forms a real, inverted image of a distant object very near the focal point of the eyepiece. Because the object is essentially at infinity, this point at which I_1 forms is the focal point of the objective. The eyepiece then forms, at I_2, an enlarged, inverted image of the image at I_1. To provide the largest possible magnification, the image distance for the eyepiece is infinite. Therefore, the image due to the objective lens, which acts as the object for the eyepiece lens, must be located at the focal point of the eyepiece. Hence, the two lenses are separated by a distance $f_o + f_e$, which corresponds to the length of the telescope tube.

The angular magnification of the telescope is given by θ/θ_o, where θ_o is the angle subtended by the object at the objective and θ is the angle subtended by the final image at the viewer's eye. Consider Figure 36.42a, in which the object is a very great distance to the left of the figure. The angle θ_o (to the *left* of the objective) subtended by the object at the objective is the same as the angle (to the *right* of the objective) subtended by the first image at the objective. Therefore,

$$\tan \theta_o \approx \theta_o \approx -\frac{h'}{f_o}$$

where the negative sign indicates that the image is inverted.

The angle θ subtended by the final image at the eye is the same as the angle that a ray coming from the tip of I_1 and traveling parallel to the principal axis makes with the principal axis after it passes through the lens. Therefore,

$$\tan \theta \approx \theta \approx \frac{h'}{f_e}$$

We have not used a negative sign in this equation because the final image is not inverted; the object creating this final image I_2 is I_1, and both it and I_2 point in the same direction. Therefore, the angular magnification of the telescope can be expressed as

$$m = \frac{\theta}{\theta_o} = \frac{h'/f_e}{-h'/f_o} = -\frac{f_o}{f_e} \tag{36.27}$$

This result shows that the angular magnification of a telescope equals the ratio of the objective focal length to the eyepiece focal length. The negative sign indicates that the image is inverted.

When you look through a telescope at such relatively nearby objects as the Moon and the planets, magnification is important. Individual stars in our galaxy, however, are so far away that they always appear as small points of light no matter how great the magnification. To gather as much light as possible, large research telescopes used to study very distant objects must have a large diameter. It is difficult and expensive to manufacture large lenses for refracting telescopes. Another difficulty with large lenses is that their weight leads to sagging, which is an additional source of aberration.

These problems associated with large lenses can be partially overcome by replacing the objective with a concave mirror, which results in the second type of telescope, the **reflecting telescope.** Because light is reflected from the mirror and does not pass through a lens, the mirror can have rigid supports on the back side. Such supports eliminate the problem of sagging.

Figure 36.43a shows the design for a typical reflecting telescope. The incoming light rays are reflected by a parabolic mirror at the base. These reflected rays converge toward point A in the figure, where an image would be formed. Before this image is formed, however, a small, flat mirror M reflects the light toward an opening in the tube's side and it passes into an eyepiece. This particular design is said to have a Newtonian focus because Newton developed it. Figure 36.43b shows such a telescope. Notice that the light never passes through glass (except through the small eyepiece) in the reflecting telescope. As a result, problems associated with chromatic aberration are virtually eliminated. The reflecting telescope can be made even shorter by orienting the flat mirror so that it reflects the light back

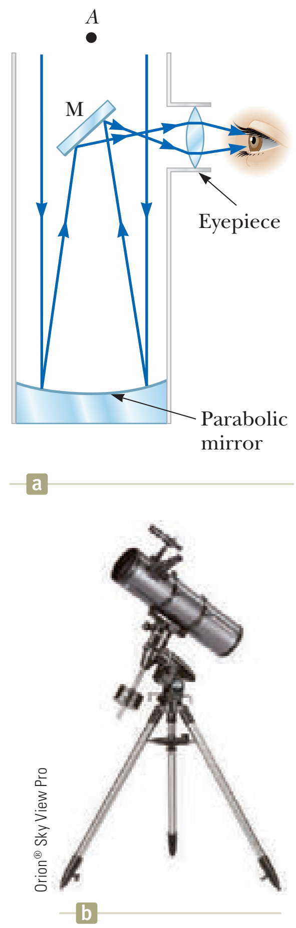
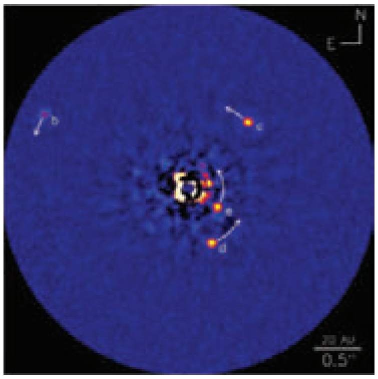


Figure 36.43 (a) A Newtonian-focus reflecting telescope. (b) A reflecting telescope. This type of telescope is shorter than that in Figure 36.42b.



