is the photon energy reduced from the full 14.4 keV available as a result of having to share energy and momentum with the recoiling atom? (*b*) In certain crystals the atoms are so tightly bound that the entire crystal recoils when a gamma-ray photon is emitted, instead of the individual atom. This phenomenon is known as the **Mössbauer effect.** By how much is the photon energy reduced in this situation if the excited $^{57}_{26}Fe$ nucleus is part of a 1.0-g crystal? (*c*) The essentially recoil-free emission of gamma rays in situations like that of *b* means that it is possible to construct a source of virtually monoenergetic and hence monochromatic photons. Such a source was used in the experiment described in Sec. 2.9. What is the original frequency and the change in frequency of a 14.4-keV gamma-ray photon after it has fallen 20 m near the earth's surface?

54. Find the Schwarzschild radius of the earth, whose mass is 5.98×10^{24} kg.

55. The gravitational potential energy U relative to infinity of a body of mass m at a distance R from the center of a body of mass M is $U = -GmM/R$. (*a*) If R is the radius of the body of mass M, find the escape speed v_e of the body, which is the minimum speed needed to leave it permanently. (*b*) Obtain a formula for the Schwarzschild radius of the body by setting $v_e = c$, the speed of light, and solving for R. (Of course, a relativistic calculation is correct here, but it is interesting to see what a classical calculation produces.)

CHAPTER 3

Wave Properties of Particles



In a scanning electron microscope, an electron beam that scans a specimen causes secondary electrons to be ejected in numbers that vary with the angle of the surface. A suitable data display suggests the three-dimensional form of the specimen. The high resolution of this image of a red spider mite on a leaf is a consequence of the wave nature of moving electrons.

3.1 DE BROGLIE WAVES
A moving body behaves in certain ways as though it has a wave nature

3.2 WAVES OF WHAT?
Waves of probability

3.3 DESCRIBING A WAVE
A general formula for waves

3.4 PHASE AND GROUP VELOCITIES
A group of waves need not have the same velocity as the waves themselves

3.5 PARTICLE DIFFRACTION
An experiment that confirms the existence of de Broglie waves

3.6 PARTICLE IN A BOX
Why the energy of a trapped particle is quantized

3.7 UNCERTAINTY PRINCIPLE I
We cannot know the future because we cannot know the present

3.8 UNCERTAINTY PRINCIPLE II
A particle approach gives the same result

3.9 APPLYING THE UNCERTAINTY PRINCIPLE
A useful tool, not just a negative statement

Looking back, it may seem odd that two decades passed between the 1905 discovery of the particle properties of waves and the 1924 speculation that particles might show wave behavior. It is one thing, however, to suggest a revolutionary concept to explain otherwise mysterious data and quite another to suggest an equally revolutionary concept without a strong experimental mandate. The latter is just what Louis de Broglie did in 1924 when he proposed that moving objects have wave as well as particle characteristics. So different was the scientific climate at the time from that around the turn of the century that de Broglie's ideas soon received respectful attention, whereas the earlier quantum theory of light of Planck and Einstein had been largely ignored despite its striking empirical support. The existence of de Broglie waves was experimentally demonstrated by 1927, and the duality principle they represent provided the starting point for Schrödinger's successful development of quantum mechanics in the previous year.

3.1 DE BROGLIE WAVES

A moving body behaves in certain ways as though it has a wave nature

A photon of light of frequency ν has the momentum

$$p = \frac{h\nu}{c} = \frac{h}{\lambda}$$

since $\lambda\nu = c$. The wavelength of a photon is therefore specified by its momentum according to the relation

Photon wavelength
$$\lambda = \frac{h}{p} \tag{3.1}$$

De Broglie suggested that Eq. (3.1) is a completely general one that applies to material particles as well as to photons. The momentum of a particle of mass m and velocity v is $p = \gamma mv$, and its **de Broglie wavelength** is accordingly

De Broglie wavelength
$$\lambda = \frac{h}{\gamma mv} \tag{3.2}$$

Louis de Broglie (1892–1987), although coming from a French family long identified with diplomacy and the military and initially a student of history, eventually followed his older brother Maurice in a career in physics. His doctoral thesis in 1924 contained the proposal that moving bodies have wave properties that complement their particle properties: these "seemingly incompatible conceptions can each represent an aspect of the truth. . . . They may serve in turn to represent the facts without ever entering into direct conflict." Part of de Broglie's inspiration came from Bohr's theory of the hydrogen atom, in which the electron is supposed to follow only certain orbits around the nucleus. "This fact suggested to me the idea that electrons . . . could not be considered simply as particles but that periodicity must be assigned to them also." Two years later Erwin Schrödinger used the concept of de Broglie waves to develop a general theory that he and others applied to explain a wide variety of atomic phenomena. The existence of de Broglie waves was confirmed in diffraction experiments with electron beams in 1927, and in 1929 de Broglie received the Nobel Prize.

The greater the particle's momentum, the shorter its wavelength. In Eq. (3.2) γ is the relativistic factor

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

As in the case of em waves, the wave and particle aspects of moving bodies can never be observed at the same time. We therefore cannot ask which is the "correct" description. All that can be said is that in certain situations a moving body resembles a wave and in others it resembles a particle. Which set of properties is most conspicuous depends on how its de Broglie wavelength compares with its dimensions and the dimensions of whatever it interacts with.

Example 3.1

Find the de Broglie wavelengths of (*a*) a 46-g golf ball with a velocity of 30 m/s, and (*b*) an electron with a velocity of 10^7 m/s.

Solution

(*a*) Since $v \ll c$, we can let $\gamma = 1$. Hence

$$\lambda = \frac{h}{mv} = \frac{6.63 \times 10^{-34}\ \text{J}\cdot\text{s}}{(0.046\ \text{kg})(30\ \text{m/s})} = 4.8 \times 10^{-34}\ \text{m}$$

The wavelength of the golf ball is so small compared with its dimensions that we would not expect to find any wave aspects in its behavior.

(*b*) Again $v \ll c$, so with $m = 9.1 \times 10^{-31}$ kg, we have

$$\lambda = \frac{h}{mv} = \frac{6.63 \times 10^{-34}\ \text{J}\cdot\text{s}}{(9.1 \times 10^{-31}\ \text{kg})(10^7\ \text{m/s})} = 7.3 \times 10^{-11}\ \text{m}$$

The dimensions of atoms are comparable with this figure—the radius of the hydrogen atom, for instance, is 5.3×10^{-11} m. It is therefore not surprising that the wave character of moving electrons is the key to understanding atomic structure and behavior.

Example 3.2

Find the kinetic energy of a proton whose de Broglie wavelength is 1.000 fm = 1.000×10^{-15} m, which is roughly the proton diameter.

Solution

A relativistic calculation is needed unless pc for the proton is much smaller than the proton rest energy of $E_0 = 0.938$ GeV. To find out, we use Eq. (3.2) to determine pc:

$$pc = (\gamma mv)c = \frac{hc}{\lambda} = \frac{(4.136 \times 10^{-15}\ \text{eV}\cdot\text{s})(2.998 \times 10^8\ \text{m/s})}{1.000 \times 10^{-15}\ \text{m}} = 1.240 \times 10^9\ \text{eV}$$

$$= 1.2410\ \text{GeV}$$

Since $pc > E_0$ a relativistic calculation is required. From Eq. (1.24) the total energy of the proton is

$$E = \sqrt{E_0^2 + p^2c^2} = \sqrt{(0.938\ \text{GeV})^2 + (1.2340\ \text{GeV})^2} = 1.555\ \text{GeV}$$

The corresponding kinetic energy is

$$\text{KE} = E - E_0 = (1.555 - 0.938)\ \text{GeV} = 0.617\ \text{GeV} = 617\ \text{MeV}$$

De Broglie had no direct experimental evidence to support his conjecture. However, he was able to show that it accounted in a natural way for the energy quantization—the restriction to certain specific energy values—that Bohr had had to postulate in his 1913 model of the hydrogen atom. (This model is discussed in Chap. 4.) Within a few years Eq. (3.2) was verified by experiments involving the diffraction of electrons by crystals. Before we consider one of these experiments, let us look into the question of what kind of wave phenomenon is involved in the matter waves of de Broglie.

3.2 WAVES OF WHAT?

Waves of probability

In water waves, the quantity that varies periodically is the height of the water surface. In sound waves, it is pressure. In light waves, electric and magnetic fields vary. What is it that varies in the case of matter waves?

The quantity whose variations make up matter waves is called the **wave function**, symbol Ψ (the Greek letter psi). The value of the wave function associated with a moving body at the particular point x, y, z in space at the time t is related to the likelihood of finding the body there at the time.

Max Born (1882–1970) grew up in Breslau, then a German city but today part of Poland, and received a doctorate in applied mathematics at Göttingen in 1907. Soon afterward he decided to concentrate on physics, and was back in Göttingen in 1909 as a lecturer. There he worked on various aspects of the theory of crystal lattices, his "central interest" to which he often returned in later years. In 1915, at Planck's recommendation, Born became professor of physics in Berlin where, among his other activities, he played piano to Einstein's violin. After army service in World War I and a period at Frankfurt University, Born was again in Göttingen, now as professor of physics. There a remarkable center of theoretical physics developed under his leadership: Heisenberg and Pauli were among his assistants and Fermi, Dirac, Wigner, and Goeppert were among those who worked with him, just to name future Nobel Prize winners. In those days, Born wrote, "There was complete freedom of teaching and learning in German universities, with no class examinations, and no control of students. The University just offered lectures and the student had to decide for himself which he wished to attend."

Born was a pioneer in going from "the bright realm of classical physics into the still dark and unexplored underworld of the new quantum mechanics;" he was the first to use the latter term. From Born came the basic concept that the wave function Ψ of a particle is related to the probability of finding it. He began with an idea of Einstein, who "sought to make the duality of particles (light quanta or photons) and waves comprehensible by interpreting the square of the optical wave amplitude as probability density for the occurrence of photons. This idea could at once be extended to the Ψ-function: $|\Psi|^2$ must represent the probability density for electrons (or other particles). To assert this was easy; but how was it to be proved? For this purpose atomic scattering processes suggested themselves." Born's development of the quantum theory of atomic scattering (collisions of atoms with various particles) not only verified his "new way of thinking about the phenomena of nature" but also founded an important branch of theoretical physics.

Born left Germany in 1933 at the start of the Nazi period, like so many other scientists. He became a British subject and was associated with Cambridge and then Edinburg universities until he retired in 1953. Finding the Scottish climate harsh and wishing to contribute to the democratization of postwar Germany, Born spent the rest of his life in Bad Pyrmont, a town near Göttingen. His textbooks on modern physics and on optics were standard works on these subjects for many years.

The wave function Ψ itself, however, has no direct physical significance. There is a simple reason why Ψ cannot by interpreted in terms of an experiment. The probability that something be in a certain place at a given time must lie between 0 (the object is definitely not there) and 1 (the object is definitely there). An intermediate probability, say 0.2, means that there is a 20% chance of finding the object. But the amplitude of a wave can be negative as well as positive, and a negative probability, say -0.2, is meaningless. Hence Ψ by itself cannot be an observable quantity.

This objection does not apply to $|\Psi|^2$, the square of the absolute value of the wave function, which is known as **probability density:**

> The probability of experimentally finding the body described by the wave function Ψ at the point x, y, z, at the time t is proportional to the value of $|\Psi|^2$ there at t.

A large value of $|\Psi|^2$ means the strong possibility of the body's presence, while a small value of $|\Psi|^2$ means the slight possibility of its presence. As long as $|\Psi|^2$ is not actually 0 somewhere, however, there is a definite chance, however small, of detecting it there. This interpretation was first made by Max Born in 1926.

There is a big difference between the probability of an event and the event itself. Although we can speak of the wave function Ψ that describes a particle as being spread out in space, this does not mean that the particle itself is thus spread out. When an experiment is performed to detect electrons, for instance, a whole electron is either found at a certain time and place or it is not; there is no such thing as a 20 percent of an electron. However, it is entirely possible for there to be a 20 percent chance that the electron be found at that time and place, and it is this likelihood that is specified by $|\Psi|^2$.

W. L. Bragg, the pioneer in x-ray diffraction, gave this loose but vivid interpretation: "The dividing line between the wave and particle nature of matter and radiation is the moment 'now.' As this moment steadily advances through time it coagulates a wavy future into a particle past. . . . Everything in the future is a wave, everything in the past is a particle." If "the moment 'now' " is understood to be the time a measurement is performed, this is a reasonable way to think about the situation. (The philosopher Søren Kierkegaard may have been anticipating this aspect of modern physics when he wrote, "Life can only be understood backwards, but it must be lived forwards.")

Alternatively, if an experiment involves a great many identical objects all described by the same wave function Ψ, the *actual density* (number per unit volume) of objects at x, y, z at the time t is proportional to the corresponding value of $|\Psi|^2$. It is instructive to compare the connection between Ψ and the density of particles it describes with the connection discussed in Sec. 2.4 between the electric field E of an electromagnetic wave and the density N of photons associated with the wave.

While the wavelength of the de Broglie waves associated with a moving body is given by the simple formula $\lambda = h/\gamma mv$, to find their amplitude Ψ as a function of position and time is often difficult. How to calculate Ψ is discussed in Chap. 5 and the ideas developed there are applied to the structure of the atom in Chap. 6. Until then we can assume that we know as much about Ψ as each situation requires.

3.3 DESCRIBING A WAVE

A general formula for waves

How fast do de Broglie waves travel? Since we associate a de Broglie wave with a moving body, we expect that this wave has the same velocity as that of the body. Let us see if this is true.

If we call the de Broglie wave velocity v_p, we can apply the usual formula

$$v_p = \nu\lambda$$

to find v_p. The wavelength λ is simply the de Broglie wavelength $\lambda = h/\gamma mv$. To find the frequency, we equate the quantum expression $E = h\nu$ with the relativistic formula for total energy $E = \gamma mc^2$ to obtain

$$h\nu = \gamma mc^2$$

$$\nu = \frac{\gamma mc^2}{h}$$

The de Broglie wave velocity is therefore

De Broglie phase velocity

$$v_p = \nu\lambda = \left(\frac{\gamma mc^2}{h}\right)\left(\frac{h}{\gamma mv}\right) = \frac{c^2}{v} \tag{3.3}$$

Because the particle velocity v must be less than the velocity of light c, the de Broglie waves always travel faster than light! In order to understand this unexpected result, we must look into the distinction between **phase velocity** and **group velocity**. (Phase velocity is what we have been calling wave velocity.)

Let us begin by reviewing how waves are described mathematically. For simplicity we consider a string stretched along the x axis whose vibrations are in the y direction, as in Fig. 3.1, and are simple harmonic in character. If we choose $t = 0$ when the displacement y of the string at $x = 0$ is a maximum, its displacement at any future time t at the same place is given by the formula

$$y = A \cos 2\pi\nu t \tag{3.4}$$

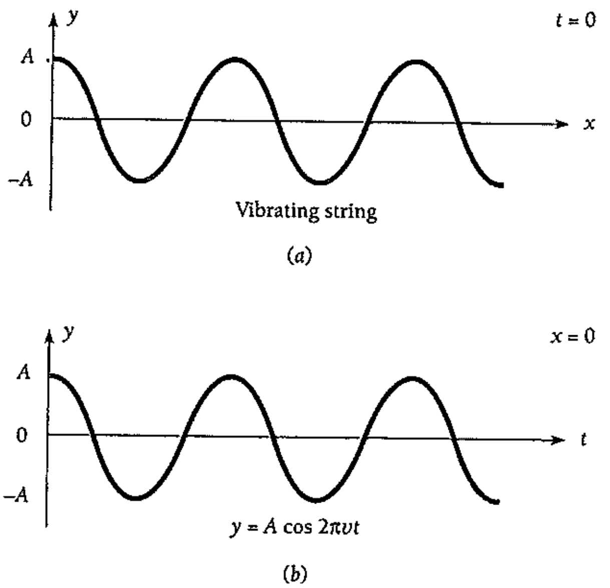


Figure 3.1 (*a*) The appearance of a wave in a stretched string at a certain time. (*b*) How the displacement of a point on the string varies with time.

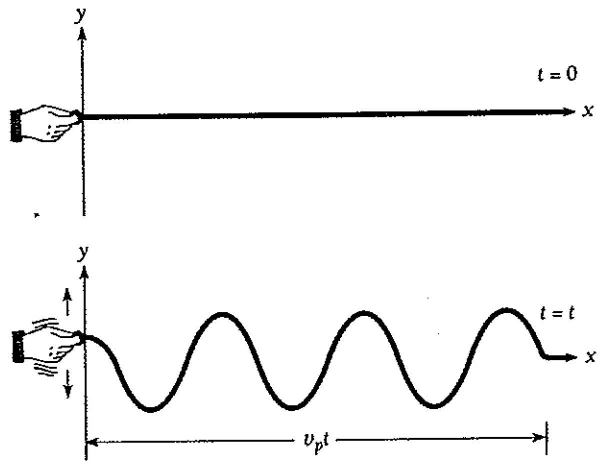


Figure 3.2 Wave propagation.

where A is the amplitude of the vibrations (that is, their maximum displacement on either side of the x axis) and ν their frequency.

Equation (3.4) tells us what the displacement of a single point on the string is as a function of time t. A complete description of wave motion in a stretched string, however, should tell us what y is at *any* point on the string at *any* time. What we want is a formula giving y as a function of both x and t.

To obtain such a formula, let us imagine that we shake the string at $x = 0$ when $t = 0$, so that a wave starts to travel down the string in the $+x$ direction (Fig. 3.2). This wave has some speed v_p that depends on the properties of the string. The wave travels the distance $x = v_pt$ in the time t, so the time interval between the formation of the wave at $x = 0$ and its arrival at the point x is x/v_p. Hence the displacement y of the string at x at any time t is exactly the same as the value of y at $x = 0$ *at the earlier time* $t - x/v_p$. By simply replacing t in Eq. (3.4) with $t - x/v_p$, then, we have the desired formula giving y in terms of both x and t:

Wave formula
$$y = A \cos 2\pi\nu\left(t - \frac{x}{v_p}\right) \tag{3.5}$$

As a check, we note that Eq. (3.5) reduces to Eq. (3.4) at $x = 0$.

Equation (3.5) may be rewritten

$$y = A \cos 2\pi\left(\nu t - \frac{\nu x}{v_p}\right)$$

Since the wave speed v_p is given by $v_p = \nu\lambda$ we have

Wave formula
$$y = A \cos 2\pi\left(\nu t - \frac{x}{\lambda}\right) \tag{3.6}$$

Equation (3.6) is often more convenient to use than Eq. (3.5).

Perhaps the most widely used description of a wave, however, is still another form of Eq. (3.5). The quantities **angular frequency** ω and **wave number** k are defined by the formulas

Angular frequency $$\omega = 2\pi\nu \tag{3.7}$$

Wave number $$k = \frac{2\pi}{\lambda} = \frac{\omega}{v_p} \tag{3.8}$$

The unit of ω is the radian per second and that of k is the radian per meter. Angular frequency gets its name from uniform circular motion, where a particle that moves around a circle ν times per second sweeps out $2\pi\nu$ rad/s. The wave number is equal to the number of radians corresponding to a wave train 1 m long, since there are 2π rad in one complete wave.

In terms of ω and k, Eq. (3.5) becomes

Wave formula $$y = A \cos (\omega t - kx) \tag{3.9}$$

In three dimensions k becomes a vector **k** normal to the wave fronts and x is replaced by the radius vector **r**. The scalar product $\mathbf{k} \cdot \mathbf{r}$ is then used instead of kx in Eq. (3.9).

3.4 PHASE AND GROUP VELOCITIES

A group of waves need not have the same velocity as the waves themselves

The amplitude of the de Broglie waves that correspond to a moving body reflects the probability that it will be found at a particular place at a particular time. It is clear that de Broglie waves cannot be represented simply by a formula resembling Eq. (3.9), which describes an indefinite series of waves all with the same amplitude A. Instead, we expect the wave representation of a moving body to correspond to a **wave packet**, or **wave group**, like that shown in Fig. 3.3, whose waves have amplitudes upon which the likelihood of detecting the body depends.

A familiar example of how wave groups come into being is the case of **beats**. When two sound waves of the same amplitude but of slightly different frequencies are produced simultaneously, the sound we hear has a frequency equal to the average of the two original frequencies and its amplitude rises and falls periodically. The amplitude fluctuations occur as many times per second as the difference between the two original frequencies. If the original sounds have frequencies of, say, 440 and 442 Hz, we will hear a fluctuating sound of frequency 441 Hz with two loudness peaks, called beats, per second. The production of beats is illustrated in Fig. 3.4.

A way to mathematically describe a wave group, then, is in terms of a superposition of individual waves of different wavelengths whose interference with one another results in the variation in amplitude that defines the group shape. If the velocities of the waves are the same, the velocity with which the wave group travels is the common phase velocity. However, if the phase velocity varies with wavelength, the different individual waves do not proceed together. This situation is called **dispersion**. As a result the wave group has a velocity different from the phase velocities of the waves that make it up. This is the case with de Broglie waves.

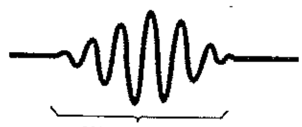


Figure 3.3 A wave group.

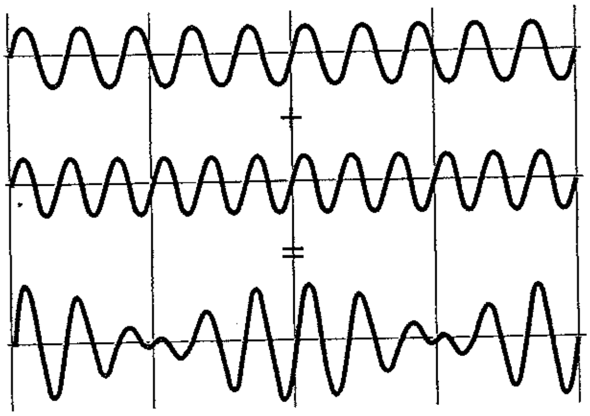


Figure 3.4 Beats are produced by the superposition of two waves with different frequencies.

It is not hard to find the velocity v_g with which a wave group travels. Let us suppose that the wave group arises from the combination of two waves that have the same amplitude A but differ by an amount $\Delta\omega$ in angular frequency and an amount Δk in wave number. We may represent the original waves by the formulas

$$y_1 = A\cos(\omega t - kx)$$
$$y_2 = A\cos[(\omega + \Delta\omega)t - (k + \Delta k)x]$$

The resultant displacement y at any time t and any position x is the sum of y_1 and y_2. With the help of the identity

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

and the relation

$$\cos(-\theta) = \cos\theta$$

we find that

$$\begin{aligned} y &= y_1 + y_2 \\ &= 2A\cos\tfrac{1}{2}[(2\omega + \Delta\omega)t - (2k + \Delta k)x]\cos\tfrac{1}{2}(\Delta\omega\, t - \Delta k\, x) \end{aligned}$$

Since $\Delta\omega$ and Δk are small compared with ω and k respectively,

$$2\omega + \Delta\omega \approx 2\omega$$
$$2k + \Delta k \approx 2k$$

and so

Beats $$y = 2A\cos(\omega t - kx)\cos\left(\frac{\Delta\omega}{2}t - \frac{\Delta k}{2}x\right) \tag{3.10}$$

Equation (3.10) represents a wave of angular frequency ω and wave number k that has superimposed upon it a modulation of angular frequency $\frac{1}{2}\Delta\omega$ and of wave number $\frac{1}{2}\Delta k$.

The effect of the modulation is thus to produce successive wave groups, as in Fig. 3.4. The phase velocity v_p is

Phase velocity

$$v_p = \frac{\omega}{k} \tag{3.11}$$

and the velocity v_g of the wave groups is

Group velocity

$$v_g = \frac{\Delta\omega}{\Delta k} \tag{3.12}$$

When ω and k have continuous spreads instead of the two values in the preceding discussion, the group velocity is instead given by

Group velocity

$$v_g = \frac{d\omega}{dk} \tag{3.13}$$

Depending on how phase velocity varies with wave number in a particular situation, the group velocity may be less or greater than the phase velocities of its member waves. If the phase velocity is the same for all wavelengths, as is true for light waves in empty space, the group and phase velocities are the same.

The angular frequency and wave number of the de Broglie waves associated with a body of mass m moving with the velocity v are

Angular frequency of de Broglie waves

$$\omega = 2\pi\nu = \frac{2\pi\gamma mc^2}{h} = \frac{2\pi mc^2}{h\sqrt{1 - v^2/c^2}} \tag{3.14}$$

Wave number of de Broglie waves

$$k = \frac{2\pi}{\lambda} = \frac{2\pi\gamma mv}{h} = \frac{2\pi mv}{h\sqrt{1 - v^2/c^2}} \tag{3.15}$$

Both ω and k are functions of the body's velocity v.