National Research Council Canada, C. Marois & Keck Observatory

Figure 36.44 A direct optical image of a solar system around the star HR8799, developed at the Keck Observatory in Hawaii.

toward the objective mirror and the light enters an eyepiece in a hole in the middle of the mirror.

The largest reflecting telescopes in the world are at the Keck Observatory on Mauna Kea, Hawaii. The site includes two telescopes with diameters of 10 m, each containing 36 hexagonally shaped, computer-controlled mirrors that work together to form a large reflecting surface. In addition, the two telescopes can work together to provide a telescope with an effective diameter of 85 m. In contrast, the largest refracting telescope in the world, at the Yerkes Observatory in Williams Bay, Wisconsin, has a diameter of only 1 m.

Figure 36.44 shows a remarkable optical image from the Keck Observatory of a solar system around the star HR8799, located 129 light-years from the Earth. The planets labeled b, c, and d were seen in 2008 and the innermost planet, labeled e, was observed in December 2010. This photograph represents the first direct image of another solar system and was made possible by the adaptive optics technology used in the Keck Observatory.

Summary

Definitions

The **lateral magnification** M of the image due to a mirror or lens is defined as the ratio of the image height h' to the object height h. It is equal to the negative of the ratio of the image distance q to the object distance p:

$$M \equiv \frac{\text{image height}}{\text{object height}} = \frac{h'}{h} = -\frac{q}{p} \qquad \textbf{(36.1, 36.2, 36.17)}$$

The **angular magnification** m is the ratio of the angle subtended by an object with a lens in use (angle θ in Fig. 36.40b) to the angle subtended by the object placed at the near point with no lens in use (angle θ_0 in Fig. 36.40a):

$$m \equiv \frac{\theta}{\theta_0} \qquad \textbf{(36.22)}$$

The ratio of the focal length of a camera lens to the diameter of the lens is called the ***f*-number** of the lens:

$$f\text{-number} \equiv \frac{f}{D} \qquad \textbf{(36.20)}$$

Concepts and Principles

In the paraxial ray approximation, the object distance p and image distance q for a spherical mirror of radius R are related by the **mirror equation:**

$$\frac{1}{p} + \frac{1}{q} = \frac{2}{R} = \frac{1}{f} \qquad \textbf{(36.4, 36.6)}$$

where $f = R/2$ is the **focal length** of the mirror.

An image can be formed by refraction from a spherical surface of radius R. The object and image distances for refraction from such a surface are related by

$$\frac{n_1}{p} + \frac{n_2}{q} = \frac{n_2 - n_1}{R} \qquad \textbf{(36.8)}$$

where the light is incident in the medium for which the index of refraction is n_1 and is refracted in the medium for which the index of refraction is n_2.

The inverse of the **focal length** f of a thin lens surrounded by air is given by the **lens-makers' equation:**

$$\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \qquad \textbf{(36.15)}$$

Converging lenses have positive focal lengths, and **diverging lenses** have negative focal lengths.

For a thin lens, and in the paraxial ray approximation, the object and image distances are related by the **thin lens equation:**

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f} \qquad \textbf{(36.16)}$$

The maximum magnification of a single lens of focal length f used as a simple magnifier is

$$m_{\max} = 1 + \frac{25 \text{ cm}}{f} \qquad (36.24)$$

The overall magnification of the image formed by a compound microscope is

$$M = -\frac{L}{f_o}\left(\frac{25 \text{ cm}}{f_e}\right) \qquad (36.26)$$

where f_o and f_e are the focal lengths of the objective and eyepiece lenses, respectively, and L is the distance between the lenses.

The angular magnification of a refracting telescope can be expressed as

$$m = -\frac{f_o}{f_e} \qquad (36.27)$$

where f_o and f_e are the focal lengths of the objective and eyepiece lenses, respectively. The angular magnification of a reflecting telescope is given by the same expression where f_o is the focal length of the objective mirror.