The group velocity v_g of the de Broglie waves associated with the body is

$$v_g = \frac{d\omega}{dk} = \frac{d\omega/dv}{dk/dv}$$

Now

$$\frac{d\omega}{dv} = \frac{2\pi mv}{h(1 - v^2/c^2)^{3/2}}$$

$$\frac{dk}{dv} = \frac{2\pi m}{h(1 - v^2/c^2)^{3/2}}$$

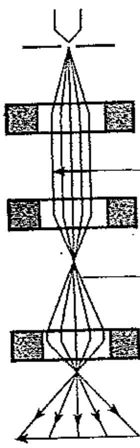


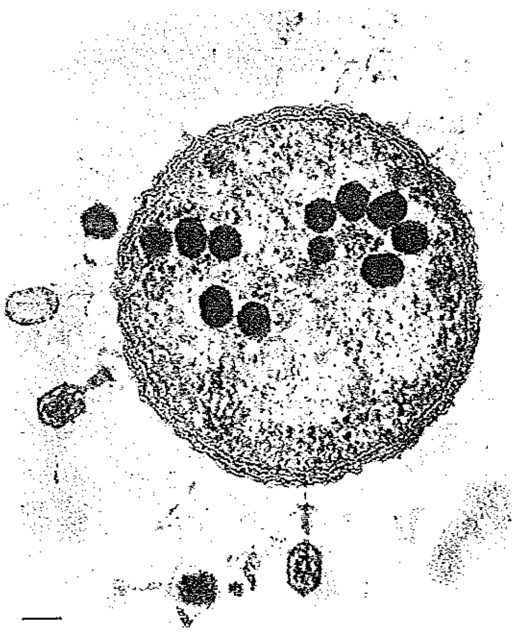
Figure 3.5 Because the wavelengths of the fast electrons in an electron microscope are shorter than those of the light waves in an optical microscope, the electron microscope can produce sharp images at higher magnifications. The electron beam in an electron microscope is focused by magnetic fields.

Electron Microscopes

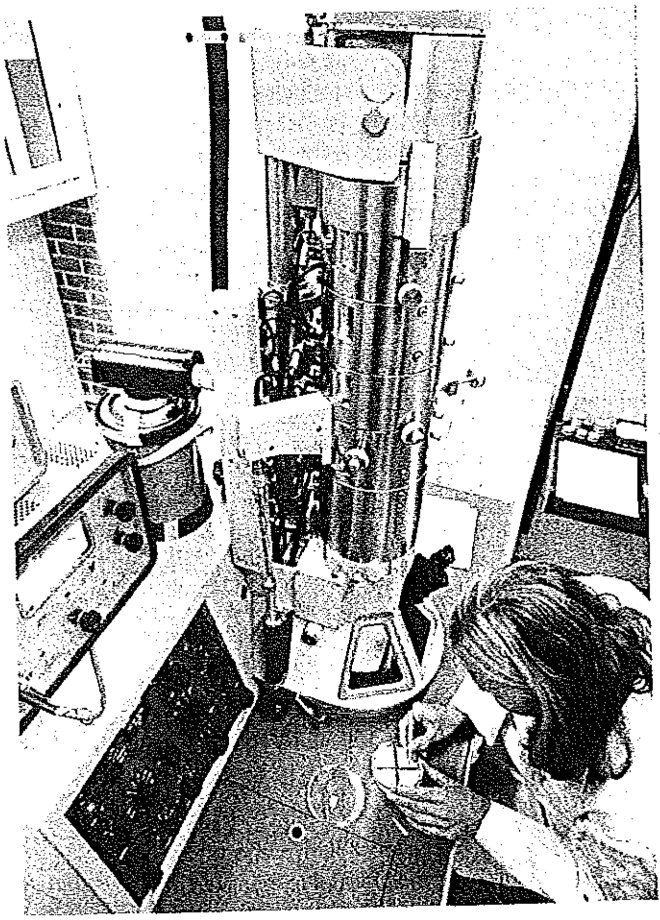
The wave nature of moving electrons is the basis of the electron microscope, the first of which was built in 1932. The resolving power of any optical instrument, which is limited by diffraction, is proportional to the wavelength of whatever is used to illuminate the specimen. In the case of a good microscope that uses visible light, the maximum useful magnification is about 500×; higher magnifications give larger images but do not reveal any more detail. Fast electrons, however, have wavelengths very much shorter than those of visible light and are easily controlled by electric and magnetic fields because of their charge. X-rays also have short wavelengths, but it is not (yet?) possible to focus them adequately.

In an electron microscope, current-carrying coils produce magnetic fields that act as lenses to focus an electron beam on a specimen and then produce an enlarged image on a fluorescent screen or photographic plate (Fig. 3.5). To prevent the beam from being scattered and thereby blurring the image, a thin specimen is used and the entire system is evacuated.

The technology of magnetic "lenses" does not permit the full theoretical resolution of electron waves to be realized in practice. For instance, 100-keV electrons have wavelengths of 0.0037 nm, but the actual resolution they can provide in an electron microscope may be only about 0.1 nm. However, this is still a great improvement on the ~200-nm resolution of an optical microscope, and magnifications of over 1,000,000× have been achieved with electron microscopes.



Electron micrograph showing bacteriophage viruses in an *Escherichia coli* bacterium. The bacterium is approximately 1 μm across.



An electron microscope.

and so the group velocity turns out to be

De Broglie group velocity

$$v_g = v \tag{3.16}$$

The de Broglie wave group associated with a moving body travels with the same velocity as the body.

The phase velocity v_p of de Broglie waves is, as we found earlier,

De Broglie phase velocity

$$v_p = \frac{\omega}{k} = \frac{c^2}{v} \tag{3.3}$$

This exceeds both the velocity of the body v and the velocity of light c, since $v < c$. However, v_p has no physical significance because the motion of the wave group, not the motion of the individual waves that make up the group, corresponds to the motion of the body, and $v_g < c$ as it should be. The fact that $v_p > c$ for de Broglie waves therefore does not violate special relativity.

Example 3.3

An electron has a de Broglie wavelength of 2.00 pm = 2.00×10^{-12} m. Find its kinetic energy and the phase and group velocities of its de Broglie waves.

Solution

(*a*) The first step is to calculate pc for the electron, which is

$$pc = \frac{hc}{\lambda} = \frac{(4.136 \times 10^{-15}\ \text{eV}\cdot\text{s})(3.00 \times 10^{8}\ \text{m/s})}{2.00 \times 10^{-12}\ \text{m}} = 6.20 \times 10^{5}\ \text{eV}$$

$$= 620\ \text{keV}$$

The rest energy of the electron is $E_0 = 511$ keV, so

$$\text{KE} = E - E_0 = \sqrt{E_0^2 + (pc)^2} - E_0 = \sqrt{(511\ \text{keV})^2 + (620\ \text{keV})^2} - 511\ \text{keV}$$

$$= 803\ \text{keV} - 511\ \text{keV} = 292\ \text{keV}$$

(*b*) The electron velocity can be found from

$$E = \frac{E_0}{\sqrt{1 - v^2/c^2}}$$

to be

$$v = c\sqrt{1 - \frac{E_0^2}{E^2}} = c\sqrt{1 - \left(\frac{511\ \text{keV}}{803\ \text{keV}}\right)^2} = 0.771c$$

Hence the phase and group velocities are respectively

$$v_p = \frac{c^2}{v} = \frac{c^2}{0.771c} = 1.30c$$

$$v_g = v = 0.771c$$

3.5 PARTICLE DIFFRACTION

An experiment that confirms the existence of de Broglie waves

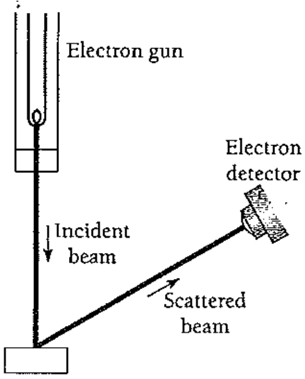


Figure 3.6 The Davisson-Germer experiment.

A wave effect with no analog in the behavior of Newtonian particles is diffraction. In 1927 Clinton Davisson and Lester Germer in the United States and G. P. Thomson in England independently confirmed de Broglie's hypothesis by demonstrating that electron beams are diffracted when they are scattered by the regular atomic arrays of crystals. (All three received Nobel Prizes for their work. J. J. Thomson, G. P.'s father, had earlier won a Nobel Prize for verifying the particle nature of the electron: the wave-particle duality seems to have been the family business.) We shall look at the experiment of Davisson and Germer because its interpretation is more direct.

Davisson and Germer were studying the scattering of electrons from a solid using an apparatus like that sketched in Fig. 3.6. The energy of the electrons in the primary beam, the angle at which they reach the target, and the position of the detector could all be varied. Classical physics predicts that the scattered electrons will emerge in all directions with only a moderate dependence of their intensity on scattering angle and even less on the energy of the primary electrons. Using a block of nickel as the target, Davisson and Germer verified these predictions.

In the midst of their work an accident occurred that allowed air to enter their apparatus and oxidize the metal surface. To reduce the oxide to pure nickel, the target was baked in a hot oven. After this treatment, the target was returned to the apparatus and the measurements resumed.

Now the results were very different. Instead of a continuous variation of scattered electron intensity with angle, distinct maxima and minima were observed whose positions depended upon the electron energy! Typical polar graphs of electron intensity after the accident are shown in Fig. 3.7. The method of plotting is such that the intensity at any angle is proportional to the distance of the curve at that angle from the point of scattering. If the intensity were the same at all scattering angles, the curves would be circles centered on the point of scattering.

Two questions come to mind immediately: What is the reason for this new effect? Why did it not appear until after the nickel target was baked?

De Broglie's hypothesis suggested that electron waves were being diffracted by the target, much as x-rays are diffracted by planes of atoms in a crystal. This idea received

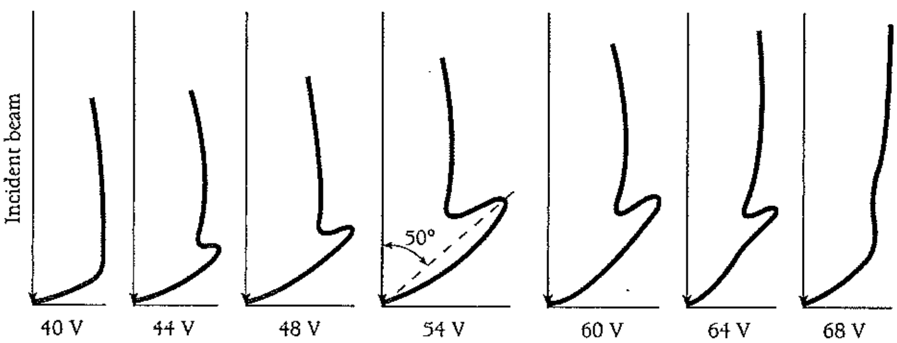


Figure 3.7 Results of the Davisson-Germer experiment, showing how the number of scattered electrons varied with the angle between the incoming beam and the crystal surface. The Bragg planes of atoms in the crystal were not parallel to the crystal surface, so the angles of incidence and scattering relative to one family of these planes were both 65° (see Fig. 3.8).

support when it was realized that heating a block of nickel at high temperature causes the many small individual crystals of which it is normally composed to form into a single large crystal, all of whose atoms are arranged in a regular lattice.

Let us see whether we can verify that de Broglie waves are responsible for the findings of Davisson and Germer. In a particular case, a beam of 54-eV electrons was directed perpendicularly at the nickel target and a sharp maximum in the electron distribution occurred at an angle of 50° with the original beam. The angles of incidence and scattering relative to the family of Bragg planes shown in Fig. 3.8 are both 65°. The spacing of the planes in this family, which can be measured by x-ray diffraction, is 0.091 nm. The Bragg equation for maxima in the diffraction pattern is

$$n\lambda = 2d \sin \theta \qquad (2.13)$$

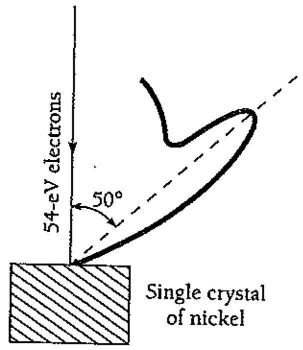


Figure 3.8 The diffraction of the de Broglie waves by the target is responsible for the results of Davisson and Germer.

Here $d = 0.091$ nm and $\theta = 65°$. For $n = 1$ the de Broglie wavelength λ of the diffracted electrons is

$$\lambda = 2d \sin \theta = (2)(0.091 \text{ nm})(\sin 65°) = 0.165 \text{ nm}$$

Now we use de Broglie's formula $\lambda = h/\gamma mv$ to find the expected wavelength of the electrons. The electron kinetic energy of 54 eV is small compared with its rest energy mc^2 of 0.51 MeV, so we can let $\gamma = 1$. Since

$$\text{KE} = \tfrac{1}{2}mv^2$$

the electron momentum mv is

$$\begin{aligned} mv &= \sqrt{2m\text{KE}} \\ &= \sqrt{(2)(9.1 \times 10^{-31} \text{ kg})(54 \text{ eV})(1.6 \times 10^{-19} \text{ J/eV})} \\ &= 4.0 \times 10^{-24} \text{ kg} \cdot \text{m/s} \end{aligned}$$

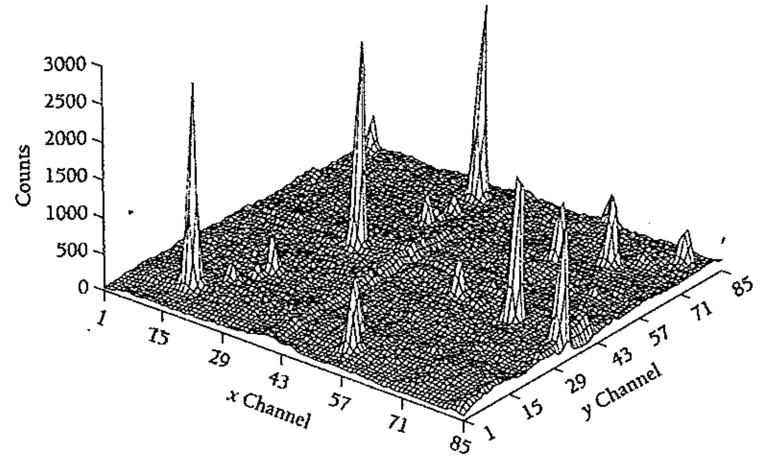
The electron wavelength is therefore

$$\lambda = \frac{h}{mv} = \frac{6.63 \times 10^{-34} \text{ J} \cdot \text{s}}{4.0 \times 10^{-24} \text{ kg} \cdot \text{m/s}} = 1.66 \times 10^{-10} \text{ m} = 0.166 \text{ nm}$$

which agrees well with the observed wavelength of 0.165 nm. The Davisson-Germer experiment thus directly verifies de Broglie's hypothesis of the wave nature of moving bodies.

Analyzing the Davisson-Germer experiment is actually less straightforward than indicated above because the energy of an electron increases when it enters a crystal by an amount equal to the work function of the surface. Hence the electron speeds in the experiment were greater inside the crystal and the de Broglie wavelengths there shorter than the values outside. Another complication arises from interference between waves diffracted by different families of Bragg planes, which restricts the occurrence of maxima to certain combinations of electron energy and angle of incidence rather than merely to any combination that obeys the Bragg equation.

Electrons are not the only bodies whose wave behavior can be demonstrated. The diffraction of neutrons and of whole atoms when scattered by suitable crystals has been observed, and in fact neutron diffraction, like x-ray and electron diffraction, has been used for investigating crystal structures.



Neutron diffraction by a quartz crystal. The peaks represent directions in which constructive interference occurred. *(Courtesy Frank J. Rotella and Arthur J. Schultz, Argonne National Laboratory)*

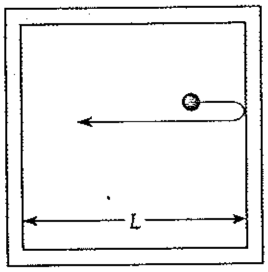


Figure 3.9 A particle confined to a box of width L. The particle is assumed to move back and forth along a straight line between the walls of the box.

3.6 PARTICLE IN A BOX

Why the energy of a trapped particle is quantized

The wave nature of a moving particle leads to some remarkable consequences when the particle is restricted to a certain region of space instead of being able to move freely.

The simplest case is that of a particle that bounces back and forth between the walls of a box, as in Fig. 3.9. We shall assume that the walls of the box are infinitely hard, so the particle does not lose energy each time it strikes a wall, and that its velocity is sufficiently small so that we can ignore relativistic considerations. Simple as it is, this model situation requires fairly elaborate mathematics in order to be properly analyzed, as we shall learn in Chap. 5. However, even a relatively crude treatment can reveal the essential results.

From a wave point of view, a particle trapped in a box is like a standing wave in a string stretched between the box's walls. In both cases the wave variable (transverse displacement for the string, wave function Ψ for the moving particle) must be 0 at the walls, since the waves stop there. The possible de Broglie wavelengths of the particle in the box therefore are determined by the width L of the box, as in Fig. 3.10. The longest wavelength is specified by $\lambda = 2L$, the next by $\lambda = L$, then $\lambda = 2L/3$, and so forth. The general formula for the permitted wavelengths is

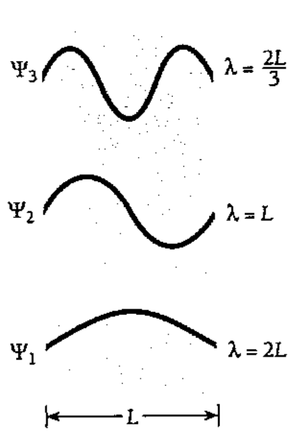


Figure 3.10 Wave functions of a particle trapped in a box L wide.

De Broglie wavelengths of trapped particle

$$\lambda_n = \frac{2L}{n} \qquad n = 1, 2, 3, \ldots \tag{3.17}$$

Because $mv = h/\lambda$, the restrictions on de Broglie wavelength λ imposed by the width of the box are equivalent to limits on the momentum of the particle and, in turn, to limits on its kinetic energy. The kinetic energy of a particle of momentum mv is

$$\text{KE} = \tfrac{1}{2}mv^2 = \frac{(mv)^2}{2m} = \frac{h^2}{2m\lambda^2}$$

The permitted wavelengths are $\lambda_n = 2L/n$, and so, because the particle has no potential energy in this model, the only energies it can have are

Particle in a box $$E_n = \frac{n^2h^2}{8mL^2} \qquad n = 1, 2, 3, \ldots \tag{3.18}$$

Each permitted energy is called an **energy level**, and the integer n that specifies an energy level E_n is called its **quantum number**.

We can draw three general conclusions from Eq. (3.18). These conclusions apply to *any* particle confined to a certain region of space (even if the region does not have a well-defined boundary), for instance an atomic electron held captive by the attraction of the positively charged nucleus.

1 A trapped particle cannot have an arbitrary energy, as a free particle can. The fact of its confinement leads to restrictions on its wave function that allow the particle to have only certain specific energies and no others. Exactly what these energies are depends on the mass of the particle and on the details of how it is trapped.

2 A trapped particle cannot have zero energy. Since the de Broglie wavelength of the particle is $\lambda = h/mv$, a speed of $v = 0$ means an infinite wavelength. But there is no way to reconcile an infinite wavelength with a trapped particle, so such a particle must have at least some kinetic energy. The exclusion of $E = 0$ for a trapped particle, like the limitation of E to a set of discrete values, is a result with no counterpart in classical physics, where all non-negative energies, including zero, are allowed.

3 Because Planck's constant is so small—only 6.63×10^{-34} J · s—quantization of energy is conspicuous only when m and L are also small. This is why we are not aware of energy quantization in our own experience. Two examples will make this clear.

Example 3.4

An electron is in a box 0.10 nm across, which is the order of magnitude of atomic dimensions. Find its permitted energies.

Solution

Here $m = 9.1 \times 10^{-31}$ kg and $L = 0.10$ nm $= 1.0 \times 10^{-10}$ m, so that the permitted electron energies are

$$E_n = \frac{(n^2)(6.63 \times 10^{-34}\ \text{J} \cdot \text{s})^2}{(8)(9.1 \times 10^{-31}\ \text{kg})(1.0 \times 10^{-10}\ \text{m})^2} = 6.0 \times 10^{-18} n^2\ \text{J}$$
$$= 38n^2\ \text{eV}$$

The minimum energy the electron can have is 38 eV, corresponding to $n = 1$. The sequence of energy levels continues with $E_2 = 152$ eV, $E_3 = 342$ eV, $E_4 = 608$ eV, and so on (Fig. 3.11). If such a box existed, the quantization of a trapped electron's energy would be a prominent feature of the system. (And indeed energy quantization is prominent in the case of an atomic electron.)

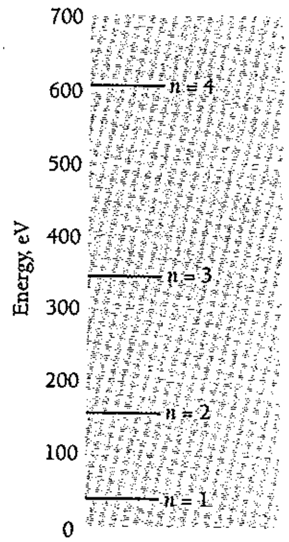


Figure 3.11 Energy levels of an electron confined to a box 0.1 nm wide.

Example 3.5

A 10-g marble is in a box 10 cm across. Find its permitted energies.

Solution

With $m = 10$ g $= 1.0 \times 10^{-2}$ kg and $L = 10$ cm $= 1.0 \times 10^{-1}$ m,

$$E_n = \frac{(n^2)(6.63 \times 10^{-34}\ \text{J} \cdot \text{s})^2}{(8)(1.0 \times 10^{-2}\ \text{kg})(1.0 \times 10^{-1}\ \text{m})^2}$$
$$= 5.5 \times 10^{-64} n^2\ \text{J}$$

The minimum energy the marble can have is 5.5×10^{-64} J, corresponding to $n = 1$. A marble with this kinetic energy has a speed of only 3.3×10^{-31} m/s and therefore cannot be experimentally distinguished from a stationary marble. A reasonable speed a marble might have is, say, $\frac{1}{3}$ m/s—which corresponds to the energy level of quantum number $n = 10^{30}$! The permissible energy levels are so very close together, then, that there is no way to determine whether the marble can take on only those energies predicted by Eq. (3.18) or any energy whatever. Hence in the domain of everyday experience, quantum effects are imperceptible, which accounts for the success of Newtonian mechanics in this domain.

3.7 UNCERTAINTY PRINCIPLE 1

We cannot know the future because we cannot know the present

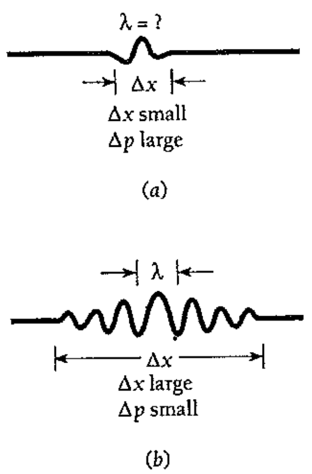


Figure 3.12 (*a*) A narrow de Broglie wave group. The position of the particle can be precisely determined, but the wavelength (and hence the particle's momentum) cannot be established because there are not enough waves to measure accurately. (*b*) A wide wave group. Now the wavelength can be precisely determined but not the position of the particle.

To regard a moving particle as a wave group implies that there are fundamental limits to the accuracy with which we can measure such "particle" properties as position and momentum.

To make clear what is involved, let us look at the wave group of Fig. 3.3. The particle that corresponds to this wave group may be located anywhere within the group at a given time. Of course, the probability density $|\Psi|^2$ is a maximum in the middle of the group, so it is most likely to be found there. Nevertheless, we may still find the particle anywhere that $|\Psi|^2$ is not actually 0.

The narrower its wave group, the more precisely a particle's position can be specified (Fig. 3.12*a*). However, the wavelength of the waves in a narrow packet is not well defined; there are not enough waves to measure λ accurately. This means that since $\lambda = h/\gamma mv$, the particle's momentum γmv is not a precise quantity. If we make a series of momentum measurements, we will find a broad range of values.

On the other hand, a wide wave group, such as that in Fig. 3.12*b*, has a clearly defined wavelength. The momentum that corresponds to this wavelength is therefore a precise quantity, and a series of measurements will give a narrow range of values. But where is the particle located? The width of the group is now too great for us to be able to say exactly where the particle is at a given time.

Thus we have the **uncertainty principle:**

> It is impossible to know both the exact position and exact momentum of an object at the same time.

This principle, which was discovered by Werner Heisenberg in 1927, is one of the most significant of physical laws.