Objective Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. The faceplate of a diving mask can be ground into a corrective lens for a diver who does not have perfect vision. The proper design allows the person to see clearly both under water and in the air. Normal eyeglasses have lenses with both the front and back surfaces curved. Should the lenses of a diving mask be curved (a) on the outer surface only, (b) on the inner surface only, or (c) on both surfaces?

2. Lulu looks at her image in a makeup mirror. It is enlarged when she is close to the mirror. As she backs away, the image becomes larger, then impossible to identify when she is 30.0 cm from the mirror, then upside down when she is beyond 30.0 cm, and finally small, clear, and upside down when she is much farther from the mirror. **(i)** Is the mirror (a) convex, (b) plane, or (c) concave? **(ii)** Is the magnitude of its focal length (a) 0, (b) 15.0 cm, (c) 30.0 cm, (d) 60.0 cm, or (e) ∞?

3. An object is located 50.0 cm from a converging lens having a focal length of 15.0 cm. Which of the following statements is true regarding the image formed by the lens? (a) It is virtual, upright, and larger than the object. (b) It is real, inverted, and smaller than the object. (c) It is virtual, inverted, and smaller than the object. (d) It is real, inverted, and larger than the object. (e) It is real, upright, and larger than the object.

4. **(i)** When an image of an object is formed by a converging lens, which of the following statements is *always* true? More than one statement may be correct. (a) The image is virtual. (b) The image is real. (c) The image is upright. (d) The image is inverted. (e) None of those statements is always true. **(ii)** When the image of an object is formed by a diverging lens, which of the statements is *always* true?

5. A converging lens in a vertical plane receives light from an object and forms an inverted image on a screen. An opaque card is then placed next to the lens, covering only the upper half of the lens. What happens to the image on the screen? (a) The upper half of the image disappears. (b) The lower half of the image disappears. (c) The entire image disappears. (d) The entire image is still visible, but is dimmer. (e) No change in the image occurs.

6. If Josh's face is 30.0 cm in front of a concave shaving mirror creating an upright image 1.50 times as large as the object, what is the mirror's focal length? (a) 12.0 cm (b) 20.0 cm (c) 70.0 cm (d) 90.0 cm (e) none of those answers

7. Two thin lenses of focal lengths f_1 = 15.0 and f_2 = 10.0 cm, respectively, are separated by 35.0 cm along a common axis. The f_1 lens is located to the left of the f_2 lens. An object is now placed 50.0 cm to the left of the f_1 lens, and a final image due to light passing though both lenses forms. By what factor is the final image different in size from the object? (a) 0.600 (b) 1.20 (c) 2.40 (d) 3.60 (e) none of those answers

8. If you increase the aperture diameter of a camera by a factor of 3, how is the intensity of the light striking the film affected? (a) It increases by factor of 3. (b) It decreases by a factor of 3. (c) It increases by a factor of 9. (d) It decreases by a factor of 9. (e) Increasing the aperture size doesn't affect the intensity.

9. A person spearfishing from a boat sees a stationary fish a few meters away in a direction about 30° below the horizontal. To spear the fish, and assuming the spear does not change direction when it enters the water, should the person (a) aim above where he sees the fish, (b) aim below the fish, or (c) aim precisely at the fish?

10. Model each of the following devices in use as consisting of a single converging lens. Rank the cases according to the ratio of the distance from the object to the lens to the focal length of the lens, from the largest ratio to the smallest. (a) a film-based movie projector showing a movie (b) a magnifying glass being used to examine a postage stamp (c) an astronomical refracting telescope being used to make a sharp image of stars on an electronic detector (d) a searchlight being used to produce a beam of parallel rays from a point source (e) a camera lens being used to photograph a soccer game

11. A converging lens made of crown glass has a focal length of 15.0 cm when used in air. If the lens is immersed in water, what is its focal length? (a) negative

(b) less than 15.0 cm (c) equal to 15.0 cm (d) greater than 15.0 cm (e) none of those answers

12. A converging lens of focal length 8 cm forms a sharp image of an object on a screen. What is the smallest possible distance between the object and the screen? (a) 0 (b) 4 cm (c) 8 cm (d) 16 cm (e) 32 cm

13. **(i)** When an image of an object is formed by a plane mirror, which of the following statements is *always* true? More than one statement may be correct. (a) The image is virtual. (b) The image is real. (c) The image is upright. (d) The image is inverted. (e) None of those statements is always true. **(ii)** When the image of an object is formed by a concave mirror, which of the preceding statements are *always* true? **(iii)** When the image of an object is formed by a convex mirror, which of the preceding statements are *always* true?

14. An object, represented by a gray arrow, is placed in front of a plane mirror. Which of the diagrams in Figure OQ36.14 correctly describes the image, represented by the pink arrow?

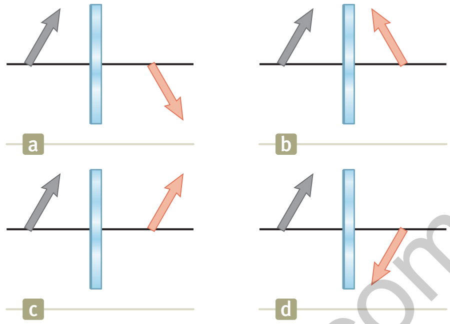


Figure OQ36.14

Conceptual Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. A converging lens of short focal length can take light diverging from a small source and refract it into a beam of parallel rays. A Fresnel lens as shown in Figure 36.27 is used in a lighthouse for this purpose. A concave mirror can take light diverging from a small source and reflect it into a beam of parallel rays. (a) Is it possible to make a Fresnel mirror? (b) Is this idea original, or has it already been done?

2. Explain this statement: "The focal point of a lens is the location of the image of a point object at infinity." (a) Discuss the notion of infinity in real terms as it applies to object distances. (b) Based on this statement, can you think of a simple method for determining the focal length of a converging lens?

3. Why do some emergency vehicles have the symbol ƎƆИA⅃UꓭMA written on the front?

4. Explain why a mirror cannot give rise to chromatic aberration.

5. (a) Can a converging lens be made to diverge light if it is placed into a liquid? (b) **What If?** What about a converging mirror?

6. Explain why a fish in a spherical goldfish bowl appears larger than it really is.

7. In Figure 36.26a, assume the gray object arrow is replaced by one that is much taller than the lens. (a) How many rays from the top of the object will strike the lens? (b) How many principal rays can be drawn in a ray diagram?

8. Lenses used in eyeglasses, whether converging or diverging, are always designed so that the middle of the lens curves away from the eye like the center lenses of Figures 36.25a and 36.25b. Why?

9. Suppose you want to use a converging lens to project the image of two trees onto a screen. As shown in Figure CQ36.9, one tree is a distance x from the lens and the other is at $2x$. You adjust the screen so that the near tree is in focus. If you now want the far tree to be in focus, do you move the screen toward or away from the lens?

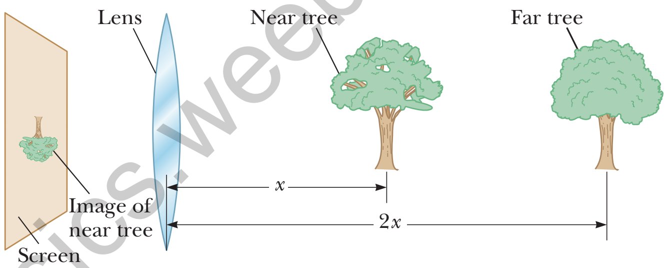


Figure CQ36.9

10. Consider a spherical concave mirror with the object located to the left of the mirror beyond the focal point. Using ray diagrams, show that the image moves to the left as the object approaches the focal point.

11. In Figures CQ36.11a and CQ36.11b, which glasses correct nearsightedness and which correct farsightedness?

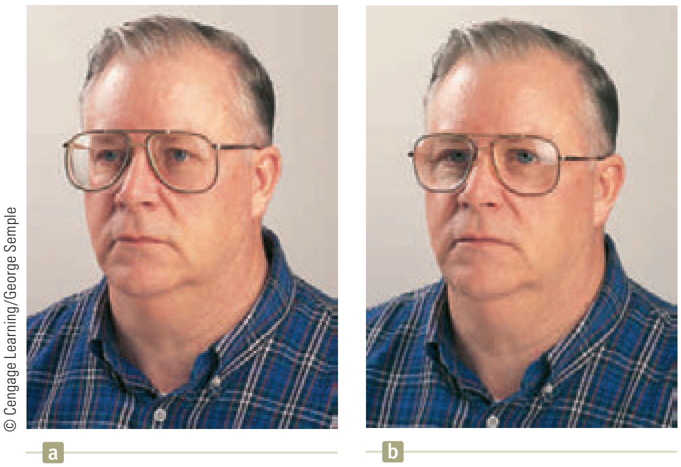


Figure CQ36.11 Conceptual Questions 11 and 12.