A formal analysis supports the above conclusion and enables us to put it on a quantitative basis. The simplest example of the formation of wave groups is that given in Sec. 3.4, where two wave trains slightly different in angular frequency ω and wave number k were superposed to yield the series of groups shown in Fig. 3.4. A moving body corresponds to a single wave group, not a series of them, but a single wave group can also be thought of in terms of the superposition of trains of harmonic waves. However, an infinite number of wave trains with different frequencies, wave numbers, and amplitudes is required for an isolated group of arbitrary shape, as in Fig. 3.13.

At a certain time t, the wave group $\Psi(x)$ can be represented by the **Fourier integral**

$$\Psi(x) = \int_0^\infty g(k) \cos kx \, dk \tag{3.19}$$

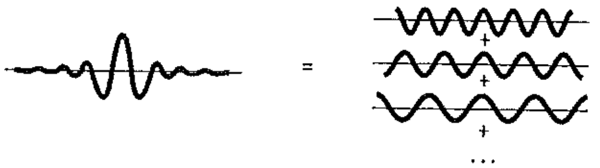


Figure 3.13 An isolated wave group is the result of superposing an infinite number of waves with different wavelengths. The narrower the wave group, the greater the range of wavelengths involved. A narrow de Broglie wave group thus means a well-defined position (Δx smaller) but a poorly defined wavelength and a large uncertainty Δp in the momentum of the particle the group represents. A wide wave group means a more precise momentum but a less precise position.

where the function $g(k)$ describes how the amplitudes of the waves that contribute to $\Psi(x)$ vary with wave number k. This function is called the **Fourier transform** of $\Psi(x)$, and it specifies the wave group just as completely as $\Psi(x)$ does. Figure 3.14 contains graphs of the Fourier transforms of a pulse and of a wave group. For comparison, the Fourier transform of an infinite train of harmonic waves is also included. There is only a single wave number in this case, of course.

Strictly speaking, the wave numbers needed to represent a wave group extend from $k = 0$ to $k = \infty$, but for a group whose length Δx is finite, the waves whose amplitudes $g(k)$ are appreciable have wave numbers that lie within a finite interval Δk. As Fig. 3.14 indicates, the narrower the group, the broader the range of wave numbers needed to describe it, and vice versa.

The relationship between the distance Δx and the wave-number spread Δk depends upon the shape of the wave group and upon how Δx and Δk are defined. The minimum value of the product $\Delta x\, \Delta k$ occurs when the envelope of the group has the familiar bell shape of a Gaussian function. In this case the Fourier transform happens to be a Gaussian function also. If Δx and Δk are taken as the standard deviations of the respective functions $\Psi(x)$ and $g(k)$, then this minimum value is $\Delta x\, \Delta k = \frac{1}{2}$. Because wave groups in general do not have Gaussian forms, it is more realistic to express the relationship between Δx and Δk as

$$\Delta x\, \Delta k \geq \frac{1}{2} \tag{3.20}$$

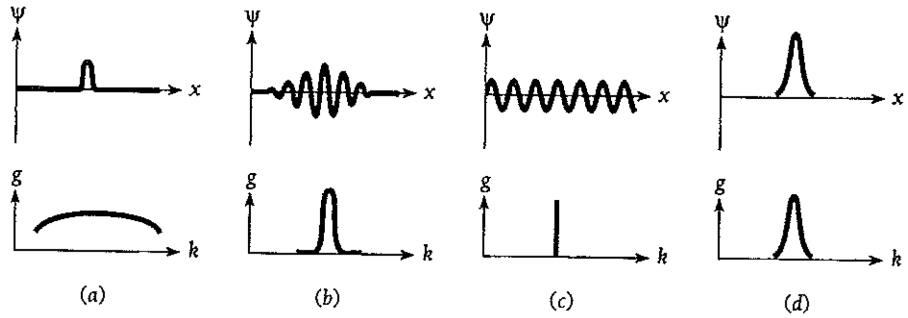


Figure 3.14 The wave functions and Fourier transforms for (*a*) a pulse, (*b*) a wave group, (*c*) a wave train, and (*d*) a Gaussian distribution. A brief disturbance needs a broader range of frequencies to describe it than a disturbance of greater duration. The Fourier transform of a Gaussian function is also a Gaussian function.

Gaussian Function

When a set of measurements is made of some quantity x in which the experimental errors are random, the result is often a **Gaussian distribution** whose form is the bell-shaped curve shown in Fig. 3.15. The **standard deviation** σ of the measurements is a measure of the spread of x values about the mean of x_0, where σ equals the square root of the average of the squared deviations from x_0. If N measurements were made,

Standard deviation $$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - x_0)^2}$$

The width of a Gaussian curve at half its maximum value is 2.35σ.

The *Gaussian function* $f(x)$ that describes the above curve is given by

Gaussian function $$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-x_0)^2/2\sigma^2}$$

where $f(x)$ is the probability that the value x be found in a particular measurement. Gaussian functions occur elsewhere in physics and mathematics as well. (Gabriel Lippmann had this to say about the Gaussian function: "Experimentalists think that it is a mathematical theorem while mathematicians believe it to be an experimental fact.")

The probability that a measurement lie inside a certain range of x values, say between x_1 and x_2, is given by the area of the $f(x)$ curve between these limits. This area is the integral

$$P_{x_1x_2} = \int_{x_1}^{x_2} f(x)\,dx$$

An interesting questions is what fraction of a series of measurements has values within a standard deviation of the mean value x_0. In this case $x_1 = x_0 - \sigma$ and $x_2 = x_0 + \sigma$, and

$$P_{x_0\pm\sigma} = \int_{x_0-\sigma}^{x_0+\sigma} f(x)\,dx = 0.683$$

Hence 68.3 percent of the measurements fall in this interval, which is shaded in Fig. 3.15. A similar calculation shows that 95.4 percent of the measurements fall within two standard deviations of the mean value.

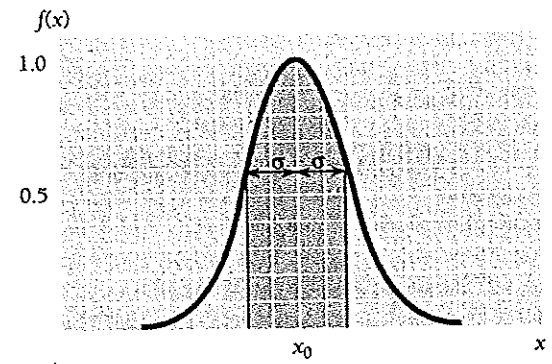


Figure 3.15 A Gaussian distribution. The probability of finding a value of x is given by the Gaussian function $f(x)$. The mean value of x is x_0, and the total width of the curve at half its maximum value is 2.35σ, where σ is the standard deviation of the distribution. The total probability of finding a value of x within a standard deviation of x_0 is equal to the shaded area and is 68.3 percent.

The de Broglie wavelength of a particle of momentum p is $\lambda = h/p$ and the corresponding wave number is

$$k = \frac{2\pi}{\lambda} = \frac{2\pi p}{h}$$

In terms of wave number the particle's momentum is therefore

$$p = \frac{hk}{2\pi}$$

Hence an uncertainty Δk in the wave number of the de Broglie waves associated with the particle results in an uncertainty Δp in the particle's momentum according to the formula

$$\Delta p = \frac{h\,\Delta k}{2\pi}$$

Since $\Delta x\,\Delta k \geq \frac{1}{2}$, $\Delta k \geq 1/(2\Delta x)$ and

Uncertainty principle

$$\Delta x\,\Delta p \geq \frac{h}{4\pi} \tag{3.21}$$

This equation states that the product of the uncertainty Δx in the position of an object at some instant and the uncertainty Δp in its momentum component in the x direction at the same instant is equal to or greater than $h/4\pi$.

If we arrange matters so that Δx is small, corresponding to a narrow wave group, then Δp will be large. If we reduce Δp in some way, a broad wave group is inevitable and Δx will be large.

Werner Heisenberg (1901–1976) was born in Duisberg, Germany, and studied theoretical physics at Munich, where he also became an enthusiastic skier and mountaineer. At Göttingen in 1924 as an assistant to Max Born, Heisenberg became uneasy about mechanical models of the atom: "Any picture of the atom that our imagination is able to invent is for that very reason defective," he later remarked. Instead he conceived an abstract approach using matrix algebra. In 1925, together with Born and Pascual Jordan, Heisenberg developed this approach into a consistent theory of quantum mechanics, but it was so difficult to understand and apply that it had very little impact on physics at the time. Schrödinger's wave formulation of quantum mechanics the following year was much more successful; Schrödinger and others soon showed that the wave and matrix versions of quantum mechanics were mathematically equivalent.

In 1927, working at Bohr's institute in Copenhagen, Heisenberg developed a suggestion by Wolfgang Pauli into the uncertainty principle. Heisenberg initially felt that this principle was a consequence of the disturbances inevitably produced by any measuring process. Bohr, on the other hand, thought that the basic cause of the uncertainties was the wave-particle duality, so that they were built into the natural world rather than solely the result of measurement. After much argument Heisenberg came around to Bohr's view. (Einstein, always skeptical about quantum mechanics, said after a lecture by Heisenberg on the uncertainty principle: "Marvelous, what ideas the young people have these days. But I don't believe a word of it.") Heisenberg received the Nobel Prize in 1932.

Heisenberg was one of the very few distinguished scientists to remain in Germany during the Nazi period. In World War II he led research there on atomic weapons, but little progress had been made by the war's end. Exactly why remains unclear, although there is no evidence that Heisenberg, as he later claimed, had moral qualms about creating such weapons and more or less deliberately dragged his feet. Heisenberg recognized early that "an explosive of unimaginable consequences" could be developed, and he and his group should have been able to have gotten farther than they did. In fact, alarmed by the news that Heisenberg was working on an atomic bomb, the U.S. government sent the former Boston Red Sox catcher Moe Berg to shoot Heisenberg during a lecture in neutral Switzerland in 1944. Berg, sitting in the second row, found himself uncertain from Heisenberg's remarks about how advanced the German program was, and kept his gun in his pocket.

These uncertainties are due not to inadequate apparatus but to the imprecise character in nature of the quantities involved. Any instrumental or statistical uncertainties that arise during a measurement only increase the product $\Delta x\, \Delta p$. Since we cannot know exactly both where a particle is right now and what its momentum is, we cannot say anything definite about where it will be in the future or how fast it will be moving then. *We cannot know the future for sure because we cannot know the present for sure.* But our ignorance is not total: we can still say that the particle is more likely to be in one place than another and that its momentum is more likely to have a certain value than another.

H-Bar

The quantity $h/2\pi$ appears often in modern physics because it turns out to be the basic unit of angular momentum. It is therefore customary to abbreviate $h/2\pi$ by the symbol $\hbar$ ("h-bar"):

$$\hbar = \frac{h}{2\pi} = 1.054 \times 10^{-34}\ \text{J} \cdot \text{s}$$

In the remainder of this book $\hbar$ is used in place of $h/2\pi$. In terms of $\hbar$, the uncertainty principle becomes

Uncertainty principle

$$\Delta x\, \Delta p \geq \frac{\hbar}{2} \tag{3.22}$$

Example 3.6

A measurement establishes the position of a proton with an accuracy of $\pm 1.00 \times 10^{-11}$ m. Find the uncertainty in the proton's position 1.00 s later. Assume $v \ll c$.

Solution

Let us call the uncertainty in the proton's position Δx_0 at the time $t = 0$. The uncertainty in its momentum at this time is therefore, from Eq. (3.22),

$$\Delta p \geq \frac{\hbar}{2\Delta x_0}$$

Since $v \ll c$, the momentum uncertainty is $\Delta p = \Delta(mv) = m\, \Delta v$ and the uncertainty in the proton's velocity is

$$\Delta v = \frac{\Delta p}{m} \geq \frac{\hbar}{2m\, \Delta x_0}$$

The distance x the proton covers in the time t cannot be known more accurately than

$$\Delta x = t\, \Delta v \geq \frac{\hbar t}{2m\, \Delta x_0}$$

Hence Δx is inversely proportional to Δx_0: the *more* we know about the proton's position at $t = 0$, the *less* we know about its later position at $t > 0$. The value of Δx at $t = 1.00$ s is

$$\Delta x \geq \frac{(1.054 \times 10^{-34}\ \text{J} \cdot \text{s})(1.00\ \text{s})}{(2)(1.672 \times 10^{-27}\ \text{kg})(1.00 \times 10^{-11}\ \text{m})}$$

$$\geq 3.15 \times 10^{3}\ \text{m}$$

This is 3.15 km—nearly 2 mi! What has happened is that the original wave group has spread out to a much wider one (Fig. 3.16). This occurred because the phase velocities of the component waves vary with wave number and a large range of wave numbers must have been present to produce the narrow original wave group. See Fig. 3.14.

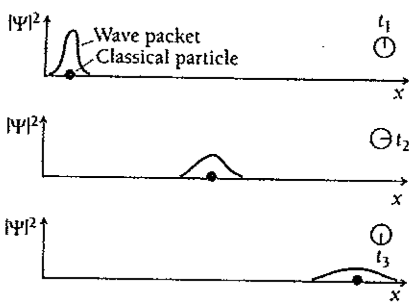


Figure 3.16 The wave packet that corresponds to a moving packet is a composite of many individual waves, as in Fig. 3.13. The phase velocities of the individual waves vary with their wave lengths. As a result, as the particle moves, the wave packet spreads out in space. The narrower the original wavepacket—that is, the more precisely we know its position at that time—the more it spreads out because it is made up of a greater span of waves with different phase velocities.

3.8 UNCERTAINTY PRINCIPLE II

A particle approach gives the same result

The uncertainty principle can be arrived at from the point of view of the particle properties of waves as well as from the point of view of the wave properties of particles.

We might want to measure the position and momentum of an object at a certain moment. To do so, we must touch it with something that will carry the required information back to us. That is, we must poke it with a stick, shine light on it, or perform some similar act. The measurement process itself thus requires that the object be interfered with in some way. If we consider such interferences in detail, we are led to the same uncertainty principle as before even without taking into account the wave nature of moving bodies.

Suppose we look at an electron using light of wavelength λ, as in Fig. 3.17. Each photon of this light has the momentum h/λ. When one of these photons bounces off the electron (which must happen if we are to "see" the electron), the electron's

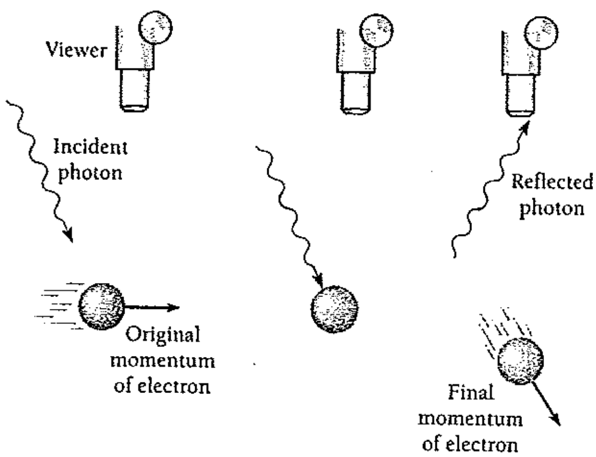


Figure 3.17 An electron cannot be observed without changing its momentum.

original momentum will be changed. The exact amount of the change Δp cannot be predicted, but it will be of the same order of magnitude as the photon momentum h/λ. Hence

$$\Delta p \approx \frac{h}{\lambda} \tag{3.23}$$

The longer the wavelength of the observing photon, the smaller the uncertainty in the electron's momentum.

Because light is a wave phenomenon as well as a particle phenomenon, we cannot expect to determine the electron's location with perfect accuracy regardless of the instrument used. A reasonable estimate of the minimum uncertainty in the measurement might be one photon wavelength, so that

$$\Delta x \geq \lambda \tag{3.24}$$

The shorter the wavelength, the smaller the uncertainty in location. However, if we use light of short wavelength to increase the accuracy of the position measurement, there will be a corresponding decrease in the accuracy of the momentum measurement because the higher photon momentum will disturb the electron's motion to a greater extent. Light of long wavelength will give a more accurate momentum but a less accurate position.

Combining Eqs. (3.23) and (3.24) gives

$$\Delta x \, \Delta p \geq h \tag{3.25}$$

This result is consistent with Eq. (3.22), $\Delta x \, \Delta p \geq \hbar/2$.

Arguments like the preceding one, although superficially attractive, must be approached with caution. The argument above implies that the electron can possess a definite position and momentum at any instant and that it is the measurement process that introduces the indeterminacy in $\Delta x \, \Delta p$. On the contrary, *this indeterminacy is inherent in the nature of a moving body*. The justification for the many "derivations" of this kind is first, they show it is impossible to imagine a way around the uncertainty principle; and second, they present a view of the principle that can be appreciated in a more familiar context than that of wave groups.

3.9 APPLYING THE UNCERTAINTY PRINCIPLE

A useful tool, not just a negative statement

Planck's constant h is so small that the limitations imposed by the uncertainty principle are significant only in the realm of the atom. On such a scale, however, this principle is of great help in understanding many phenomena. It is worth keeping in mind that the lower limit of $\hbar/2$ for $\Delta x \, \Delta p$ is rarely attained. More usually $\Delta x \, \Delta p \geq \hbar$, or even (as we just saw) $\Delta x \, \Delta p \geq h$.

Example 3.7

A typical atomic nucleus is about 5.0×10^{-15} m in radius. Use the uncertainty principle to place a lower limit on the energy an electron must have if it is to be part of a nucleus.

Solution

Letting $\Delta x = 5.0 \times 10^{-5}$ m we have

$$\Delta p \geq \frac{\hbar}{2\Delta x} \geq \frac{1.054 \times 10^{-34}\ \mathrm{J \cdot s}}{(2)(5.0 \times 10^{-15}\ \mathrm{m})} \geq 1.1 \times 10^{-20}\ \mathrm{kg \cdot m/s}$$

If this is the uncertainty in a nuclear electron's momentum, the momentum p itself must be at least comparable in magnitude. An electron with such a momentum has a kinetic energy KE many times greater than its rest energy mc^2. From Eq. (1.24) we see that we can let KE $= pc$ here to a sufficient degree of accuracy. Therefore

$$\mathrm{KE} = pc \geq (1.1 \times 10^{-20}\ \mathrm{kg \cdot m/s})(3.0 \times 10^{8}\ \mathrm{m/s}) \geq 3.3 \times 10^{-12}\ \mathrm{J}$$

Since 1 eV $= 1.6 \times 10^{-19}$ J, the kinetic energy of an electron must exceed 20 MeV if it is to be inside a nucleus. Experiments show that the electrons emitted by certain unstable nuclei never have more than a small fraction of this energy, from which we conclude that nuclei cannot contain electrons. The electron an unstable nucleus may emit comes into being at the moment the nucleus decays (see Secs. 11.3 and 12.5).

Example 3.8

A hydrogen atom is 5.3×10^{-11} m in radius. Use the uncertainty principle to estimate the minimum energy an electron can have in this atom.

Solution

Here we find that with $\Delta x = 5.3 \times 10^{-11}$ m.

$$\Delta p \geq \frac{\hbar}{2\Delta x} \geq 9.9 \times 10^{-25}\ \mathrm{kg \cdot m/s}$$

An electron whose momentum is of this order of magnitude behaves like a classical particle, and its kinetic energy is

$$\mathrm{KE} = \frac{p^2}{2m} \geq \frac{(9.9 \times 10^{-25}\ \mathrm{kg \cdot m/s})^2}{(2)(9.1 \times 10^{-31}\ \mathrm{kg})} \geq 5.4 \times 10^{-19}\ \mathrm{J}$$

which is 3.4 eV. The kinetic energy of an electron in the lowest energy level of a hydrogen atom is actually 13.6 eV.

Energy and Time

Another form of the uncertainty principle concerns energy and time. We might wish to measure the energy E emitted during the time interval Δt in an atomic process. If the energy is in the form of em waves, the limited time available restricts the accuracy with which we can determine the frequency ν of the waves. Let us assume that the minimum uncertainty in the number of waves we count in a wave group is one wave. Since the frequency of the waves under study is equal to the number of them we count divided by the time interval, the uncertainty $\Delta\nu$ in our frequency measurement is

$$\Delta\nu \geq \frac{1}{\Delta t}$$

The corresponding energy uncertainty is

$$\Delta E = h\,\Delta\nu$$

and so

$$\Delta E \geq \frac{h}{\Delta t} \qquad \text{or} \qquad \Delta E\,\Delta t \geq h$$

A more precise calculation based on the nature of wave groups changes this result to

Uncertainties in energy and time

$$\Delta E\,\Delta t \geq \frac{\hbar}{2} \tag{3.26}$$

Equation (3.26) states that the product of the uncertainty ΔE in an energy measurement and the uncertainty Δt in the time at which the measurement is made is equal to or greater than $\hbar/2$. This result can be derived in other ways as well and is a general one not limited to em waves.

Example 3.9

An "excited" atom gives up its excess energy by emitting a photon of characteristic frequency, as described in Chap. 4. The average period that elapses between the excitation of an atom and the time it radiates is 1.0×10^{-8} s. Find the inherent uncertainty in the frequency of the photon.

Solution

The photon energy is uncertain by the amount

$$\Delta E \geq \frac{\hbar}{2\Delta t} \geq \frac{1.054 \times 10^{-34}\ \text{J}\cdot\text{s}}{2(1.0 \times 10^{-8}\ \text{s})} \geq 5.3 \times 10^{-27}\ \text{J}$$

The corresponding uncertainty in the frequency of light is

$$\Delta\nu = \frac{\Delta E}{h} \geq 8 \times 10^{6}\ \text{Hz}$$

This is the irreducible limit to the accuracy with which we can determine the frequency of the radiation emitted by an atom. As a result, the radiation from a group of excited atoms does not appear with the precise frequency ν. For a photon whose frequency is, say, 5.0×10^{14} Hz, $\Delta\nu/\nu = 1.6 \times 10^{-8}$. In practice, other phenomena such as the doppler effect contribute more than this to the broadening of spectral lines.

EXERCISES

It is only the first step that takes the effort. —Marquise du Deffand

3.1 De Broglie Waves

1. A photon and a particle have the same wavelength. Can anything be said about how their linear momenta compare? About how the photon's energy compares with the particle's total energy? About how the photon's energy compares with the particle's kinetic energy?

2. Find the de Broglie wavelength of (*a*) an electron whose speed is 1.0×10^8 m/s, and (*b*) an electron whose speed is 2.0×10^8 m/s.

3. Find the de Broglie wavelength of a 1.0-mg grain of sand blown by the wind at a speed of 20 m/s.

4. Find the de Broglie wavelength of the 40-keV electrons used in a certain electron microscope.

5. By what percentage will a nonrelativistic calculation of the de Broglie wavelength of a 100-keV electron be in error?

6. Find the de Broglie wavelength of a 1.00-MeV proton. Is a relativistic calculation needed?

7. The atomic spacing in rock salt, NaCl, is 0.282 nm. Find the kinetic energy (in eV) of a neutron with a de Broglie wavelength of 0.282 nm. Is a relativistic calculation needed? Such neutrons can be used to study crystal structure.

8. Find the kinetic energy of an electron whose de Broglie wavelength is the same as that of a 100-keV x-ray.

9. Green light has a wavelength of about 550 nm. Through what potential difference must an electron be accelerated to have this wavelength?

10. Show that the de Broglie wavelength of a particle of mass m and kinetic energy KE is given by

$$\lambda = \frac{hc}{\sqrt{\text{KE}(\text{KE} + 2mc^2)}}$$

11. Show that if the total energy of a moving particle greatly exceeds its rest energy, its de Broglie wavelength is nearly the same as the wavelength of a photon with the same total energy.

12. (*a*) Derive a relativistically correct formula that gives the de Broglie wavelength of a charged particle in terms of the potential difference V through which it has been accelerated. (*b*) What is the nonrelativistic approximation of this formula, valid for $eV \ll mc^2$?

3.4 Phase and Group Velocities

13. An electron and a proton have the same velocity. Compare the wavelengths and the phase and group velocities of their de Broglie waves.

14. An electron and a proton have the same kinetic energy. Compare the wavelengths and the phase and group velocities of their de Broglie waves.

15. Verify the statement in the text that, if the phase velocity is the same for all wavelengths of a certain wave phenomenon (that is, there is no dispersion), the group and phase velocities are the same.

16. The phase velocity of ripples on a liquid surface is $\sqrt{2\pi S/\lambda\rho}$, where S is the surface tension and ρ the density of the liquid. Find the group velocity of the ripples.

17. The phase velocity of ocean waves is $\sqrt{g\lambda/2\pi}$, where g is the acceleration of gravity. Find the group velocity of ocean waves.

18. Find the phase and group velocities of the de Broglie waves of an electron whose speed is $0.900c$.

19. Find the phase and group velocities of the de Broglie waves of an electron whose kinetic energy is 500 keV.

20. Show that the group velocity of a wave is given by $v_g = d\nu/d(1/\lambda)$.

21. (*a*) Show that the phase velocity of the de Broglie waves of a particle of mass m and de Broglie wavelength λ is given by

$$v_p = c\sqrt{1 + \left(\frac{mc\lambda}{h}\right)^2}$$

(*b*) Compare the phase and group velocities of an electron whose de Broglie wavelength is exactly 1×10^{-13} m.

22. In his original paper, de Broglie suggested that $E = h\nu$ and $p = h/\lambda$, which hold for electromagnetic waves, are also valid for moving particles. Use these relationships to show that the group velocity v_g of a de Broglie wave group is given by dE/dp, and with the help of Eq. (1.24), verify that $v_g = v$ for a particle of velocity v.

3.5 Particle Diffraction

23. What effect on the scattering angle in the Davisson-Germer experiment does increasing the electron energy have?

24. A beam of neutrons that emerges from a nuclear reactor contains neutrons with a variety of energies. To obtain neutrons with an energy of 0.050 eV, the beam is passed through a crystal whose atomic planes are 0.20 nm apart. At what angles relative to the original beam will the desired neutrons be diffracted?

25. In Sec. 3.5 it was mentioned that the energy of an electron entering a crystal increases, which reduces its de Broglie wavelength. Consider a beam of 54-eV electrons directed at a nickel target. The potential energy of an electron that enters the target changes by 26 eV. (*a*) Compare the electron speeds outside and inside the target. (*b*) Compare the respective de Broglie wavelengths.

26. A beam of 50-keV electrons is directed at a crystal and diffracted electrons are found at an angle of 50° relative to the original beam. What is the spacing of the atomic planes of the crystal? A relativistic calculation is needed for λ.

3.6 Particle in a Box

27. Obtain an expression for the energy levels (in MeV) of a neutron confined to a one-dimensional box 1.00×10^{-14} m wide. What is the neutron's minimum energy? (The diameter of an atomic nucleus is of this order of magnitude.)

28. The lowest energy possible for a certain particle trapped in a certain box is 1.00 eV. (*a*) What are the next two higher energies the particle can have? (*b*) If the particle is an electron, how wide is the box?

29. A proton in a one-dimensional box has an energy of 400 keV in its first excited state. How wide is the box?

3.7 Uncertainty Principle I
3.8 Uncertainty Principle II
3.9 Applying the Uncertainty Principle

30. Discuss the prohibition of $E = 0$ for a particle trapped in a box L wide in terms of the uncertainty principle. How does the minimum momentum of such a particle compare with the momentum uncertainty required by the uncertainty principle if we take $\Delta x = L$?

31. The atoms in a solid possess a certain minimum **zero-point** energy even at 0 K, while no such restriction holds for the molecules in an ideal gas. Use the uncertainty principle to explain these statements.

32. Compare the uncertainties in the velocities of an electron and a proton confined in a 1.00-nm box.

33. The position and momentum of a 1.00-keV electron are simultaneously determined. If its position is located to within 0.100 nm, what is the percentage of uncertainty in its momentum?

34. (*a*) How much time is needed to measure the kinetic energy of an electron whose speed is 10.0 m/s with an uncertainty of no more than 0.100 percent? How far will the electron have traveled in this period of time? (*b*) Make the same calculations for a 1.00-g insect whose speed is the same. What do these sets of figures indicate?

35. How accurately can the position of a proton with $v \ll c$ be determined without giving it more than 1.00 keV of kinetic energy?

36. (*a*) Find the magnitude of the momentum of a particle in a box in its nth state. (*b*) The minimum change in the particle's momentum that a measurement can cause corresponds to a change of ± 1 in the quantum number n. If $\Delta x = L$, show that $\Delta p\, \Delta x \geq \hbar/2$.

37. A marine radar operating at a frequency of 9400 MHz emits groups of electromagnetic waves 0.0800 μs in duration. The time needed for the reflections of these groups to return indicates the distance to a target. (*a*) Find the length of each group and the number of waves it contains. (*b*) What is the approximate minimum bandwidth (that is, spread of frequencies) the radar receiver must be able to process?

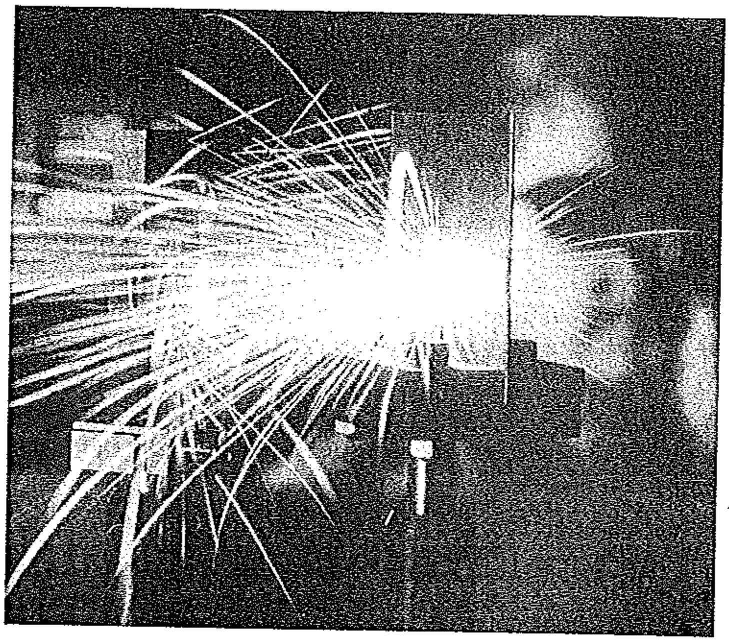
38. An unstable elementary particle called the eta meson has a rest mass of 549 MeV/c^2 and a mean lifetime of 7.00×10^{-19} s. What is the uncertainty in its rest mass?

39. The frequency of oscillation of a harmonic oscillator of mass m and spring constant C is $\nu = \sqrt{C/m}/2\pi$. The energy of the oscillator is $E = p^2/2m + Cx^2/2$, where p is its momentum when its displacement from the equilibrium position is x. In classical physics the minimum energy of the oscillator is $E_{min} = 0$. Use the uncertainty principle to find an expression for E in terms of x only and show that the minimum energy is actually $E_{min} = h\nu/2$ by setting $dE/dx = 0$ and solving for E_{min}.

40. (*a*) Verify that the uncertainty principle can be expressed in the form $\Delta L\, \Delta\theta \geq \hbar/2$, where ΔL is the uncertainty in the angular momentum of a particle and $\Delta\theta$ is the uncertainty in its angular position. (*Hint*: Consider a particle of mass m moving in a circle of radius r at the speed v, for which $L = mvr$.) (*b*) At what uncertainty in L will the angular position of a particle become completely indeterminate?

CHAPTER 4

Atomic Structure



Solid-state infrared laser cutting 1.6-mm steel sheet. This laser uses an yttrium-aluminum-garnet crystal doped with neodymium. The neodymium is pumped with radiation from small semiconductor lasers, a highly efficient method.

4.1 THE NUCLEAR ATOM
An atom is largely empty space

4.2 ELECTRON ORBITS
The planetary model of the atom and why it fails

4.3 ATOMIC SPECTRA
Each element has a characteristic line spectrum

4.4 THE BOHR ATOM
Electron waves in the atom

4.5 ENERGY LEVELS AND SPECTRA
A photon is emitted when an electron jumps from one energy level to a lower level

4.6 CORRESPONDENCE PRINCIPLE
The greater the quantum number, the closer quantum physics approaches classical physics

4.7 NUCLEAR MOTION
The nuclear mass affects the wavelengths of spectral lines

4.8 ATOMIC EXCITATION
How atoms absorb and emit energy

4.9 THE LASER
How to produce light waves all in step

APPENDIX: RUTHERFORD SCATTERING

Far in the past people began to suspect that matter, despite appearing continuous, has a definite structure on a microscopic level beyond the direct reach of our senses. This suspicion did not take on a more concrete form until a little over a century and a half ago. Since then the existence of atoms and molecules, the ultimate particles of matter in its common forms, has been amply demonstrated, and their own ultimate particles, electrons, protons, and neutrons, have been identified and studied as well. In this chapter and in others to come our chief concern will be the structure of the atom, since it is this structure that is responsible for nearly all the properties of matter that have shaped the world around us.

Every atom consists of a small nucleus of protons and neutrons with a number of electrons some distance away. It is tempting to think that the electrons circle the nucleus as planets do the sun, but classical electromagnetic theory denies the possibility of stable electron orbits. In an effort to resolve this paradox, Niels Bohr applied quantum ideas to atomic structure in 1913 to obtain a model which, despite its inadequacies and later replacement by a quantum-mechanical description of greater accuracy and usefulness, still remains a convenient mental picture of the atom. Bohr's theory of the hydrogen atom is worth examining both for this reason and because it provides a valuable transition to the more abstract quantum theory of the atom.

4.1 THE NUCLEAR ATOM

An atom is largely empty space

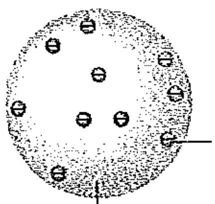


Figure 4.1 The Thomson model of the atom. The Rutherford scattering experiment showed it to be incorrect.

Most scientists of the late nineteenth century accepted the idea that the chemical elements consist of atoms, but they knew almost nothing about the atoms themselves. One clue was the discovery that all atoms contain electrons. Since electrons carry negative charges whereas atoms are neutral, positively charged matter of some kind must be present in atoms. But what kind? And arranged in what way?

One suggestion, made by the British physicist J. J. Thomson in 1898, was that atoms are just positively charged lumps of matter with electrons embedded in them, like raisins in a fruitcake (Fig. 4.1). Because Thomson had played an important role in discovering the electron, his idea was taken seriously. But the real atom turned out to be quite different.

The most direct way to find out what is inside a fruitcake is to poke a finger into it, which is essentially what Hans Geiger and Ernest Marsden did in 1911. At the suggestion of Ernest Rutherford, they used as probes the fast **alpha particles** emitted by certain radioactive elements. Alpha particles are helium atoms that have lost two electrons each, leaving them with a charge of $+2e$.

Geiger and Marsden placed a sample of an alpha-emitting substance behind a lead screen with a small hole in it, as in Fig. 4.2, so that a narrow beam of alpha particles was produced. This beam was directed at a thin gold foil. A zinc sulfide screen, which gives off a visible flash of light when struck by an alpha particle, was set on the other side of the foil with a microscope to see the flashes.

It was expected that the alpha particles would go right through the foil with hardly any deflection. This follows from the Thomson model, in which the electric charge inside an atom is assumed to be uniformly spread through its volume. With only weak electric forces exerted on them, alpha particles that pass through a thin foil ought to be deflected only slightly, 1° or less.

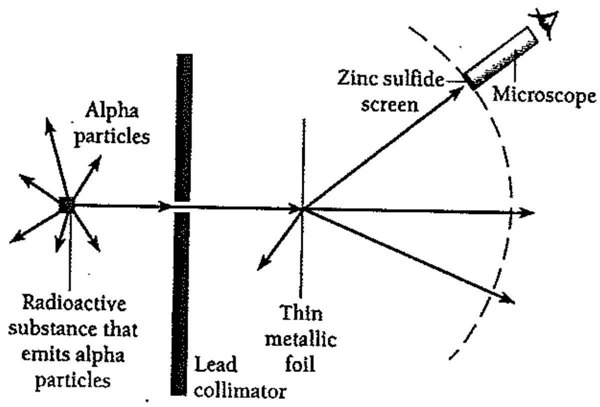


Figure 4.2 The Rutherford scattering experiment.

What Geiger and Marsden actually found was that although most of the alpha particles indeed were not deviated by much, a few were scattered through very large angles. Some were even scattered in the backward direction. As Rutherford remarked, "It was as incredible as if you fired a 15-inch shell at a piece of tissue paper and it came back and hit you."

Alpha particles are relatively heavy (almost 8000 electron masses) and those used in this experiment had high speeds (typically 2×10^7 m/s), so it was clear that powerful forces were needed to cause such marked deflections. The only way to

Ernest Rutherford (1871–1937), a native of New Zealand, was on his family's farm digging potatoes when he learned that he had won a scholarship for graduate study at Cambridge University in England. "This is the last potato I will every dig," he said, throwing down his spade. Thirteen years later he received the Nobel Prize in chemistry.

At Cambridge, Rutherford was a research student under J. J. Thomson, who would soon announce the discovery of the electron. Rutherford's own work was on the newly found phenomenon of radioactivity, and he quickly distinguished between alpha and beta particles, two of the emissions of radioactive materials. In 1898 he went to McGill University in Canada, where he found that alpha particles are the nuclei of helium atoms and that the radioactive decay of an element gives rise to another element. Working with the chemist Frederick Soddy and others, Rutherford traced the successive transformations of radioactive elements, such as uranium and radium, until they end up as stable lead.

In 1907 Rutherford returned to England as professor of physics at Manchester, where in 1911 he showed that the nuclear model of the atom was the only one that could explain the observed scattering of alpha particles by thin metal foils. Rutherford's last important discovery, reported in 1919, was the disintegration of nitrogen nuclei when bombarded with alpha particles, the first example of the artificial transmutation of elements into other elements. After other similar experiments, Rutherford suggested that all nuclei contain hydrogen nuclei, which he called protons. He also proposed that a neutral particle was present in nuclei as well.

In 1919 Rutherford became director of the Cavendish Laboratory at Cambridge, where under his stimulus great strides in understanding the nucleus continued to be made. James Chadwick discovered the neutron there in 1932. The Cavendish Laboratory was the site of the first accelerator for producing high-energy particles. With the help of this accelerator, fusion reactions in which light nuclei unite to form heavier nuclei were observed for the first time.

Rutherford was not infallible: only a few years before the discovery of fission and the building of the first nuclear reactor, he dismissed the idea of practical uses for nuclear energy as "moonshine." He died in 1937 of complications of a hernia and was buried near Newton in Westminster Abbey.

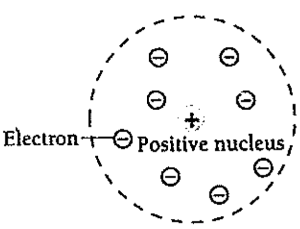


Figure 4.3 The Rutherford model of the atom.

explain the results, Rutherford found, was to picture an atom as being composed of a tiny nucleus in which its positive charge and nearly all its mass are concentrated, with the electrons some distance away (Fig. 4.3). With an atom being largely empty space, it is easy to see why most alpha particles go right through a thin foil. However, when an alpha particle happens to come near a nucleus, the intense electric field there scatters it through a large angle. The atomic electrons, being so light, do not appreciably affect the alpha particles.

The experiments of Geiger and Marsden and later work of a similar kind also supplied information about the nuclei of the atoms that composed the various target foils. The deflection of an alpha particle when it passes near a nucleus depends on the magnitude of the nuclear charge. Comparing the relative scattering of alpha particles by different foils thus provides a way to find the nuclear charges of the atoms involved.

All the atoms of any one element turned out to have the same unique nuclear charge, and this charge increased regularly from element to element in the periodic table. The nuclear charges always turned out to be multiples of $+e$; the number Z of unit positive charges in the nuclei of an element is today called the **atomic number** of the element. We know now that protons, each with a charge $+e$, provide the charge on a nucleus, so the atomic number of an element is the same as the number of protons in the nuclei of its atoms.

Ordinary matter, then, is mostly empty space. The solid wood of a table, the steel that supports a bridge, the hard rock underfoot, all are simply collections of tiny charged particles comparatively farther away from one another than the sun is from the planets. If all the actual matter, electrons and nuclei, in our bodies could somehow be packed closely together, we would shrivel to specks just visible with a microscope.

Rutherford Scattering Formula

The formula that Rutherford obtained for alpha particle scattering by a thin foil on the basis of the nuclear model of the atom is

Rutherford scattering formula

$$N(\theta) = \frac{N_i n t Z^2 e^4}{(8\pi\epsilon_0)^2 r^2 \, \mathrm{KE}^2 \sin^4(\theta/2)} \tag{4.1}$$

This formula is derived in the Appendix to this chapter. The symbols in Eq. (4.1) have the following meanings:

$N(\theta)$ = number of alpha particles per unit area that reach the screen at a scattering angle of θ
N_i = total number of alpha particles that reach the screen
n = number of atoms per unit volume in the foil
Z = atomic number of the foil atoms
r = distance of the screen from the foil
KE = kinetic energy of the alpha particles
t = foil thickness

The predictions of Eq. (4.1) agreed with the measurements of Geiger and Marsden, which supported the hypothesis of the nuclear atom. This is why Rutherford is credited

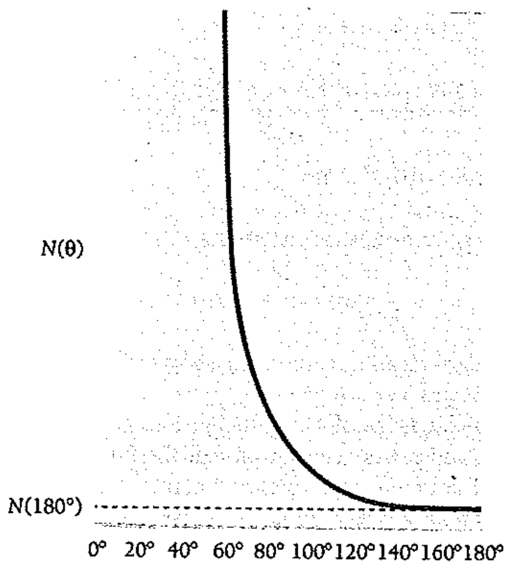


Figure 4.4 Rutherford scattering. $N(\theta)$ is the number of alpha particles per unit area that reach the screen at a scattering angle of θ; $N(180°)$ is this number for backward scattering. The experimental findings follow this curve, which is based on the nuclear model of the atom.

with the "discovery" of the nucleus. Because $N(\theta)$ is inversely proportional to $\sin^4 (\theta/2)$ the variation of $N(\theta)$ with θ is very pronounced (Fig. 4.4): only 0.14 percent of the incident alpha particles are scattered by more than 1°.

Nuclear Dimensions

In his derivation of Eq. (4.1) Rutherford assumed that the size of a target nucleus is small compared with the minimum distance R to which incident alpha particles approach the nucleus before being deflected away. Rutherford scattering therefore gives us a way to find an upper limit to nuclear dimensions.

Let us see what the distance of closest approach R was for the most energetic alpha particles employed in the early experiments. An alpha particle will have its smallest R when it approaches a nucleus head on, which will be followed by a 180° scattering. At the instant of closest approach the initial kinetic energy KE of the particle is entirely converted to electric potential energy, and so at that instant

$$\mathrm{KE}_{\mathrm{initial}} = \mathrm{PE} = \frac{1}{4\pi\epsilon_0}\frac{2Ze^2}{R}$$

since the charge of the alpha particle is $2e$ and that of the nucleus is Ze. Hence

Distance of closest approach

$$R = \frac{2Ze^2}{4\pi\epsilon_0 \mathrm{KE}_{\mathrm{initial}}} \qquad (4.2)$$

The maximum KE found in alpha particles of natural origin is 7.7 MeV, which is 1.2×10^{-12} J. Since $1/4\pi\epsilon_0 = 9.0 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2$,

$$R = \frac{(2)(9.0 \times 10^9 \text{ N} \cdot \text{m}^2/\text{C}^2)(1.6 \times 10^{-19} \text{ C})^2 Z}{1.2 \times 10^{-12} \text{ J}}$$

$$= 3.8 \times 10^{-16} Z \text{ m}$$

The atomic number of gold, a typical foil material, is $Z = 79$, so that

$$R \text{ (Au)} = 3.0 \times 10^{-14} \text{ m}$$

The radius of the gold nucleus is therefore less than 3.0×10^{-14} m, well under 10^{-4} the radius of the atom as a whole.

In more recent years particles of much higher energies than 7.7 MeV have been artificially accelerated, and it has been found that the Rutherford scattering formula does indeed eventually fail to agree with experiment. These experiments and the information they provide on actual nuclear dimensions are discussed in Chap. 11. The radius of the gold nucleus turns out to be about $\frac{1}{5}$ of the value of R (Au) found above.

Neutron Stars

The density of nuclear matter is about 2.4×10^{17} kg/m^3, which is equivalent to 4 billion tons per cubic inch. As discussed in Sec. 9.11, neutron stars are stars whose atoms have been so compressed that most of their protons and electrons have fused into neutrons, which are the most stable form of matter under enormous pressures. The densities of neutron stars are comparable to those of nuclei: a neutron star packs the mass of one or two suns into a sphere only about 10 km in radius. If the earth were this dense, it would fit into a large apartment house.

4.2 ELECTRON ORBITS

The planetary model of the atom and why it fails

The Rutherford model of the atom, so convincingly confirmed by experiment, pictures a tiny, massive, positively charged nucleus surrounded at a relatively great distance by enough electrons to render the atom electrically neutral as a whole. The electrons cannot be stationary in this model, because there is nothing that can keep them in place against the electric force pulling them to the nucleus. If the electrons are in motion, however, dynamically stable orbits like those of the planets around the sun are possible (Fig. 4.5).

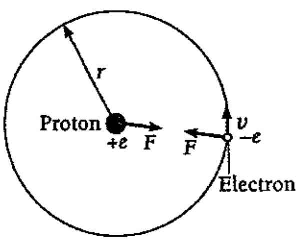


Figure 4.5 Force balance in the hydrogen atom.

Let us look at the classical dynamics of the hydrogen atom, whose single electron makes it the simplest of all atoms. We assume a circular electron orbit for convenience, though it might as reasonably be assumed to be elliptical in shape. The centripetal force

$$F_c = \frac{mv^2}{r}$$

holding the electron in an orbit r from the nucleus is provided by the electric force

$$F_e = \frac{1}{4\pi\epsilon_0}\frac{e^2}{r^2}$$

between them. The condition for a dynamically stable orbit is

$$F_c = F_e$$

$$\frac{mv^2}{r} = \frac{1}{4\pi\epsilon_0}\frac{e^2}{r^2} \tag{4.3}$$

The electron velocity v is therefore related to its orbit radius r by the formula

Electron velocity

$$v = \frac{e}{\sqrt{4\pi\epsilon_0 mr}} \tag{4.4}$$

The total energy E of the electron in a hydrogen atom is the sum of its kinetic and potential energies, which are

$$\text{KE} = \frac{1}{2}mv^2 \qquad \text{PE} = -\frac{e^2}{4\pi\epsilon_0 r}$$

(The minus sign follows from the choice of PE $= 0$ at $r = \infty$, that is, when the electron and proton are infinitely far apart.) Hence

$$E = \text{KE} + \text{PE} = \frac{mv^2}{2} - \frac{e^2}{4\pi\epsilon_0 r}$$

Substituting for v from Eq. (4.4) gives

$$E = \frac{e^2}{8\pi\epsilon_0 r} - \frac{e^2}{4\pi\epsilon_0 r}$$

Total energy of hydrogen atom

$$E = -\frac{e^2}{8\pi\epsilon_0 r} \tag{4.5}$$

The total energy of the electron is negative. This holds for every atomic electron and reflects the fact that it is bound to the nucleus. If E were greater than zero, an electron would not follow a closed orbit around the nucleus.

Actually, of course, the energy E is not a property of the electron alone but is a property of the system of electron + nucleus. The effect of the sharing of E between the electron and the nucleus is considered in Sec. 4.7.

Example 4.1

Experiments indicate that 13.6 eV is required to separate a hydrogen atom into a proton and an electron; that is, its total energy is $E = -13.6$ eV. Find the orbital radius and velocity of the electron in a hydrogen atom.

Solution

Since 13.6 eV = 2.2×10^{-18} J, from Eq. (4.5)

$$r = -\frac{e^2}{8\pi\epsilon_0 E} = -\frac{(1.6 \times 10^{-19}\ \text{C})^2}{(8\pi)(8.85 \times 10^{-12}\ \text{F/m})(-2.2 \times 10^{-18}\ \text{J})}$$
$$= 5.3 \times 10^{-11}\ \text{m}$$

An atomic radius of this magnitude agrees with estimates made in other ways. The electron's velocity can be found from Eq. (4.4):

$$v = \frac{e}{\sqrt{4\pi\epsilon_0 m r}} = \frac{1.6 \times 10^{-19}\ \text{C}}{\sqrt{(4\pi)(8.85 \times 10^{-12}\ \text{F/m})(9.1 \times 10^{-31}\ \text{kg})(5.3 \times 10^{-11}\ \text{m})}}$$
$$= 2.2 \times 10^6\ \text{m/s}$$

Since $v \ll c$, we can ignore special relativity when considering the hydrogen atom.

The Failure of Classical Physics

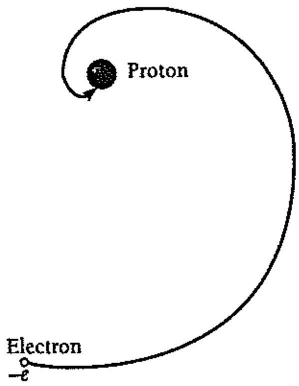


Figure 4.6 An atomic electron should, classically, spiral rapidly into the nucleus as it radiates energy due to its acceleration.