12. Bethany tries on either her hyperopic grandfather's or her myopic brother's glasses and complains, "Everything looks blurry." Why do the eyes of a person wearing glasses not look blurry? (See Fig. CQ36.11.)

13. In a Jules Verne novel, a piece of ice is shaped to form a magnifying lens, focusing sunlight to start a fire. Is that possible?

14. A solar furnace can be constructed by using a concave mirror to reflect and focus sunlight into a furnace enclosure. What factors in the design of the reflecting mirror would guarantee very high temperatures?

15. Figure CQ36.15 shows a lithograph by M. C. Escher titled *Hand with Reflection Sphere (Self-Portrait in Spherical Mirror)*. Escher said about the work: "The picture shows a spherical mirror, resting on a left hand. But as a print is the reverse of the original drawing on stone, it was my right hand that you see depicted. (Being left-handed, I needed my left hand to make the drawing.) Such a globe reflection collects almost one's whole surroundings in one disk-shaped image. The whole room, four walls, the floor, and the ceiling, everything, albeit distorted, is compressed into that one small circle. Your own head, or more exactly the point between your eyes, is the absolute center. No matter how you turn or twist yourself, you can't get out of that central point. You are immovably the focus, the unshakable core, of your world." Comment on the accuracy of Escher's description.

Figure CQ36.15

16. If a cylinder of solid glass or clear plastic is placed above the words LEAD OXIDE and viewed from the side as shown in Figure CQ36.16, the word LEAD appears inverted, but the word OXIDE does not. Explain.

Figure CQ36.16

17. Do the equations $1/p + 1/q = 1/f$ and $M = -q/p$ apply to the image formed by a flat mirror? Explain your answer.

Problems

ENHANCED WebAssign The problems found in this chapter may be assigned online in Enhanced WebAssign

1. straightforward; **2.** intermediate; **3.** challenging

1. full solution available in the *Student Solutions Manual/Study Guide*

AMT Analysis Model tutorial available in Enhanced WebAssign

GP Guided Problem

M Master It tutorial available in Enhanced WebAssign

W Watch It video solution available in Enhanced WebAssign

Section 36.1 Images Formed by Flat Mirrors

1. AMT M W Determine the minimum height of a vertical flat mirror in which a person 178 cm tall can see his or her full image. *Suggestion:* Drawing a ray diagram would be helpful.

2. AMT In a choir practice room, two parallel walls are 5.30 m apart. The singers stand against the north wall. The organist faces the south wall, sitting 0.800 m away from it. To enable her to see the choir, a flat mirror 0.600 m wide is mounted on the south wall, straight in front of her. What width of the north wall can the organist see? *Suggestion:* Draw a top-view diagram to justify your answer.

3. (a) Does your bathroom mirror show you older or younger than you actually are? (b) Compute an order-of-magnitude estimate for the age difference based on data you specify.

4. A person walks into a room that has two flat mirrors on opposite walls. The mirrors produce multiple images of the person. Consider *only* the images formed in the mirror on the left. When the person is 2.00 m from the mirror on the left wall and 4.00 m from the mirror on the right wall, find the distance from the person to the first three images seen in the mirror on the left wall.

5. A periscope (Fig. P36.5) is useful for viewing objects that cannot be seen directly. It can be used in submarines and when watching golf matches or parades from behind a crowd of people. Suppose the object is a distance p_1 from the upper mirror and the centers

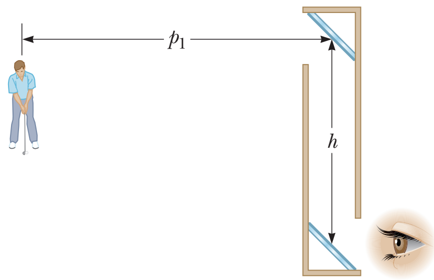


Figure P36.5

of the two flat mirrors are separated by a distance h. (a) What is the distance of the final image from the lower mirror? (b) Is the final image real or virtual? (c) Is it upright or inverted? (d) What is its magnification? (e) Does it appear to be left–right reversed?