The analysis above is a straightforward application of Newton's laws of motion and Coulomb's law of electric force—both pillars of classical physics—and is in accord with the experimental observation that atoms are stable. However, it is *not* in accord with electromagnetic theory—another pillar of classical physics—which predicts that accelerated electric charges radiate energy in the form of em waves. An electron pursuing a curved path is accelerated and therefore should continuously lose energy, spiraling into the nucleus in a fraction of a second (Fig. 4.6).

But atoms do not collapse. This contradiction further illustrates what we saw in the previous two chapters: The laws of physics that are valid in the macroworld do not always hold true in the microworld of the atom.

Is Rutherford's Analysis Valid?

An interesting question comes up at this point. When he derived his scattering formula, Rutherford used the same laws of physics that prove such dismal failures when applied to atomic stability. Might it not be that this formula is not correct and that in reality the atom does not resemble Rutherford's model of a small central nucleus surrounded by distant electrons? This is not a trivial point. It is a curious coincidence that the quantum-mechanical analysis of alpha particle scattering by thin foils yields precisely the same formula that Rutherford found.

To verify that a classical calculation ought to be at least approximately correct, we note that the de Broglie wavelength of an alpha particle whose speed is 2.0×10^7 m/s is

$$\lambda = \frac{h}{mv} = \frac{6.63 \times 10^{-34}\ \text{J} \cdot \text{s}}{(6.6 \times 10^{-27}\ \text{kg})(2.0 \times 10^7\ \text{m/s})}$$
$$= 5.0 \times 10^{-15}\ \text{m}$$

As we saw in Sec. 4.1, the closest an alpha particle with this wavelength ever gets to a gold nucleus is 3.0×10^{-14} m, which is six de Broglie wavelengths. It is therefore just reasonable to regard the alpha particle as a classical particle in the interaction. We are correct in thinking of the atom in terms of Rutherford's model, though the dynamics of the atomic electrons—which is another matter—requires a nonclassical approach.

Classical physics fails to provide a meaningful analysis of atomic structure because it approaches nature in terms of "pure" particles and "pure" waves. In reality particles and waves have many properties in common, though the smallness of Planck's constant makes the wave-particle duality imperceptible in the macroworld. The usefulness of classical physics decreases as the scale of the phenomena under study decreases, and we must allow for the particle behavior of waves and the wave behavior of particles to understand the atom. In the rest of this chapter we shall see how the Bohr atomic model, which combines classical and modern notions, accomplishes part of the latter task. Not until we consider the atom from the point of view of quantum mechanics, which makes no compromise with the intuitive notions we pick up in our daily lives, will we find a really successful theory of the atom.

4.3 ATOMIC SPECTRA

Each element has a characteristic line spectrum

Atomic stability is not the only thing that a successful theory of the atom must account for. The existence of spectral lines is another important aspect of the atom that finds no explanation in classical physics.

We saw in Chap. 2 that condensed matter (solids and liquids) at all temperatures emits em radiation in which all wavelengths are present, though with different intensities. The observed features of this radiation were explained by Planck without reference to exactly how it was produced by the radiating material or to the nature of the material. From this it follows that we are witnessing the collective behavior of a great many interacting atoms rather than the characteristic behavior of the atoms of a particular element.

At the other extreme, the atoms or molecules in a rarefied gas are so far apart on the average that they only interact during occasional collisions. Under these circumstances we would expect any emitted radiation to be characteristic of the particular atoms or molecules present, which turns out to be the case.

When an atomic gas or vapor at somewhat less than atmospheric pressure is suitably "excited," usually by passing an electric current through it, the emitted radiation has a spectrum which contains certain specific wavelengths only. An idealized arrangement for observing such atomic spectra is shown in Fig. 4.7; actual spectrometers use diffraction

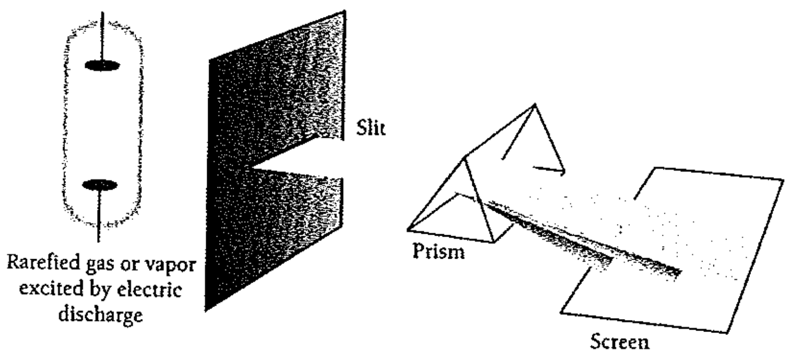


Figure 4.7 An idealized spectrometer.

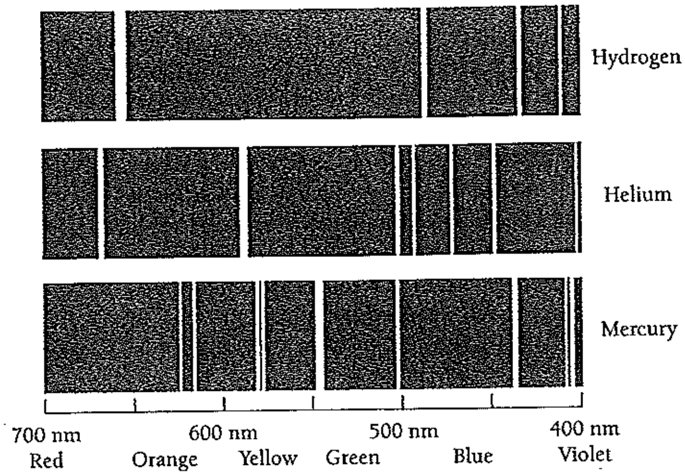


Figure 4.8 Some of the principal lines in the emission spectra of hydrogen, helium, and mercury.

Gas atoms excited by electric currents in these tubes radiate light of wavelengths characteristic of the gas used.

gratings. Figure 4.8 shows the **emission line spectra** of several elements. Every element displays a unique line spectrum when a sample of it in the vapor phase is excited. Spectroscopy is therefore a useful tool for analyzing the composition of an unknown substance.

When white light is passed through a gas, the gas is found to absorb light of certain of the wavelengths present in its emission spectrum. The resulting **absorption line spectrum** consists of a bright background crossed by dark lines that correspond to the missing wavelengths (Fig. 4.9); emission spectra consist of bright lines on a dark background. The spectrum of sunlight has dark lines in it because the luminous part of the

Figure 4.9 The dark lines in the absorption spectrum of an element correspond to bright lines in its emission spectrum.

sun, which radiates very nearly like a blackbody heated to 5800 K, is surrounded by an envelope of cooler gas that absorbs light of certain wavelengths only. Most other stars have spectra of this kind.

The number, intensity, and exact wavelengths of the lines in the spectrum of an element depend upon temperature, pressure, the presence of electric and magnetic fields, and the motion of the source. It is possible to tell by examining its spectrum not only what elements are present in a light source but much about their physical state. An astronomer, for example, can establish from the spectrum of a star which elements its atmosphere contains, whether they are ionized, and whether the star is moving toward or away from the earth.

Spectral Series

A century ago the wavelengths in the spectrum of an element were found to fall into sets called **spectral series.** The first such series was discovered by J. J. Balmer in 1885 in the course of a study of the visible part of the hydrogen spectrum. Figure 4.10 shows the **Balmer series.** The line with the longest wavelength, 656.3 nm, is designated H_α, the next, whose wavelength is 486.3 nm, is designated H_β, and so on. As the wave-length decreases, the lines are found closer together and weaker in intensity until the **series limit** at 364.6 nm is reached, beyond which there are no further separate lines but only a faint continuous spectrum. Balmer's formula for the wavelengths of this series is

Balmer $$\frac{1}{\lambda} = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \quad n = 3, 4, 5, \ldots \tag{4.6}$$

The quantity R, known as the **Rydberg constant,** has the value

Rydberg constant $$R = 1.097 \times 10^7 \text{ m}^{-1} = 0.01097 \text{ nm}^{-1}$$

The H_α line corresponds to $n = 3$, the H_β line to $n = 4$, and so on. The series limit corresponds to $n = \infty$, so that it occurs at a wavelength of $4/R$, in agreement with experiment.

The Balmer series contains wavelengths in the visible portion of the hydrogen spectrum. The spectral lines of hydrogen in the ultraviolet and infrared regions fall into several other series. In the ultraviolet the **Lyman series** contains the wavelengths given by the formula

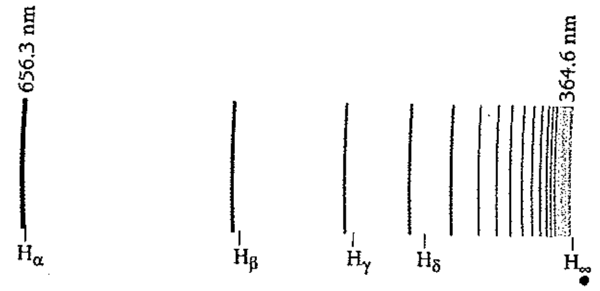


Figure 4.10 The Balmer series of hydrogen. The H_α line is red, the H_β line is blue, the H_γ and H_δ lines are violet, and the other lines are in the near ultraviolet.

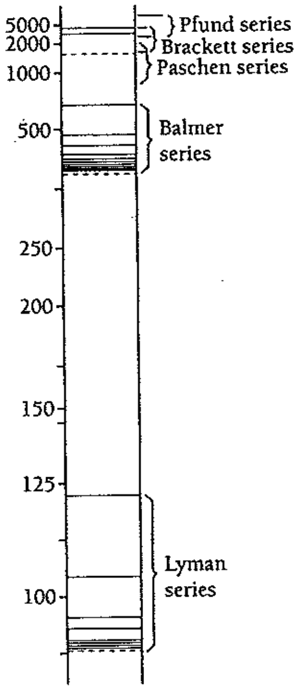


Figure 4.11 The spectral series of hydrogen. The wavelengths in each series are related by simple formulas.

Lyman $$\frac{1}{\lambda} = R\left(\frac{1}{1^2} - \frac{1}{n^2}\right) \qquad n = 2, 3, 4, \ldots \tag{4.7}$$

In the infrared, three spectral series have been found whose lines have the wavelengths specified by the formulas

Paschen $$\frac{1}{\lambda} = R\left(\frac{1}{3^2} - \frac{1}{n^2}\right) \qquad n = 4, 5, 6, \ldots \tag{4.8}$$

Brackett $$\frac{1}{\lambda} = R\left(\frac{1}{4^2} - \frac{1}{n^2}\right) \qquad n = 5, 6, 7, \ldots \tag{4.9}$$

Pfund $$\frac{1}{\lambda} = R\left(\frac{1}{5^2} - \frac{1}{n^2}\right) \qquad n = 6, 7, 8, \ldots \tag{4.10}$$

These spectral series of hydrogen are plotted in terms of wavelength in Fig. 4.11; the Brackett series evidently overlaps the Paschen and Pfund series. The value of R is the same in Eqs. (4.6) to (4.10).

These observed regularities in the hydrogen spectrum, together with similar regularities in the spectra of more complex elements, pose a definitive test for any theory of atomic structure.

4.4 THE BOHR ATOM

Electron waves in the atom

The first theory of the atom to meet with any success was put forward in 1913 by Niels Bohr. The concept of matter waves leads in a natural way to this theory, as de Broglie found, and this is the route that will be followed here. Bohr himself used a different approach, since de Broglie's work came a decade later, which makes his achievement all the more remarkable. The results are exactly the same, however.

We start by examining the wave behavior of an electron in orbit around a hydrogen nucleus. (In this chapter, since the electron velocities are much smaller than c, we will assume that $\gamma = 1$ and for simplicity omit γ from the various equations.) The de Broglie wavelength of this electron is

$$\lambda = \frac{h}{mv}$$

where the electron velocity v is that given by Eq. (4.4):

$$v = \frac{e}{\sqrt{4\pi\epsilon_0 mr}}$$

Hence

Orbital electron wavelength $$\lambda = \frac{h}{e}\sqrt{\frac{4\pi\epsilon_0 r}{m}} \tag{4.11}$$

By substituting 5.3×10^{-11} m for the radius r of the electron orbit (see Example 4.1), we find the electron wavelength to be

$$\lambda = \frac{6.63 \times 10^{-34}\,\mathrm{J \cdot s}}{1.6 \times 10^{-19}\mathrm{C}} \sqrt{\frac{(4\pi)(8.85 \times 10^{-12}\,\mathrm{C^2/N \cdot m^2})(5.3 \times 10^{-11}\mathrm{m})}{9.1 \times 10^{-31}\,\mathrm{kg}}}$$

$$= 33 \times 10^{-11}\ \mathrm{m}$$

This wavelength is exactly the same as the circumference of the electron orbit,

$$2\pi r = 33 \times 10^{-11}\ \mathrm{m}$$

The orbit of the electron in a hydrogen atom corresponds to one complete electron wave joined on itself (Fig. 4.12)!

The fact that the electron orbit in a hydrogen atom is one electron wavelength in circumference provides the clue we need to construct a theory of the atom. If we consider the vibrations of a wire loop (Fig. 4.13), we find that their wavelengths always fit an integral number of times into the loop's circumference so that each wave joins smoothly with the next. If the wire were perfectly elastic, these vibrations would continue indefinitely. Why are these the only vibrations possible in a wire loop? If a fractional number of wavelengths is placed around the loop, as in Fig. 4.14, destructive

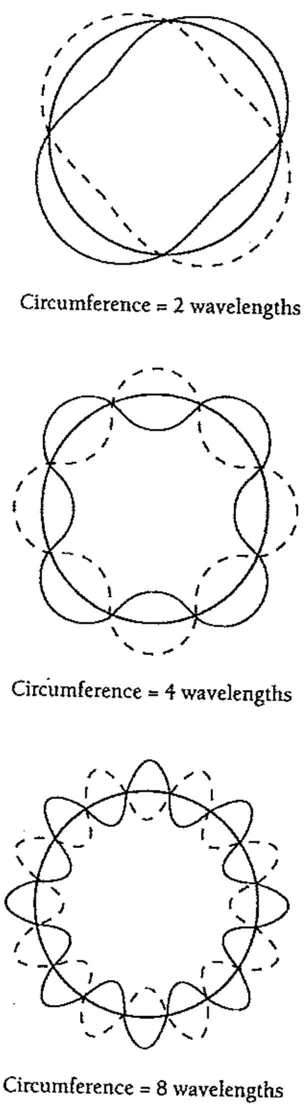


Figure 4.13 Some modes of vibration of a wire loop. In each case a whole number of wavelengths fit into the circumference of the loop.

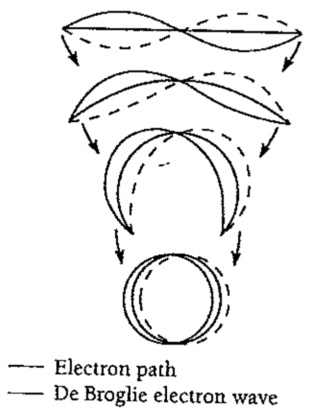


Figure 4.12 The orbit of the electron in a hydrogen atom corresponds to a complete electron de Broglie wave joined on itself.

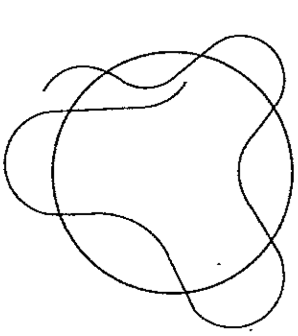


Figure 4.14 A fractional number of wavelengths cannot persist because destructive interference will occur.

Niels Bohr (1884–1962) was born and spent most of his life in Copenhagen, Denmark. After receiving his doctorate at the university there in 1911, Bohr went to England to broaden his scientific horizons. At Rutherford's laboratory in Manchester, Bohr was introduced to the just-discovered nuclear model of the atom, which was in conflict with the existing principles of physics. Bohr realized that it was "hopeless" to try to make sense of the atom in the framework of classical physics alone, and he felt that the quantum theory of light must somehow be the key to understanding atomic structure.

Back in Copenhagen in 1913, a friend suggested to Bohr that Balmer's formula for one set of the spectral lines of hydrogen might be relevant to his quest. "As soon as I saw Balmer's formula the whole thing was immediately clear to me," Bohr said later. To construct his theory, Bohr began with two revolutionary ideas. The first was that an atomic electron can circle its nucleus only in certain orbits, and the other was that an atom emits or absorbs a photon of light when an electron jumps from one permitted orbit to another.

What is the condition for a permitted orbit? To find out, Bohr used as a guide what became known as the correspondence principle: When quantum numbers are very large, quantum effects should not be conspicuous, and the quantum theory must then give the same results as classical physics. Applying this principle showed that the electron in a permitted orbit must have an angular momentum that is a multiple of $\hbar = h/2\pi$. A decade later Louis de Broglie explained this quantization of angular momentum in terms of the wave nature of a moving electron.

Bohr was able to account for all the spectral series of hydrogen, not just the Balmer series, but the publication of the theory aroused great controversy. Einstein, an enthusiastic supporter of the theory (which "appeared to me like a miracle—and appears to me as a miracle even today," he wrote many years later), nevertheless commented on its bold mix of classical and quantum concepts, "One ought to be ashamed of the successes [of the theory] because they have been earned according to the Jesuit maxim, 'Let not thy left hand know what the other doeth.'" Other noted physicists were more deeply disturbed: Otto Stern and Max von Laue said they would quit physics if Bohr were right. (They later changed their minds.) Bohr and others tried to extend his model to many-electron atoms with occasional success—for instance, the correct prediction of the properties of the then-unknown element hafnium—but real progress had to wait for Wolfgang Pauli's exclusion principle of 1925.

In 1916 Bohr returned to Rutherford's laboratory, where he stayed until 1919. Then an Institute of Theoretical Physics was created for him in Copenhagen, and he directed it until his death. The institute was a magnet for quantum theoreticians from all over the world, who were stimulated by the exchange of ideas at regular meetings there. Bohr received the Nobel Prize in 1922. His last important work came in 1939, when he used an analogy between a large nucleus and a liquid drop to explain why nuclear fission, which had just been discovered, occurs in certain nuclei but not in others. During World War II Bohr contributed to the development of the atomic bomb at Los Alamos, New Mexico. After the war, Bohr returned to Copenhagen, where he died in 1962.

interference will occur as the waves travel around the loop, and the vibrations will die out rapidly.

By considering the behavior of electron waves in the hydrogen atom as analogous to the vibrations of a wire loop, then, we can say that

> An electron can circle a nucleus only if its orbit contains an integral number of de Broglie wavelengths.

This statement combines both the particle and wave characters of the electron since the electron wavelength depends upon the orbital velocity needed to balance the pull of the nucleus. To be sure, the analogy between an atomic electron and the standing waves of Fig. 4.13 is hardly the last word on the subject, but it represents an illuminating step along the path to the more profound and comprehensive, but also more abstract, quantum-mechanical theory of the atom.

It is easy to express the condition that an electron orbit contain an integral number of de Broglie wavelengths. The circumference of a circular orbit of radius r is $2\pi r$, and so the condition for orbit stability is

Condition for orbit stability $$n\lambda = 2\pi r_n \qquad n = 1, 2, 3, \ldots \tag{4.12}$$

where r_n designates the radius of the orbit that contain n wavelengths. The integer n is called the **quantum number** of the orbit. Substituting for λ, the electron wavelength given by Eq. (4.11), yields

$$\frac{nh}{e}\sqrt{\frac{4\pi\epsilon_0 r_n}{m}} = 2\pi r_n$$

and so the possible electron orbits are those whose radii are given by

Orbital radii in Bohr atom $$r_n = \frac{n^2h^2\epsilon_0}{\pi me^2} \qquad n = 1, 2, 3, \ldots \tag{4.13}$$

The radius of the innermost orbit is customarily called the **Bohr radius** of the hydrogen atom and is denoted by the symbol a_0:

Bohr radius $$a_0 = r_1 = 5.292 \times 10^{-11}\ \text{m}$$

The other radii are given in terms of a_0 by the formula

$$r_n = n^2 a_0 \tag{4.14}$$

4.5 ENERGY LEVELS AND SPECTRA

A photon is emitted when an electron jumps from one energy level to a lower level

The various permitted orbits involve different electron energies. The electron energy E_n is given in terms of the orbit radius r_n by Eq. (4.5) as

$$E_n = -\frac{e^2}{8\pi\epsilon_0 r_n}$$

Substituting for r_n from Eq (4.13), we see that

Energy levels $$E_n = -\frac{me^4}{8\epsilon_0^2h^2}\left(\frac{1}{n^2}\right) = \frac{E_1}{n^2} \qquad n = 1, 2, 3, \ldots \tag{4.15}$$

$$E_1 = -2.18 \times 10^{-18}\ \text{J} = -13.6\ \text{eV}$$

The energies specified by Eq. (4.15) are called the **energy levels** of the hydrogen atom and are plotted in Fig. 4.15. These levels are all negative, which signifies that the electron does not have enough energy to escape from the nucleus. An atomic electron can have only these energies and no others. An analogy might be a person on a ladder, who can stand only on its steps and not in between.

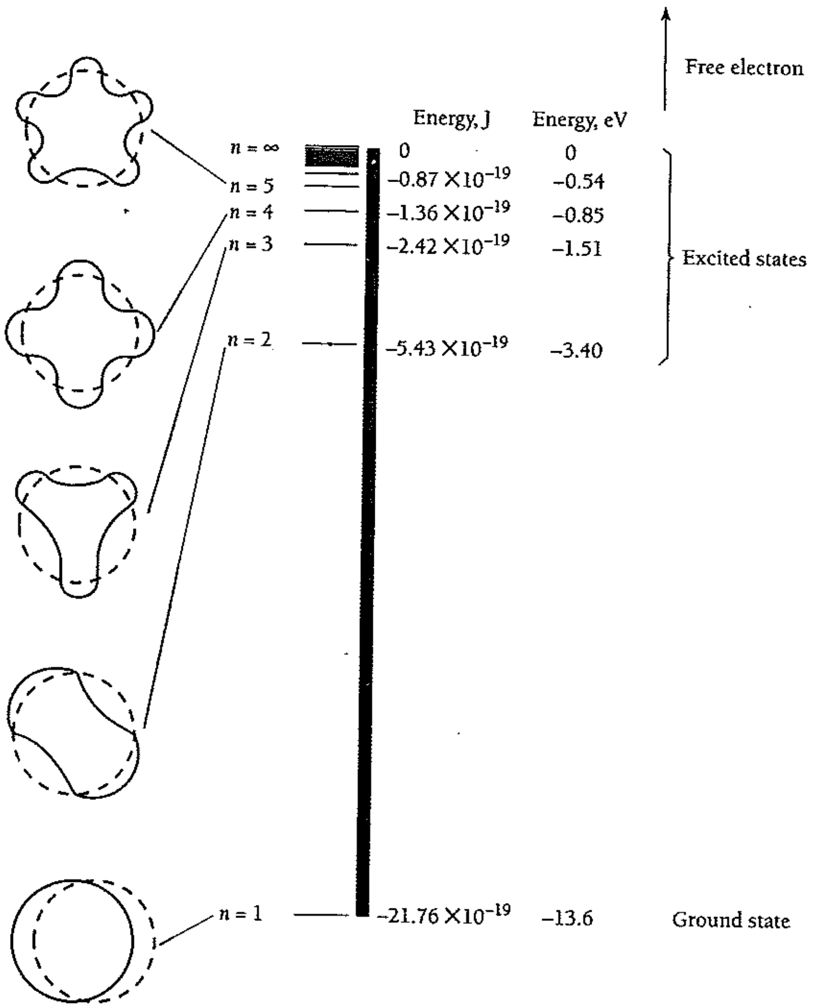


Figure 4.15 Energy levels of the hydrogen atom.

The lowest energy level E_1 is called the **ground state** of the atom, and the higher levels E_2, E_3, E_4, . . . are called **excited states.** As the quantum number n increases, the corresponding energy E_n approaches closer to 0. In the limit of $n = \infty$, $E_\infty = 0$ and the electron is no longer bound to the nucleus to form an atom. A positive energy for a nucleus-electron combination means that the electron is free and has no quantum conditions to fulfill; such a combination does not constitute an atom, of course.

The work needed to remove an electron from an atom in its ground state is called its **ionization energy.** The ionization energy is accordingly equal to $-E_1$, the energy that must be provided to raise an electron from its ground state to an energy of $E = 0$, when it is free. In the case of hydrogen, the ionization energy is 13.6 eV since the ground-state energy of the hydrogen atom is −13.6 eV. Figure 7.10 shows the ionization energies of the elements.

Example 4.2

An electron collides with a hydrogen atom in its ground state and excites it to a state of $n = 3$. How much energy was given to the hydrogen atom in this inelastic (KE not conserved) collision?

Solution

From Eq. (4.15) the energy change of a hydrogen atom that goes from an initial state of quantum number n_i to a final state of quantum number n_f is

$$\Delta E = E_f - E_i = \frac{E_1}{n_f^2} - \frac{E_1}{n_i^2} = E_1\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)$$

Here $n_i = 1$, $n_f = 3$, and $E_1 = -13.6$ eV, so

$$\Delta E = -13.6\left(\frac{1}{3^2} - \frac{1}{1^2}\right)\text{eV} = 12.1 \text{ eV}$$

Example 4.3

Hydrogen atoms in states of high quantum number have been created in the laboratory and observed in space. They are called **Rydberg atoms.** (*a*) Find the quantum number of the Bohr orbit in a hydrogen atom whose radius is 0.0100 mm. (*b*) What is the energy of a hydrogen atom in this state?

Solution

(*a*) From Eq. (4.14) with $r_n = 1.00 \times 10^{-5}$ m,

$$n = \sqrt{\frac{r_n}{a_0}} = \sqrt{\frac{1.00 \times 10^{-5} \text{ m}}{5.29 \times 10^{-11} \text{ m}}} = 435$$

(*b*) From Eq. (4.15),

$$E_n = \frac{E_1}{n^2} = \frac{-13.6 \text{ eV}}{(435)^2} = -7.19 \times 10^{-5} \text{ eV}$$

Rydberg atoms are obviously extremely fragile and are easily ionized, which is why they are found in nature only in the near-vacuum of space. The spectra of Rydberg atoms range down to radio frequencies and their existence was established from radio telescope data.

Origin of Line Spectra

We must now confront the equations developed above with experiment. An especially striking observation is that atoms exhibit line spectra in both emission and absorption. Do such spectra follow from our model?

The presence of discrete energy levels in the hydrogen atom suggests the connection. Let us suppose that when an electron in an excited state drops to a lower state, the lost energy is emitted as a single photon of light. According to our model, electrons cannot exist in an atom except in certain specific energy levels. The jump of an electron from one level to another, with the difference in energy between the levels being given off all at once in a photon rather than in some more gradual manner, fits in well with this model.

Quantization in the Atomic World

Sequences of energy levels are characteristic of all atoms, not just those of hydrogen. As in the case of a particle in a box, the confinement of an electron to a region of space leads to restrictions on its possible wave functions that in turn limit the possible energies to well-defined values only. The existence of atomic energy levels is a further example of the quantization, or graininess, of physical quantities on a microscopic scale.

In the world of our daily lives, matter, electric charge, energy, and so forth appear to be continuous. In the world of the atom, in contrast, matter is composed of elementary particles that have definite rest masses, charge always comes in multiples of $+e$ or $-e$, electromagnetic waves of frequency ν appear as streams of photons each with the energy $h\nu$, and stable systems of particles, such as atoms, can possess only certain energies. As we shall find, other quantities in nature are also quantized, and this quantization enters into every aspect of how electrons, protons, and neutrons interact to endow the matter around us (and of which we consist) with its familiar properties.

If the quantum number of the initial (higher-energy) state is n_i and the quantum number of the final (lower-energy) state is n_f, we are asserting that

$$\text{Initial energy} - \text{final energy} = \text{photon energy}$$

$$E_i - E_f = h\nu \tag{4.16}$$

where ν is the frequency of the emitted photon. From Eq. (4.15) we have

$$E_i - E_f = E_1\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right) = -E_1\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)$$

We recall that E_1 is a negative quantity (-13.6 eV, in fact), so $-E_1$ is a positive quantity. The frequency of the photon released in this transition is therefore

$$\nu = \frac{E_i - E_f}{h} = -\frac{E_1}{h}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \tag{4.17}$$

Since $\lambda = c/\nu$, $1/\lambda = \nu/c$ and

Hydrogen spectrum

$$\frac{1}{\lambda} = -\frac{E_1}{ch}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \tag{4.18}$$

Equation (4.18) states that the radiation emitted by excited hydrogen atoms should contain certain wavelengths only. These wavelengths, furthermore, fall into definite sequences that depend upon the quantum number n_f of the final energy level of the electron (Fig. 4.16). Since $n_i > n_f$ in each case, in order that there be an excess of energy to be given off as a photon, the calculated formulas for the first five series are

Lyman $\quad n_f = 1: \quad \dfrac{1}{\lambda} = -\dfrac{E_1}{ch}\left(\dfrac{1}{1^2} - \dfrac{1}{n^2}\right) \quad n = 2, 3, 4, \ldots$

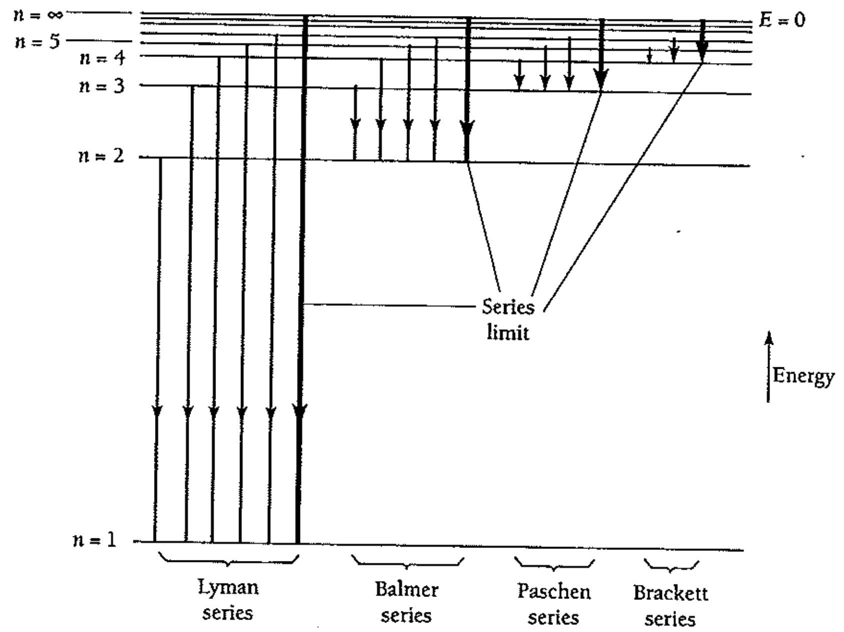


Figure 4.16 Spectral lines originate in transitions between energy levels. Shown are the spectral series of hydrogen. When $n = \infty$, the electron is free.

Balmer $\quad n_f = 2$: $\dfrac{1}{\lambda} = -\dfrac{E_1}{ch}\left(\dfrac{1}{2^2} - \dfrac{1}{n^2}\right) \quad n = 3, 4, 5, \ldots$

Paschen $\quad n_f = 3$: $\dfrac{1}{\lambda} = -\dfrac{E_1}{ch}\left(\dfrac{1}{3^2} - \dfrac{1}{n^2}\right) \quad n = 4, 5, 6, \ldots$

Brackett $\quad n_f = 4$: $\dfrac{1}{\lambda} = -\dfrac{E_1}{ch}\left(\dfrac{1}{4^2} - \dfrac{1}{n^2}\right) \quad n = 5, 6, 7, \ldots$

Pfund $\quad n_f = 5$: $\dfrac{1}{\lambda} = -\dfrac{E_1}{ch}\left(\dfrac{1}{5^2} - \dfrac{1}{n^2}\right) \quad n = 6, 7, 8, \ldots$

These sequences are identical in form with the empirical spectral series discussed earlier. The Lyman series corresponds to $n_f = 1$; the Balmer series corresponds to $n_f = 2$; the Paschen series corresponds to $n_f = 3$; the Brackett series corresponds to $n_f = 4$; and the Pfund series corresponds to $n_f = 5$.

Our final step is to compare the value of the constant term in the above equations with that of the Rydberg constant in Eqs. (4.6) to (4.10). The value of the constant term is

$$
\begin{aligned}
-\frac{E_1}{ch} &= \frac{me^4}{8\epsilon_0^2 ch^3} \\
&= \frac{(9.109 \times 10^{-31}\ \text{kg})(1.602 \times 10^{-19}\ \text{C})^4}{(8)(8.854 \times 10^{-12}\ \text{C}^2/\text{N}\cdot\text{m}^2)(2.998 \times 10^8\ \text{m/s})(6.626 \times 10^{-34}\ \text{J}\cdot\text{s})^3} \\
&= 1.097 \times 10^7\ \text{m}^{-1}
\end{aligned}
$$

which is indeed the same as R. Bohr's model of the hydrogen atom is therefore in accord with the spectral data.

Example 4.4

Find the longest wavelength present in the Balmer series of hydrogen, corresponding to the H_α line.

Solution

In the Balmer series the quantum number of the final state is $n_f = 2$. The longest wavelength in this series corresponds to the smallest energy difference between energy levels. Hence the initial state must be $n_i = 3$ and

$$\frac{1}{\lambda} = R\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) = R\left(\frac{1}{2^2} - \frac{1}{3^2}\right) = 0.139R$$

$$\lambda = \frac{1}{0.139R} = \frac{1}{0.139(1.097 \times 10^7 \text{m}^{-1})} = 6.56 \times 10^{-7}\text{m} = 656 \text{ nm}$$

This wavelength is near the red end of the visible spectrum.

4.6 CORRESPONDENCE PRINCIPLE

The greater the quantum number, the closer quantum physics approaches classical physics

Quantum physics, so different from classical physics in the microworld beyond reach of our senses, must nevertheless give the same results as classical physics in the macroworld where experiments show that the latter is valid. We have already seen that this basic requirement is true for the wave theory of moving bodies. We shall now find that it is also true for Bohr's model of the hydrogen atom.

According to electromagnetic theory, an electron moving in a circular orbit radiates em waves whose frequencies are equal to its frequency of revolution and to harmonics (that is, integral multiples) of that frequency. In a hydrogen atom the electron's speed is

$$v = \frac{e}{\sqrt{4\pi\epsilon_0 mr}}$$

according to Eq. (4.4), where r is the radius of its orbit. Hence the frequency of revolution f of the electron is

$$f = \frac{\text{electron speed}}{\text{orbit circumference}} = \frac{v}{2\pi r} = \frac{e}{2\pi\sqrt{4\pi\epsilon_0 mr^3}}$$

The radius r_n of a stable orbit is given in terms of its quantum number n by Eq. (4.13) as

$$r_n = \frac{n^2h^2\epsilon_0}{\pi me^2}$$

and so the frequency of revolution is

Frequency of revolution

$$f = \frac{me^4}{8\epsilon_0^2 h^3}\left(\frac{2}{n^3}\right) = \frac{-E_1}{h}\left(\frac{2}{n^3}\right) \tag{4.19}$$

Example 4.5

(*a*) Find the frequencies of revolution of electrons in $n = 1$ and $n = 2$ Bohr orbits. (*b*) What is the frequency of the photon emitted when an electron in an $n = 2$ orbit drops to an $n = 1$ orbit? (*c*) An electron typically spends about 10^{-8} s in an excited state before it drops to a lower state by emitting a photon. How many revolutions does an electron in an $n = 2$ Bohr orbit make in 1.00×10^{-8} s?

Solution

(*a*) From Eq. (4.19),

$$f_1 = \frac{-E_1}{h}\left(\frac{2}{1^3}\right) = \left(\frac{2.18 \times 10^{-18}\ \text{J}}{6.63 \times 10^{-34}\ \text{J}\cdot\text{s}}\right)(2) = 6.58 \times 10^{15}\ \text{rev/s}$$

$$f_2 = \frac{-E_1}{h}\left(\frac{2}{2^3}\right) = \frac{f_1}{8} = 0.823 \times 10^{15}\ \text{rev/s}$$

(*b*) From Eq. (4.17),

$$\nu = \frac{-E_1}{h}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) = \left(\frac{2.18 \times 10^{-18}\ \text{J}}{6.63 \times 10^{-34}\ \text{J}\cdot\text{s}}\right)\left(\frac{1}{1^3} - \frac{1}{2^3}\right) = 2.88 \times 10^{15}\ \text{Hz}$$

This frequency is intermediate between f_1 and f_2.

(*c*) The number of revolutions the electron makes is

$$N = f_2\,\Delta t = (8.23 \times 10^{14}\ \text{rev/s})(1.00\times 10^{-8}\ \text{s}) = 8.23 \times 10^{6}\ \text{rev}$$

The earth takes 8.23 million y to make this many revolutions around the sun.

Under what circumstances should the Bohr atom behave classically? If the electron orbit is so large that we might be able to measure it directly, quantum effects ought not to dominate. An orbit 0.01 mm across, for instance, meets this specification. As we found in Example 4.3, its quantum number is $n = 435$.

What does the Bohr theory predict such an atom will radiate? According to Eq. (4.17), a hydrogen atom dropping from the n_ith energy level to the n_fth energy level emits a photon whose frequency is

$$\nu = \frac{-E_1}{h}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)$$

Let us write n for the initial quantum number n_i and $n - p$ (where $p = 1, 2, 3, \ldots$) for the final quantum number n_f. With this substitution,

$$\nu = \frac{-E_1}{h}\left[\frac{1}{(n-p)^2} - \frac{1}{n^2}\right] = \frac{-E_1}{h}\left[\frac{2np - p^2}{n^2(n-p)^2}\right]$$

When n_i and n_f are both very large, n is much greater than p, and

$$2np - p^2 \approx 2np$$

$$(n-p)^2 \approx n^2$$

so that

Frequency of photon
$$\nu = \frac{-E_1}{h}\left(\frac{2p}{n^3}\right) \tag{4.20}$$

When $p = 1$, the frequency ν of the radiation is exactly the same as the frequency of rotation f of the orbital electron given in Eq. (4.19). Multiples of this frequency are radiated when $p = 2, 3, 4, \ldots$. Hence both quantum and classical pictures of the hydrogen atom make the same predictions in the limit of very large quantum numbers. When $n = 2$, Eq. (4.19) predicts a radiation frequency that differs from that given by Eq. (4.20) by almost 300 percent. When $n = 10{,}000$, the discrepancy is only about 0.01 percent.

The requirement that quantum physics give the same results as classical physics in the limit of large quantum numbers was called by Bohr the **correspondence principle.** It has played an important role in the development of the quantum theory of matter.

Bohr himself used the correspondence principle in reverse, so to speak, to look for the condition for orbit stability. Starting from Eq. (4.19) he was able to show that stable orbits must have electron orbital angular momenta of

Condition for orbital stability
$$mvr = \frac{nh}{2\pi} \qquad n = 1, 2, 3, \ldots \tag{4.21}$$

Since the de Broglie electron wavelength is $\lambda = h/mv$, Eq. (4.21) is the same as Eq. (4.12), $n\lambda = 2\pi r$, which states that an electron orbit must contain an integral number of wavelengths.

4.7 NUCLEAR MOTION

The nuclear mass affects the wavelengths of spectral lines

Thus far we have been assuming that the hydrogen nucleus (a proton) remains stationary while the orbital electron revolves around it. What must actually happen, of course, is that both nucleus and electron revolve around their common center of mass, which is very close to the nucleus because the nuclear mass is much greater than that of the electron (Fig. 4.17). A system of this kind is equivalent to a single particle of mass m' that revolves around the position of the heavier particle. (This equivalence is

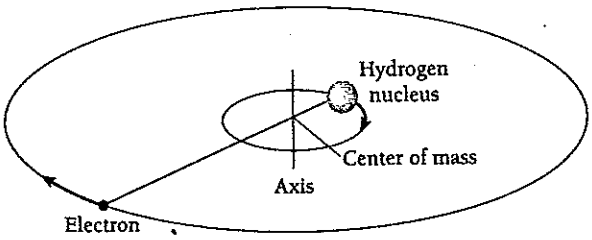


Figure 4.17 Both the electron and nucleus of a hydrogen atom revolve around a common center of mass (not to scale !).

demonstrated in Sec. 8.6.) If m is the electron mass and M the nuclear mass, then m' is given by

Reduced mass

$$m' = \frac{mM}{m + M} \tag{4.22}$$

The quantity m' is called the **reduced mass** of the electron because its value is less than m.

To take into account the motion of the nucleus in the hydrogen atom, then, all we need do is replace the electron with a particle of mass m'. The energy levels of the atom then become

Energy levels corrected for nuclear motion

$$E'_n = -\frac{m'e^4}{8\epsilon_0^2 h^2}\left(\frac{1}{n^2}\right) = \left(\frac{m'}{m}\right)\left(\frac{E_1}{n^2}\right) \tag{4.23}$$

Owing to motion of the nucleus, all the energy levels of hydrogen are changed by the fraction

$$\frac{m'}{m} = \frac{M}{M + m} = 0.99945$$

This represents an increase of 0.055 percent because the energies E_n, being smaller in absolute value, are therefore less negative.

The use of Eq. (4.23) in place of (4.15) removes a small but definite discrepancy between the predicted wavelengths of the spectral lines of hydrogen and the measured ones. The value of the Rydberg constant R to eight significant figures without correcting for nuclear motion is $1.0973731 \times 10^7\ \text{m}^{-1}$; the correction lowers it to $1.0967758 \times 10^7\ \text{m}^{-1}$.

The notion of reduced mass played an important part in the discovery of **deuterium**, a variety of hydrogen whose atomic mass is almost exactly double that of ordinary hydrogen because its nucleus contains a neutron as well as a proton. About one hydrogen atom in 6000 is a deuterium atom. Because of the greater nuclear mass, the spectral lines of deuterium are all shifted slightly to wavelengths shorter than the corresponding ones of ordinary hydrogen. Thus the H_α line of deuterium, which arises from a transition from the $n = 3$ to the $n = 2$ energy level, occurs at a wavelength of 656.1 nm, whereas the H_α line of hydrogen occurs at 656.3 nm. This difference in wavelength was responsible for the identification of deuterium in 1932 by the American chemist Harold Urey.

Example 4.6

A **positronium** "atom" is a system that consists of a positron and an electron that orbit each other. Compare the wavelengths of the spectral lines of positronium with those of ordinary hydrogen.

Solution

Here the two particles have the same mass m, so the reduced mass is

$$m' = \frac{mM}{m + M} = \frac{m^2}{2m} = \frac{m}{2}$$

where m is the electron mass. From Eq. (4.23) the energy levels of a positronium "atom" are

$$E'_n = \left(\frac{m'}{m}\right)\frac{E_1}{n^2} = \frac{E_1}{2n^2}$$

This means that the Rydberg constant—the constant term in Eq. (4.18)—for positronium is half as large as it is for ordinary hydrogen. As a result the wavelengths in the positronium spectral lines are all twice those of the corresponding lines in the hydrogen spectrum.

Example 4.7

A **muon** is an unstable elementary particle whose mass is $207m_e$ and whose charge is either $+e$ or $-e$. A negative muon (μ^-) can be captured by a nucleus to form a muonic atom. (*a*) A proton captures a μ^-. Find the radius of the first Bohr orbit of this atom. (*b*) Find the ionization energy of the atom.

Solution

(*a*) Here $m = 207m_e$ and $M = 1836m_e$, so the reduced mass is

$$m' = \frac{mM}{m+M} = \frac{(207m_e)(1836m_e)}{207m_e + 1836m_e} = 186m_e$$

According to Eq. (4.13) the orbit radius corresponding to $n = 1$ is

$$r_1 = \frac{h^2\epsilon_0}{\pi m_e e^2}$$

where $r_1 = a_0 = 5.29 \times 10^{-11}$ m. Hence the radius r' that corresponds to the reduced mass m' is

$$r'_1 = \left(\frac{m}{m'}\right) r_1 = \left(\frac{m_e}{186m_e}\right) a_0 = 2.85 \times 10^{-13}\ \text{m}$$

The muon is 186 times closer to the proton than an electron would be, so a muonic hydrogen atom is much smaller than an ordinary hydrogen atom.

(*b*) From Eq. (4.23) we have, with $n = 1$ and $E_1 = -13.6$ eV,

$$E'_1 = \left(\frac{m'}{m}\right) E_1 = 186E_1 = -2.53 \times 10^3\ \text{eV} = -2.53\ \text{keV}$$

The ionization energy is therefore 2.53 keV, 186 times that for an ordinary hydrogen atom.

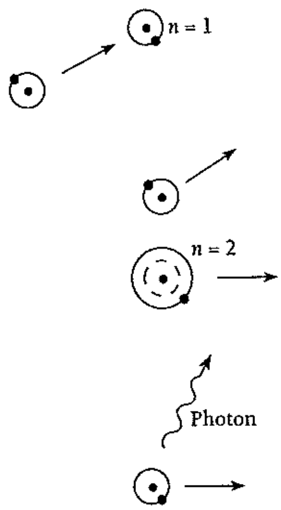


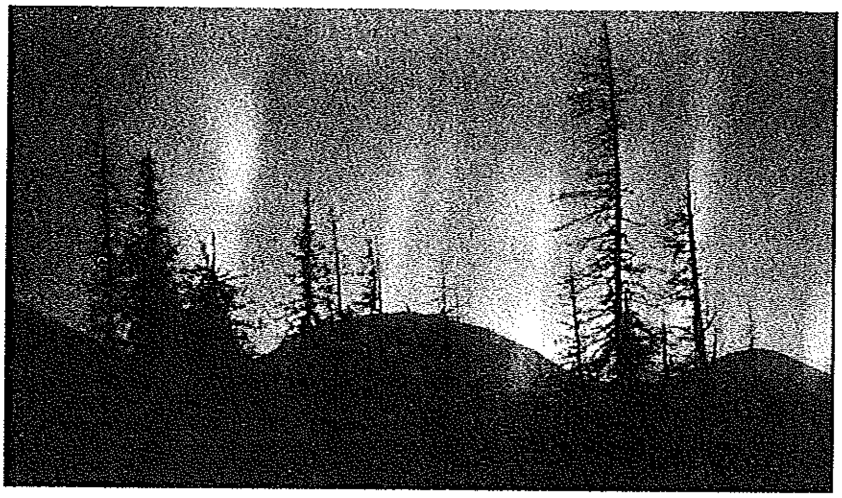
Figure 4.18 Excitation by collision. Some of the available energy is absorbed by one of the atoms, which goes into an excited energy state. The atom then emits a photon in returning to its ground (normal) state.

4.8 ATOMIC EXCITATION

How atoms absorb and emit energy

There are two main ways in which an atom can be excited to an energy above its ground state and thereby become able to radiate. One of these ways is by a collision with another particle in which part of their joint kinetic energy is absorbed by the atom. Such an excited atom will return to its ground state in an average of 10^{-8} s by emitting one or more photons (Fig. 4.18).

To produce a luminous discharge in a rarefied gas, an electric field is established that accelerates electrons and atomic ions until their kinetic energies are sufficient to



Auroras are caused by streams of fast protons and electrons from the sun that excite atoms in the upper atmosphere. The green hues of an auroral display come from oxygen, and the reds originate in both oxygen and nitrogen. This aurora occurred in Alaska.

excite atoms they collide with. Because energy transfer is a maximum when the colliding particles have the same mass (see Fig. 12.22), the electrons in such a discharge are more effective than the ions in providing energy to atomic electrons. Neon signs and mercury-vapor lamps are familiar examples of how a strong electric field applied between electrodes in a gas-filled tube leads to the emission of the characteristic spectral radiation of that gas, which happens to be reddish light in the case of neon and bluish light in the case of mercury vapor.

Another excitation mechanism is involved when an atom absorbs a photon of light whose energy is just the right amount to raise the atom to a higher energy level. For example, a photon of wavelength 121.7 nm is emitted when a hydrogen atom in the $n = 2$ state drops to the $n = 1$ state. Absorbing a photon of wavelength 121.7 nm by a hydrogen atom initially in the $n = 1$ state will therefore bring it up to the $n = 2$ state (Fig. 4.19). This process explains the origin of absorption spectra.

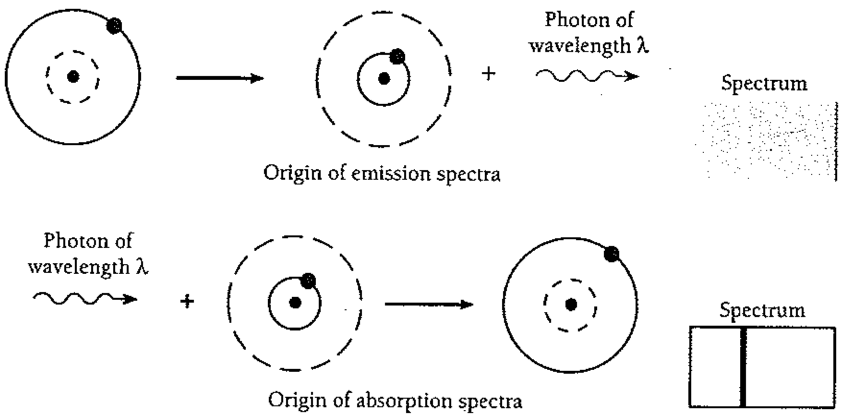


Figure 4.19 How emission and absorption spectral lines originate.

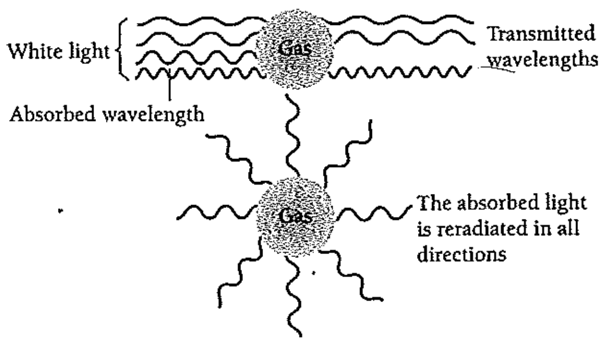


Figure 4.20 The dark lines in an absorption spectrum are never totally dark.

When white light, which contains all wavelengths, is passed through hydrogen gas, photons of those wavelengths that correspond to transitions between energy levels are absorbed. The resulting excited hydrogen atoms reradiate their excitation energy almost at once, but these photons come off in random directions with only a few in the same direction as the original beam of white light (Fig. 4.20). The dark lines in an absorption spectrum are therefore never completely black but only appear so by contrast with the bright background. We expect the lines in the absorption spectrum of any element to coincide with those in its emission spectrum that represent transitions to the ground state, which agrees with observation (see Fig. 4.9).

Franck-Hertz Experiment

Atomic spectra are not the only way to investigate energy levels inside atoms. A series of experiments based on excitation by collision was performed by James Franck and Gustav Hertz (a nephew of Heinrich Hertz) starting in 1914. These experiments demonstrated that atomic energy levels indeed exist and, furthermore, that the ones found in this way are the same as those suggested by line spectra.

Franck and Hertz bombarded the vapors of various elements with electrons of known energy, using an apparatus like that shown in Fig. 4.21. A small potential difference V_0 between the grid and collecting plate prevents electrons having energies less than a certain minimum from contributing to the current I through the ammeter. As the accelerating potential V is increased, more and more electrons arrive at the plate and I rises (Fig. 4.22).

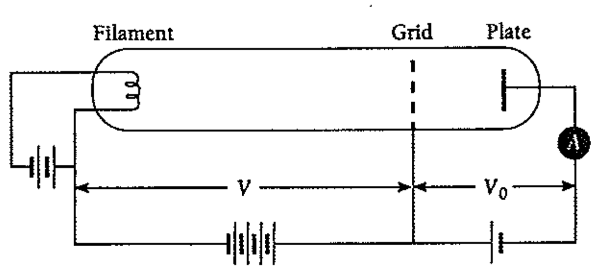


Figure 4.21 Apparatus for the Franck-Hertz experiment.

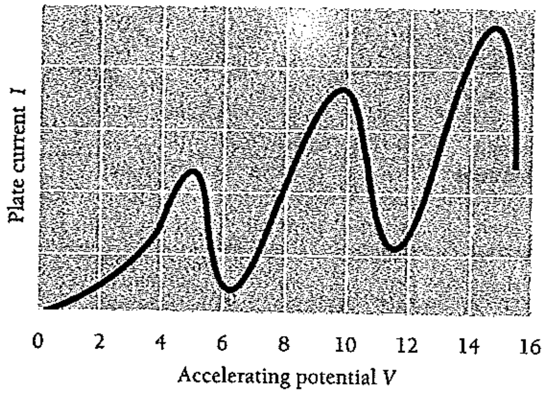


Figure 4.22 Results of the Franck-Hertz experiment, showing critical potentials in mercury vapor.

If KE is conserved when an electron collides with one of the atoms in the vapor, the electron merely bounces off in a new direction. Because an atom is much heavier than an electron, the electron loses almost no KE in the process. After a certain critical energy is reached, however, the plate current drops abruptly. This suggests that an electron colliding with one of the atoms gives up some or all of its KE to excite the atom to an energy level above its ground state. Such a collision is called inelastic, in contrast to an elastic collision in which KE is conserved. The critical electron energy equals the energy needed to raise the atom to its lowest excited state.

Then, as the accelerating potential V is raised further, the plate current again increases, since the electrons now have enough energy left to reach the plate after undergoing an inelastic collision on the way. Eventually another sharp drop in plate current occurs, which arises from the excitation of the same energy level in other atoms by the electrons. As Fig. 4.22 shows, a series of critical potentials for a given atomic vapor is obtained. Thus the higher potentials result from two or more inelastic collisions and are multiples of the lowest one.

To check that the critical potentials were due to atomic energy levels, Franck and Hertz observed the emission spectra of vapors during electron bombardment. In the case of mercury vapor, for example, they found that a minimum electron energy of 4.9 eV was required to excite the 253.6-nm spectral line of mercury—and a photon of 253.6-nm light has an energy of just 4.9 eV. The Franck-Hertz experiments were performed shortly after Bohr announced his theory of the hydrogen atom, and they independently confirmed his basic ideas.

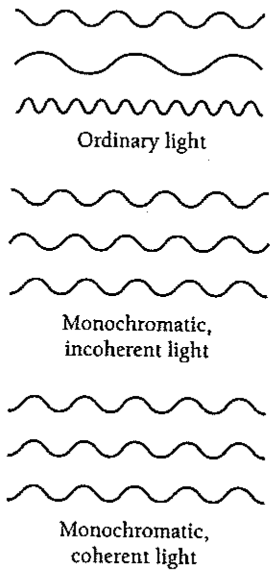


Figure 4.23 A laser produces a beam of light whose waves all have the same frequency (monochromatic) and are in phase with one another (coherent). The beam is also well collimated and so spreads out very little, even over long distances.

4.9 THE LASER

How to produce light waves all in step

The **laser** is a device that produces a light beam with some remarkable properties:

1 The light is very nearly monochromatic.

2 The light is coherent, with the waves all exactly in phase with one another (Fig.4.23).

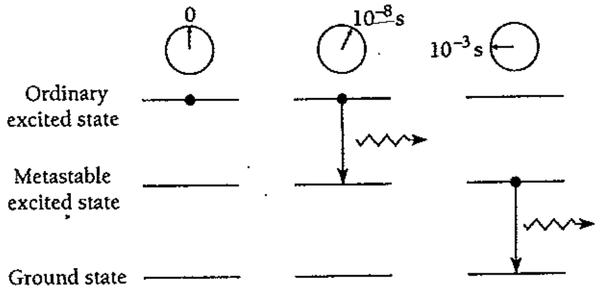


Figure 4.24 An atom can exist in a metastable energy level for a longer time before radiating than it can in an ordinary energy level.

3 A laser beam diverges hardly at all. Such a beam sent from the earth to a mirror left on the moon by the Apollo 11 expedition remained narrow enough to be detected on its return to the earth, a total distance of over three-quarters of a million kilometers. A light beam produced by any other means would have spread out too much for this to be done.
4 The beam is extremly intense, more intense by far than the light from any other source. To achieve an energy density equal to that in some laser beams, a hot object would have to be at a temperature of 10^{30} K.

The last two of these properties follow from the second of them.

The term *laser* stands for *l*ight *a*mplification by *s*timulated *e*mission of *r*adiation. The key to the laser is the presence in many atoms of one or more excited energy levels whose lifetimes may be 10^{-3} s or more instead of the usual 10^{-8} s. Such relatively long-lived states are called **metastable** (temporarily stable); see Fig. 4.24.

Three kinds of transition involving electromagnetic radiation are possible between two energy levels, E_0 and E_1, in an atom (Fig. 4.25). If the atom is initially in the lower state E_0, it can be raised to E_1 by absorbing a photon of energy $E_1 - E_0 = h\nu$. This process is called **stimulated absorption.** If the atom is initially in the upper state E_1, it can drop to E_0 by emitting a photon of energy $h\nu$. This is **spontaneous emission.**

Einstein, in 1917, was the first to point out a third possibility, **stimulated emission**, in which an incident photon of energy $h\nu$ causes a transition from E_1 to E_0. In stimulated emission, the radiated light waves are exactly in phase with the incident ones, so the result is an enhanced beam of coherent light. Einstein showed that stimulated emission has the same probability as stimulated absorption (see Sec. 9.7). That is, a photon of energy $h\nu$ incident on an atom in the upper

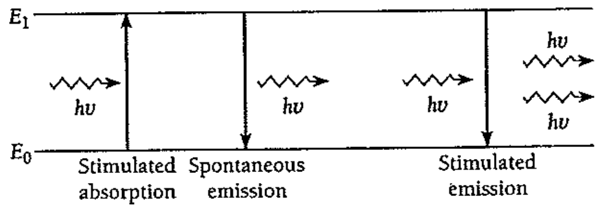


Figure 4.25 Transitions between two energy levels in an atom can occur by stimulated absorption, spontaneous emission, and stimulated emission.

Charles H. Townes (1915–) was born in Greenville, South Carolina, and attended Furman University there. After graduate study at Duke University and the California Institute of Technology, he spent 1939 to 1947 at the Bell Telephone Laboratories designing radar-controlled bombing systems. Townes then joined the physics department of Columbia University. In 1951, while sitting on a park bench, the idea for the **maser** (*m*icrowave *a*mplification by *s*timulated *e*mission of *r*adiation) occurred to him as a way to produce high-intensity microwaves, and in 1953 the first maser began operating. In this device ammonia (NH_3) molecules were raised to an excited vibrational state and then fed into a resonant cavity where, as in a laser, stimulated emission produced a cascade of photons of identical wavelength, here 1.25 cm in the microwave part of the spectrum. "Atomic clocks" of great accuracy are based on this concept, and solid-state maser amplifiers are used in such applications as radioastronomy.

In 1958 Townes and Arthur Schawlow attracted much attention with a paper showing that a similar scheme ought to be possible at optical wavelengths. Slightly earlier Gordon Gould, then a graduate student at Columbia, had come to the same conclusion, but did not publish his calculations at once since that would prevent securing a patent. Gould tried to develop the laser—his term—in private industry, but the Defense Department classified as secret the project (and his original notebooks) and denied him clearance to work on it. Finally, twenty years later, Gould succeeded in establishing his priority and received two patents on the laser, and still later, a third. The first working laser was built by Theodore Maiman at Hughes Research Laboratories in 1960. In 1964 Townes, along with two Russian laser pioneers, Aleksander Prokhorov and Nikolai Basov, was awarded a Nobel Prize. In 1981 Schawlow shared a Nobel Prize for precision spectroscopy using lasers.

Soon after its invention, the laser was spoken of as a "solution looking for a problem" because few applications were then known for it. Today, of course, lasers are widely employed for a variety of purposes.

state E_1 has the same likelihood of causing the emission of another photon of energy $h\nu$ as its likelihood of being absorbed if it is incident on an atom in the lower state E_0.

Stimulated emission involves no novel concepts. An analogy is a harmonic oscillator, for instance a pendulum, which has a sinusoidal force applied to it whose period is the same as its natural period of vibration. If the applied force is exactly in phase with the pendulum swings, the amplitude of the swings increases. This corresponds to stimulated absorption. However, if the applied force is 180° out of phase with the pendulum swings, the amplitude of the swings *decreases*. This corresponds to stimulated emission.

A **three-level laser**, the simplest kind, uses an assembly of atoms (or molecules) that have a metastable state $h\nu$ in energy above the ground state and a still higher excited state that decays to the metastable state (Fig. 4.26). What we want is more atoms in the metastable state than in the ground state. If we can arrange this and then shine light of frequency ν on the assembly, there will be more stimulated emissions from atoms in the metastable state than stimulated absorptions by atoms in the ground state. The result will be an amplification of the original light. This is the concept that underlies the operation of the laser.

The term **population inversion** describes an assembly of atoms in which the majority are in energy levels above the ground state; normally the ground state is occupied to the greatest extent.

A number of ways exist to produce a population inversion. One of them, called **optical pumping**, is illustrated in Fig. 4.27. Here an external light source is used some of whose photons have the right frequency to raise ground-state atoms to the excited state that decays spontaneously to the desired metastable state.

Why are three levels needed? Suppose there are only two levels, a metastable state $h\nu$ above the ground state. The more photons of frequency ν we pump into the assembly

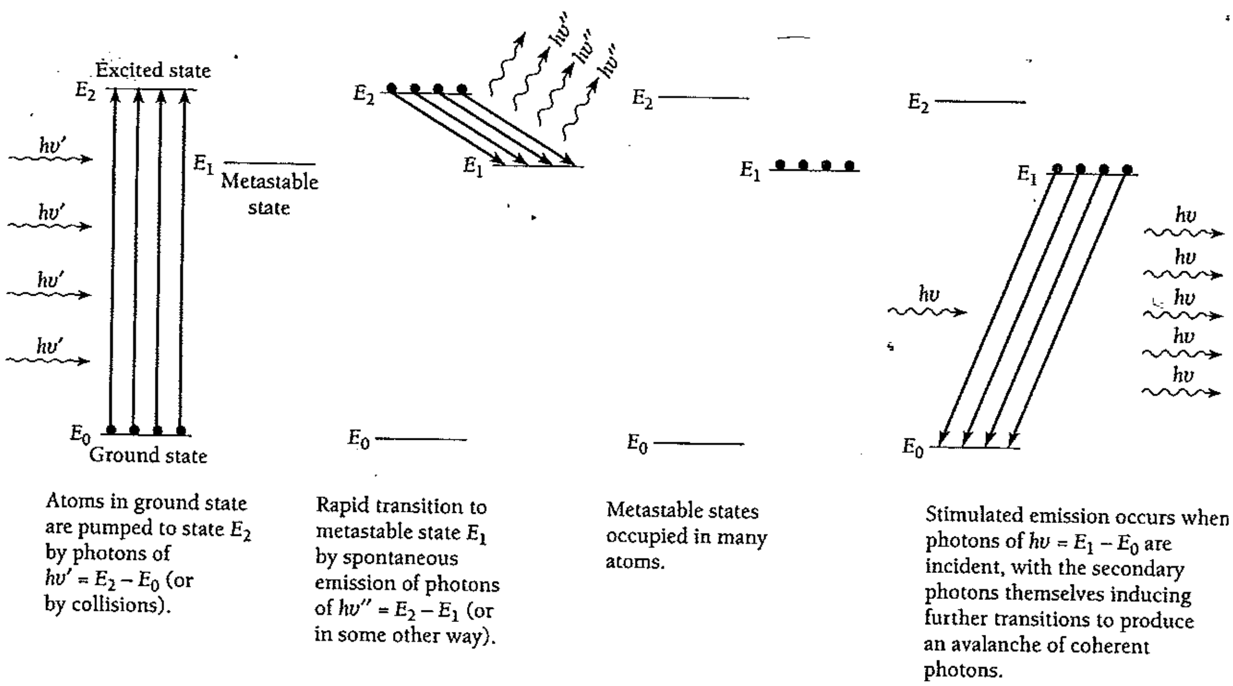


Figure 4.26 The principle of the laser.

of atoms, the more upward transitions there will be from the ground state to the metastable state. However, at the same time the pumping will stimulate downward transitions from the metastable state to the ground state. When half the atoms are in each state, the rate of stimulated emissions will equal the rate of stimulated absorptions, so the assembly cannot ever have more than half its atoms in the metastable state. In this situation laser amplification cannot occur. A population inversion is only possible when the stimulated absorptions are to a higher energy level than the metastable one from which the stimulated emission takes place, which prevents the pumping from depopulating the metastable state.

In a three-level laser, more than half the atoms must be in the metastable state for stimulated induced emission to predominate. This is not the case for a **four-level laser.**

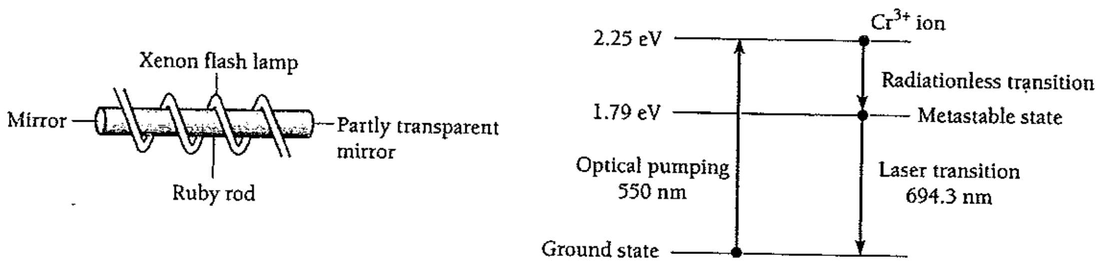


Figure 4.27 The ruby laser. In order for stimulated emission to exceed stimulated absorption, more than half the Cr^{3+} ions in the ruby rod must be in the metastable state. This laser produces a pulse of red light after each flash of the lamp.

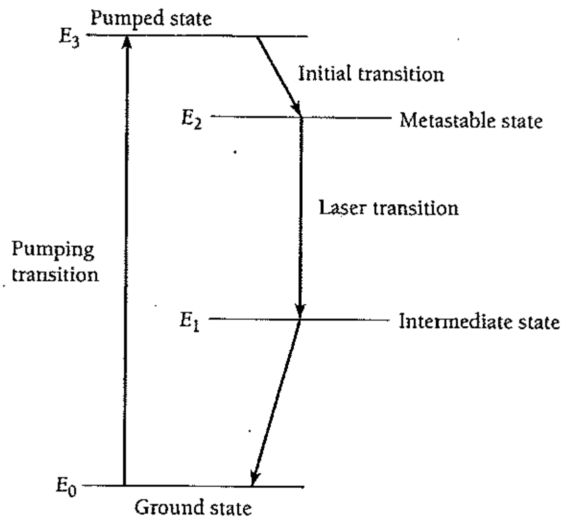
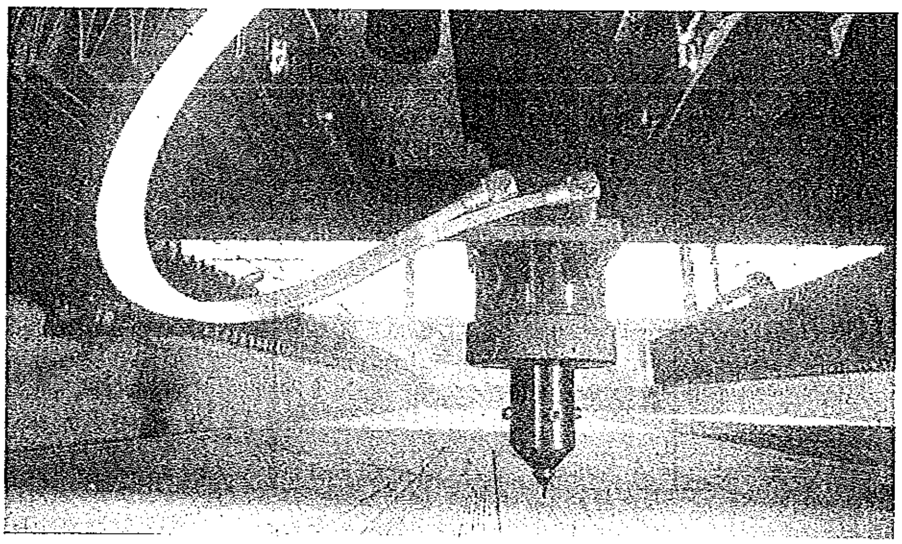


Figure 4.28 A four-level laser.

As in Fig. 4.28, the laser transition from the metastable state ends at an unstable intermediate state rather than at the ground state. Because the intermediate state decays rapidly to the ground state, very few atoms are in the intermediate state. Hence even a modest amount of pumping is enough to populate the metastable state to a greater extent than the intermediate state, as required for laser amplification.

Practical Lasers

The first successful laser, the **ruby laser,** is based on the three energy levels in the chromium ion Cr^{3+} shown in Fig. 4.27. A ruby is a crystal of aluminum oxide, Al_2O_3,



A robot arm carries a laser for cutting fabric in a clothing factory.

in which some of the Al^{3+} ions are replaced by Cr^{3+} ions, which are responsible for the red color. A Cr^{3+} ion has a metastable level whose lifetime is about 0.003 s. In the ruby laser, a xenon flash lamp excites the Cr^{3+} ions to a level of higher energy from which they fall to the metastable level by losing energy to other ions in the crystal. Photons from the spontaneous decay of some Cr^{3+} ions are reflected back and forth between the mirrored ends of the ruby rod, stimulating other excited Cr^{3+} ions to radiate. After a few microseconds the result is a large pulse of monochromatic, coherent red light from the partly transparent end of the rod.

The rod's length is made precisely an integral number of half-wavelengths long, so the radiation trapped in it forms an optical standing wave. Since the stimulated emissions are induced by the standing wave, their waves are all in step with it.

The common **helium-neon gas laser** achieves a population inversion in a different way. A mixture of about 10 parts of helium and 1 part of neon at a low pressure (~1 torr) is placed in a glass tube that has parallel mirrors, one of them partly transparent, at both ends. The spacing of the mirrors is again (as in all lasers) equal to an integral number of half-wavelengths of the laser light. An electric discharge is produced in the gas by means of electrodes outside the tube connected to a source of high-frequency alternating current, and collisions with electrons from the discharge excite He and Ne atoms to metastable states respectively 20.61 and 20.66 eV above their ground states (Fig. 4.29). Some of the excited He atoms transfer their energy to ground-state Ne atoms in collisions, with the 0.05 eV of additional energy being provided by the kinetic energy of the atoms. The purpose of the He atoms is thus to help achieve a population inversion in the Ne atoms.

The laser transition in Ne is from the metastable state at 20.66 eV to an excited state at 18.70 eV, with the emission of a 632.8-nm photon. Then another photon is spontaneously emitted in a transition to a lower metastable state; this transition yields only incoherent light. The remaining excitation energy is lost in collisions with the tube walls. Because the electron impacts that excite the He and Ne atoms occur all the time, unlike the pulsed excitation from the xenon flash lamp in a ruby laser, a He-Ne laser operates continuously. This is the laser whose narrow red beam is used in supermarkets to read bar codes. In a He-Ne laser, only a tiny

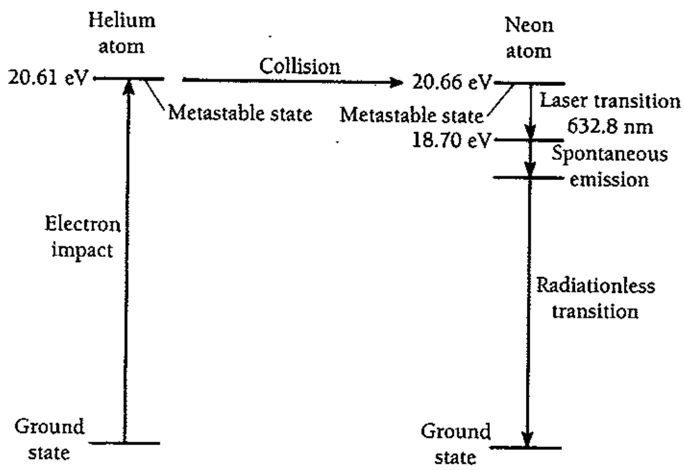


Figure 4.29 The helium-neon laser. In a four-level laser such as this, continuous operation is possible. Helium-neon lasers are commonly used to read bar codes.

Chirped Pulse Amplification

The most powerful lasers are pulsed, which produces phenomenal outputs for very short periods. The petawatt (10^{15} W) threshold was crossed in 1996 with pulses less than a trillionth of a second long—not all that much energy per pulse, but at a rate of delivery over 1000 times that of the entire electrical grid of the United States. An ingenious method called chirped pulse amplification made this possible without the laser apparatus itself being destroyed in the process. What was done was to start with a low-power laser pulse that was quite short, only 0.1 picosecond (10^{-13} s). Because the pulse was short, it consisted of a large span of wavelengths, as discussed in Sec. 3.7 (see Figs. 3.13 and 3.14). A diffraction grating then spread out the light into different paths according to wavelength, which stretched the pulse to 3 nanoseconds (3×10^{-9} s), 30,000 times longer. The result was to decrease the peak power so that laser amplifiers could boost the energy of each beam. Finally the amplified beams, each of slightly different wavelength, were recombined by another grating to produce a pulse less than 0.5 picoseconds long whose power was 1.3 petawatts.

fraction (one in millions) of the atoms present participates in the laser process at any moment.

Many other types of laser have been devised. A number of them employ molecules rather than atoms. **Chemical lasers** are based on the production by chemical reactions of molecules in metastable excited states. Such lasers are efficient and can be very powerful: one chemical laser, in which hydrogen and fluorine combine to form hydrogen fluoride, has generated an infrared beam of over 2 MW. **Dye lasers** use dye molecules whose energy levels are so close together that they can "lase" over a virtually continuous range of wavelengths (see Sec. 8.7). A dye laser can be tuned to any desired wavelength in its range. **Nd:YAG lasers,** which use the glassy solid yttrium aluminum garnet with neodymium as an impurity, are helpful in surgery because they seal small blood vessels while cutting through tissue by vaporizing water in the path of their beams. Powerful **carbon dioxide gas lasers** with outputs up to many kilowatts are used industrially for the precise cutting of almost any material, including steel, and for welding.

Tiny **semiconductor lasers** by the million process and transmit information today. (How such lasers work is described in Chap. 10.) In a compact disk player, a semiconductor laser beam is focused to a spot a micrometer (10^{-6} m) across to read data coded as pits that appear as dark spots on a reflective disk 12 cm in diameter. A compact disk can store over 600 megabytes of digital data, about 1000 times as much as the floppy disks used in personal computers. If the stored data is digitized music, the playing time can be over an hour.

Semiconductor lasers are ideal for fiber-optic transmission lines in which the electric signals that would normally be sent along copper wires are first converted into a series of pulses according to a standard code. Lasers then turn the pulses into flashes of infrared light that travel along thin (5–50 μm diameter) glass fibers and at the other end are changed back into electric signals. Over a million telephone conversations can be carried by a single fiber; by contrast, no more than 32 conversations can be carried at the same time by a pair of wires. Telephone fiber-optic systems today link many cities and exchanges within cities everywhere, and fiber-optic cables span the world's seas and oceans.

Appendix to Chapter 4

Rutherford Scattering

Rutherford's model of the atom was accepted because he was able to arrive at a formula to describe the scattering of alpha particles by thin foils on the basis of this model that agreed with the experimental results. He began by assuming that the alpha particle and the nucleus it interacts with are both small enough to be considered as point masses and charges; that the repulsive electric force between alpha particle and nucleus (which are both positively charged) is the only one acting; and that the nucleus is so massive compared with the alpha particle that it does not move during their interaction. Let us see how these assumptions lead to Eq. (4.1).

Scattering Angle

Owing to the variation of the electric force with $1/r^2$, where r is the instantaneous separation between alpha particle and nucleus, the alpha particle's path is a hyperbola with the nucleus at the outer focus (Fig. 4.30). The **impact parameter** b is the minimum distance to which the alpha particle would approach the nucleus if there were no force between them, and the **scattering angle** θ is the angle between the asymptotic direction of approach of the alpha particle and the asymptotic direction in which it recedes. Our first task is to find a relationship between b and θ.

As a result of the impulse $\int \mathbf{F}\, dt$ given it by the nucleus, the momentum of the alpha particle changes by $\Delta\mathbf{p}$ from the initial value $\mathbf{p}_1$ to the final value $\mathbf{p}_2$. That is,

$$\Delta p = p_2 - p_1 = \int F\, dt \tag{4.24}$$

Because the nucleus remains stationary during the passage of the alpha particle, by hypothesis, the alpha-particle kinetic energy is the same before and after the scattering. Hence the *magnitude* of its momentum is also the same before and after, and

$$p_1 = p_2 = mv$$

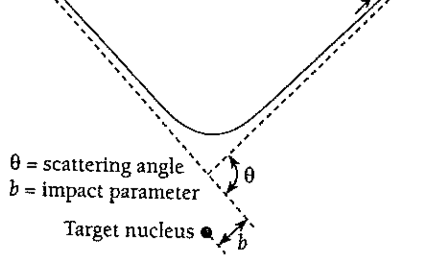


Figure 4.30 Rutherford scattering.

Here v is the alpha-particle velocity far from the nucleus.

From Fig. 4.31 we see that according to the law of sines,

$$\frac{\Delta p}{\sin\theta} = \frac{mv}{\sin\dfrac{\pi-\theta}{2}}$$

Since

$$\sin\frac{1}{2}(\pi-\theta) = \cos\frac{\theta}{2}$$

and

$$\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$$

we have for the magnitude of the momentum change

$$\Delta p = 2mv\sin\frac{\theta}{2} \tag{4.25}$$

Because the impulse $\int \mathbf{F}\,dt$ is in the same direction as the momentum change $\Delta\mathbf{p}$, its magnitude is

$$\left|\int \mathbf{F}\,dt\right| = \int F\cos\phi\,dt \tag{4.26}$$

where ϕ is the instantaneous angle between $\mathbf{F}$ and Δp along the path of the alpha particle. Inserting Eqs. (4.25) and (4.26) in Eq. (4.24),

$$2mv\sin\frac{\theta}{2} = \int_{-\infty}^{\infty} F\cos\phi\,dt$$

To change the variable on the right-hand side from t to ϕ, we note that the limits of integration will change to $-\frac{1}{2}(\pi-\theta)$ and $+\frac{1}{2}(\pi-\theta)$, corresponding to ϕ at $t=-\infty$ and $t=\infty$ respectively, and so

$$2mv\sin\frac{\theta}{2} = \int_{-(\pi-\theta)/2}^{+(\pi-\theta)/2} F\cos\phi\,\frac{dt}{d\phi}\,d\phi \tag{4.27}$$

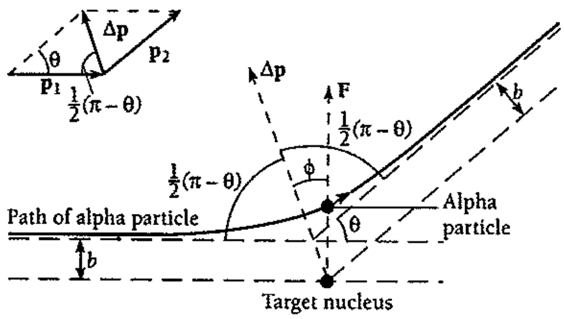


Figure 4.31 Geometrical relationships in Rutherford scattering.

The quantity $d\phi/dt$ is just the angular velocity ω of the alpha particle about the nucleus (this is evident from Fig. 4.31).

The electric force exerted by the nucleus on the alpha particle acts along the radius vector joining them, so there is no torque on the alpha particle and its angular momentum $m\omega r^2$ is constant. Hence

$$m\omega r^2 = \text{constant} = mr^2\,\frac{d\phi}{dt} = mvb$$

from which we obtain

$$\frac{dt}{d\phi} = \frac{r^2}{vb}$$

Substituting this expression for $dt/d\phi$ in Eq. (4.27) gives

$$2mv^2b\,\sin\frac{\theta}{2} = \int_{-(\pi-\theta)/2}^{+(\pi-\theta)/2} Fr^2\cos\phi\,d\phi \tag{4.28}$$

As we recall, F is the electric force exerted by the nucleus on the alpha particle. The charge on the nucleus is Ze, corresponding to the atomic number Z, and that on the alpha particle is $2e$. Therefore

$$F = \frac{1}{4\pi\epsilon_0}\frac{2Ze^2}{r^2}$$

and

$$\frac{4\pi\epsilon_0 mv^2b}{Ze^2}\sin\frac{\theta}{2} = \int_{-(\pi-\theta)/2}^{+(\pi-\theta)/2}\cos\phi\,d\phi = 2\cos\frac{\theta}{2}$$

The scattering angle θ is related to the impact parameter b by the equation

$$\cot\frac{\theta}{2} = \frac{2\pi\epsilon_0 mv^2}{Ze^2}\,b$$

It is more convenient to specify the alpha-particle energy KE instead of its mass and velocity separately; with this substitution,

Scattering angle

$$\cot\frac{\theta}{2} = \frac{4\pi\epsilon_0\text{KE}}{Ze^2}\,b \tag{4.29}$$

Figure 4.32 is a schematic representation of Eq. (4.29); the rapid decrease in θ as b increases is evident. A very near miss is required for a substantial deflection.

Rutherford Scattering Formula

Equation (4.29) cannot be directly confronted with experiment because there is no way of measuring the impact parameter corresponding to a particular observed scattering angle. An indirect strategy is required.

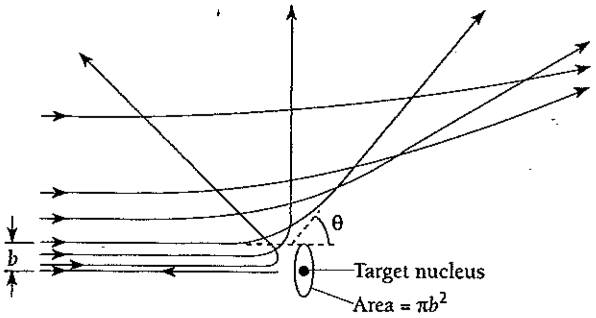


Figure 4.32 The scattering angle decreases with increasing impact parameter.

Our first step is to note that all alpha particles approaching a target nucleus with an impact parameter from 0 to b will be scattered through an angle of θ or more, where θ is given in terms of b by Eq. (4.29). This means that an alpha particle that is initially directed anywhere within the area πb^2 around a nucleus will be scattered through θ or more (Fig. 4.32). The area πb^2 is accordingly called the **cross section** for the interaction. The general symbol for cross section is σ, and so here

Cross section $$\sigma = \pi b^2 \tag{4.30}$$

Of course, the incident alpha particle is actually scattered before it reaches the immediate vicinity of the nucleus and hence does not necessarily pass within a distance b of it.

Now we consider a foil of thickness t that contains n atoms per unit volume. The number of target nuclei per unit area is nt, and an alpha-particle beam incident upon an area A therefore encounters ntA nuclei. The aggregate cross section for scatterings of θ or more is the number of target nuclei ntA multiplied by the cross section σ for such scattering per nucleus, or $ntA\sigma$. Hence the fraction f of incident alpha particles scattered by θ or more is the ratio between the aggregate cross section $ntA\sigma$ for such scattering and the total target area A. That is,

$$\begin{aligned} f &= \frac{\text{alpha particles scattered by } \theta \text{ or more}}{\text{incident alpha particles}} \\ &= \frac{\text{aggregate cross section}}{\text{target area}} = \frac{ntA\sigma}{A} \\ &= nt\pi b^2 \end{aligned}$$

Substituting for b from Eq. (4.30),

$$f = \pi nt \left(\frac{Ze^2}{4\pi\epsilon_0 \text{KE}} \right)^2 \cot^2 \frac{\theta}{2} \tag{4.31}$$

In this calculation it was assumed that the foil is sufficiently thin so that the cross sections of adjacent nuclei do not overlap and that a scattered alpha particle receives its entire deflection from an encounter with a single nucleus.

Example 4.8

Find the fraction of a beam of 7.7-MeV alpha particles that is scattered through angles of more than 45° when incident upon a gold foil 3×10^{-7} m thick. These values are typical of the alpha-particle energies and foil thicknesses used by Geiger and Marsden. For comparison, a human hair is about 10^{-4} m in diameter.

Solution

We begin by finding n, the number of gold atoms per unit volume in the foil, from the relationship

$$n = \frac{\text{atoms}}{\text{m}^3} = \frac{\text{mass/m}^3}{\text{mass/atom}}$$

Since the density of gold is 1.93×10^4 kg/m^3, its atomic mass is 197 u, and 1 u $= 1.66 \times 10^{-27}$ kg, we have

$$n = \frac{1.93 \times 10^4 \text{ kg/m}^3}{(197 \text{ u/atom})(1.66 \times 10^{-27} \text{ kg/u})}$$
$$= 5.90 \times 10^{28} \text{ atoms/m}^3$$

The atomic number Z of gold is 79, a kinetic energy of 7.7 MeV is equal to 1.23×10^{-12} J, and $\theta = 45°$; from these figures we find that

$$f = 7 \times 10^{-5}$$

of the incident alpha particles are scattered through 45° or more—only 0.007 percent! A foil this thin is quite transparent to alpha particles.

In an actual experiment, a detector measures alpha particles scattered between θ and $\theta + d\theta$, as in Fig. 4.33. The fraction of incident alpha particles so scattered is found by differentiating Eq. (4.31) with respect to θ, which gives

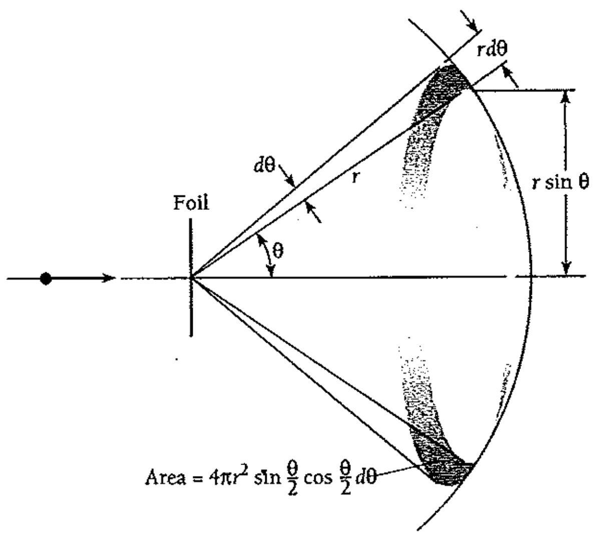


Figure 4.33 In the Rutherford experiment, particles are detected that have been scattered between θ and $\theta + d\theta$.

$$df = -\pi nt\left(\frac{Ze^2}{4\pi\epsilon_0\mathrm{KE}}\right)^2 \cot\frac{\theta}{2}\csc^2\frac{\theta}{2}\,d\theta \tag{4.32}$$

The minus sign expresses the fact that f decreases with increasing θ.

As we saw in Fig. 4.2, Geiger and Marsden placed a fluorescent screen a distance r from the foil and the scattered alpha particles were detected by means of the scintillations they caused. Those alpha particles scattered between θ and $\theta + d\theta$ reached a zone of a sphere of radius r whose width is $r\,d\theta$. The zone radius itself is $r \sin\theta$, and so the area dS of the screen struck by these particles is

$$\begin{aligned} dS &= (2\pi r \sin\theta)(r\,d\theta) = 2\pi r^2 \sin\theta\,d\theta \\ &= 4\pi r^2 \sin\frac{\theta}{2}\cos\frac{\theta}{2}\,d\theta \end{aligned}$$

If a total of N_i alpha particles strike the foil during the course of the experiment, the number scattered into $d\theta$ at θ is $N_i df$. The number $N(\theta)$ per unit area striking the screen at θ, which is the quantity actually measured, is

$$N(\theta) = \frac{N_i|df|}{dS} = \frac{N_i\pi nt\left(\dfrac{Ze^2}{4\pi\epsilon_0\mathrm{KE}}\right)^2\cot\dfrac{\theta}{2}\csc^2\dfrac{\theta}{2}\,d\theta}{4\pi r^2\sin\dfrac{\theta}{2}\cos\dfrac{\theta}{2}\,d\theta}$$

Rutherford scattering formula

$$N(\theta) = \frac{N_i n t Z^2 e^4}{(8\pi\epsilon_0)^2 r^2\,\mathrm{KE}^2 \sin^4(\theta/2)} \tag{4.1}$$

Equation (4.1) is the Rutherford scattering formula. Figure 4.4 shows how $N(\theta)$ varies with θ.

EXERCISES

It isn't that they can't see the solution. It is that they can't see the problem. —Gilbert Chesterton

4.1 The Nuclear Atom

1. The great majority of alpha particles pass through gases and thin metal foils with no deflections. To what conclusion about atomic structure does this observation lead?

2. The electric field intensity at a distance r from the center of a uniformly charged sphere of radius R and total charge Q is $Qr/4\pi\epsilon_0 R^3$ when $r < R$. Such a sphere corresponds to the Thomson model of the atom. Show that an electron in this sphere executes simple harmonic motion about its center and derive a formula for the frequency of this motion. Evaluate the frequency of the electron oscillations for the case of the hydrogen atom and compare it with the frequencies of the spectral lines of hydrogen.

3. Determine the distance of closest approach of 1.00-MeV protons incident on gold nuclei.

4.2 Electron Orbits

4. Find the frequency of revolution of the electron in the classical model of the hydrogen atom. In what region of the spectrum are electromagnetic waves of this frequency?

4.3 Atomic Spectra

5. What is the shortest wavelength present in the Brackett series of spectral lines?

6. What is the shortest wavelength present in the Paschen series of spectral lines?

4.4 The Bohr Atom

7. In the Bohr model, the electron is in constant motion. How can such an electron have a negative amount of energy?

8. Lacking de Broglie's hypothesis to guide his thinking, Bohr arrived at his model by postulating that the angular momentum of an orbital electron must be an integral multiple of $\hbar$. Show that this postulate leads to Eq. (4.13).

9. The **fine structure constant** is defined as $\alpha = e^2/2\epsilon_0 hc$. This quantity got its name because it first appeared in a theory by the German physicist Arnold Sommerfeld that tried to explain the fine structure in spectral lines (multiple lines close together instead of single lines) by assuming that elliptical as well as circular orbits are possible in the Bohr model. Sommerfeld's approach was on the wrong track, but α has nevertheless turned out to be a useful quantity in atomic physics. (*a*) Show that $\alpha = v_1/c$, where v_1 is the velocity of the electron in the ground state of the Bohr atom. (*b*) Show that the value of α is very close to 1/137 and is a pure number with no dimensions. Because the magnetic behavior of a moving charge depends on its velocity, the small value of α is representative of the relative magnitudes of the magnetic and electric aspects of electron behavior in an atom. (*c*) Show that $\alpha a_0 = \lambda_C/2\pi$, where a_0 is the radius of the ground-state Bohr orbit and λ_C is the Compton wavelength of the electron.

10. An electron at rest is released far away from a proton, toward which it moves. (*a*) Show that the de Broglie wavelength of the electron is proportional to $\sqrt{r}$, where r is the distance of the electron from the proton. (*b*) Find the wavelength of the electron when it is a_0 from the proton. How does this compare with the wavelength of an electron in a ground-state Bohr orbit? (*c*) In order for the electron to be captured by the proton to form a ground-state hydrogen atom, energy must be lost by the system. How much energy?

11. Find the quantum number that characterizes the earth's orbit around the sun. The earth's mass is 6.0×10^{24} kg, its orbital radius is 1.5×10^{11} m, and its orbital speed is 3.0×10^4 m/s.

12. Suppose a proton and an electron were held together in a hydrogen atom by gravitational forces only. Find the formula for the energy levels of such an atom, the radius of its ground-state Bohr orbit, and its ionization energy in eV.

13. Compare the uncertainty in the momentum of an electron confined to a region of linear dimension a_0 with the momentum of an electron in a ground-state Bohr orbit.

4.5 Energy Levels and Spectra

14. When radiation with a continuous spectrum is passed through a volume of hydrogen gas whose atoms are all in the ground state, which spectral series will be present in the resulting absorption spectrum?

15. What effect would you expect the rapid random motion of the atoms of an excited gas to have on the spectral lines they produce?

16. A beam of 13.0-eV electrons is used to bombard gaseous hydrogen. What series of wavelengths will be emitted?

17. A proton and an electron, both at rest initially, combine to form a hydrogen atom in the ground state. A single photon is emitted in this process. What is its wavelength?

18. How many different wavelengths would appear in the spectrum of hydrogen atoms initially in the $n = 5$ state?

19. Find the wavelength of the spectral line that corresponds to a transition in hydrogen from the $n = 10$ state to the ground state. In what part of the spectrum is this?

20. Find the wavelength of the spectral line that corresponds to a transition in hydrogen from the $n = 6$ state to the $n = 3$ state. In what part of the spectrum is this?

21. A beam of electrons bombards a sample of hydrogen. Through what potential difference must the electrons have been accelerated if the first line of the Balmer series is to be emitted?

22. How much energy is required to remove an electron in the $n = 2$ state from a hydrogen atom?

23. The longest wavelength in the Lyman series is 121.5 nm and the shortest wavelength in the Balmer series is 364.6 nm. Use the figures to find the longest wavelength of light that could ionize hydrogen.

24. The longest wavelength in the Lyman series is 121.5 nm. Use this wavelength together with the values of c and h to find the ionization energy of hydrogen.

25. An excited hydrogen atom emits a photon of wavelength λ in returning to the ground state. (*a*) Derive a formula that gives the quantum number of the initial excited state in terms of λ and R. (*b*) Use this formula to find n_i for a 102.55-nm photon.

26. An excited atom of mass m and initial speed v emits a photon in its direction of motion. If $v \ll c$, use the requirement that linear momentum and energy must both be conserved to show that the frequency of the photon is higher by $\Delta\nu/\nu \approx v/c$ than it would have been if the atom had been at rest. (See also Exercise 16 of Chap. 1.)

27. When an excited atom emits a photon, the linear momentum of the photon must be balanced by the recoil momentum of the atom. As a result, some of the excitation energy of the atom goes into the kinetic energy of its recoil. (*a*) Modify Eq. (4.16) to include this effect. (*b*) Find the ratio between the recoil energy and the photon energy for the $n = 3 \longrightarrow n = 2$ transition in hydrogen, for which $E_f - E_i = 1.9$ eV. Is the effect a major one? A nonrelativistic calculation is sufficient here.

4.6 Correspondence Principle

28. Of the following quantities, which increase and which decrease in the Bohr model as n increases? Frequency of revolution, electron speed, electron wavelength, angular momentum, potential energy, kinetic energy, total energy.

29. Show that the frequency of the photon emitted by a hydrogen atom in going from the level $n + 1$ to the level n is always intermediate between the frequencies of revolution of the electron in the respective orbits.

4.7 Nuclear Motion

30. An antiproton has the mass of a proton but a charge of $-e$. If a proton and an antiproton orbited each other, how far apart would they be in the ground state of such a system? Why might you think such a system could not occur?

31. A μ^- muon is in the $n = 2$ state of a muonic atom whose nucleus is a proton. Find the wavelength of the photon emitted when the muonic atom drops to its ground state. In what part of the spectrum is this wavelength?

32. Compare the ionization energy in positronium with that in hydrogen.

33. A mixture of ordinary hydrogen and tritium, a hydrogen isotope whose nucleus is approximately 3 times more massive than ordinary hydrogen, is excited and its spectrum observed. How far apart in wavelength will the H_α lines of the two kinds of hydrogen be?

34. Find the radius and speed of an electron in the ground state of doubly ionized lithium and compare them with the radius and speed of the electron in the ground state of the hydrogen atom. (Li^{++} has a nuclear charge of $3e$.)

35. (*a*) Derive a formula for the energy levels of a **hydrogenic atom,** which is an ion such as He^+ or Li^{2+} whose nuclear charge is $+Ze$ and which contains a single electron. (*b*) Sketch the energy levels of the He^+ ion and compare them with the energy levels of the H atom. (*c*) An electron joins a bare helium nucleus to form a He^+ ion. Find the wavelength of the photon emitted in this process if the electron is assumed to have had no kinetic energy when it combined with the nucleus.

4.9 The Laser

36. For laser action to occur, the medium used must have at least three energy levels. What must be the nature of each of these levels? Why is three the minimum number?

37. A certain ruby laser emits 1.00-J pulses of light whose wavelength is 694 nm. What is the minimum number of Cr^{3+} ions in the ruby?

38. Steam at 100°C can be thought of as an excited state of water at 100°C. Suppose that a laser could be built based upon the transition from steam to water, with the energy lost per molecule of steam appearing as a photon. What would the frequency of such a photon be? To what region of the spectrum does this correspond? The heat of vaporization of water is 2260 kJ/kg and its molar mass is 18.02 kg/kmol.

Appendix: Rutherford Scattering

39. The Rutherford scattering formula fails to agree with the data at very small scattering angles. Can you think of a reason?

40. Show that the probability for a 2.0-MeV proton to be scattered by more than a given angle when it passes through a thin foil is the same as that for a 4.0-MeV alpha particle.

41. A 5.0-MeV alpha particle approaches a gold nucleus with an impact parameter of 2.6×10^{-13} m. Through what angle will it be scattered?

42. What is the impact parameter of a 5.0-MeV alpha particle scattered by 10° when it approaches a gold nucleus?

43. What fraction of a beam of 7.7-MeV alpha particles incident upon a gold foil 3.0×10^{-7} m thick is scattered by less than 1°?

44. What fraction of a beam of 7.7-MeV alpha particles incident upon a gold foil 3.0×10^{-7} m thick is scattered by 90° or more?

45. Show that twice as many alpha particles are scattered by a foil through angles between 60° and 90° as are scattered through angles of 90° or more.

46. A beam of 8.3-MeV alpha particles is directed at an aluminum foil. It is found that the Rutherford scattering formula ceases to be obeyed at scattering angles exceeding about 60°. If the alpha-particle radius is assumed small enough to neglect here, find the radius of the aluminum nucleus.

47. In special relativity, a photon can be thought of as having a "mass" of $m = E_\nu/c^2$. This suggests that we can treat a photon that passes near the sun in the same way as Rutherford treated an alpha particle that passes near a nucleus, with an attractive gravitational force replacing the repulsive electrical force. Adapt Eq. (4.29) to this situation and find the angle of deflection θ for a photon that passes $b = R_{sun}$ from the center of the sun. The mass and radius of the sun are respectively 2.0×10^{30} kg and 7.0×10^8 m. In fact, general relativity shows that this result is exactly half the actual deflection, a conclusion supported by observations made during solar eclipses as mentioned in Sec. 1.10.