6. Two flat mirrors have their reflecting surfaces facing each other, with the edge of one mirror in contact with an edge of the other, so that the angle between the mirrors is α. When an object is placed between the mirrors, a number of images are formed. In general, if the angle α is such that $n\alpha = 360°$, where n is an integer, the number of images formed is $n - 1$. Graphically, find all the image positions for the case $n = 6$ when a point object is between the mirrors (but not on the angle bisector).

7. Two plane mirrors stand facing each other, 3.00 m apart, and a woman stands between them. The woman looks at one of the mirrors from a distance of 1.00 m and holds her left arm out to the side of her body with the palm of her hand facing the closer mirror. (a) What is the apparent position of the closest image of her left hand, measured perpendicularly from the surface of the mirror in front of her? (b) Does it show the palm of her hand or the back of her hand? (c) What is the position of the next closest image? (d) Does it show the palm of her hand or the back of her hand? (e) What is the position of the third closest image? (f) Does it show the palm of her hand or the back of her hand? (g) Which of the images are real and which are virtual?

Section 36.2 Images Formed by Spherical Mirrors

8. An object is placed 50.0 cm from a concave spherical mirror with focal length of magnitude 20.0 cm. (a) Find the location of the image. (b) What is the magnification of the image? (c) Is the image real or virtual? (d) Is the image upright or inverted?

9. M A concave spherical mirror has a radius of curvature of magnitude 20.0 cm. (a) Find the location of the image for object distances of **(i)** 40.0 cm, **(ii)** 20.0 cm, and **(iii)** 10.0 cm. For each case, state whether the image is (b) real or virtual and (c) upright or inverted. (d) Find the magnification in each case.

10. An object is placed 20.0 cm from a concave spherical mirror having a focal length of magnitude 40.0 cm. (a) Use graph paper to construct an accurate ray diagram for this situation. (b) From your ray diagram, determine the location of the image. (c) What is the magnification of the image? (d) Check your answers to parts (b) and (c) using the mirror equation.

11. M A convex spherical mirror has a radius of curvature of magnitude 40.0 cm. Determine the position of the virtual image and the magnification for object distances of (a) 30.0 cm and (b) 60.0 cm. (c) Are the images in parts (a) and (b) upright or inverted?

12. M W At an intersection of hospital hallways, a convex spherical mirror is mounted high on a wall to help people avoid collisions. The magnitude of the mirror's radius of curvature is 0.550 m. (a) Locate the image of a patient 10.0 m from the mirror. (b) Indicate whether the image is upright or inverted. (c) Determine the magnification of the image.

13. An object of height 2.00 cm is placed 30.0 cm from a convex spherical mirror of focal length of magnitude 10.0 cm. (a) Find the location of the image. (b) Indicate whether the image is upright or inverted. (c) Determine the height of the image.

14. A dentist uses a spherical mirror to examine a tooth. The tooth is 1.00 cm in front of the mirror, and the image is formed 10.0 cm behind the mirror. Determine (a) the mirror's radius of curvature and (b) the magnification of the image.

15. A large hall in a museum has a niche in one wall. On the floor plan, the niche appears as a semicircular indentation of radius 2.50 m. A tourist stands on the centerline of the niche, 2.00 m out from its deepest point, and whispers "Hello." Where is the sound concentrated after reflection from the niche?

16. *Why is the following situation impossible?* At a blind corner in an outdoor shopping mall, a convex mirror is mounted so pedestrians can see around the corner before arriving there and bumping into someone traveling in the perpendicular direction. The installers of the mirror failed to take into account the position of the Sun, and the mirror focuses the Sun's rays on a nearby bush and sets it on fire.

17. To fit a contact lens to a patient's eye, a *keratometer* can be used to measure the curvature of the eye's front surface, the cornea. This instrument places an illuminated object of known size at a known distance p from the cornea. The cornea reflects some light from the object, forming an image of the object. The magnification M of the image is measured by using a small viewing telescope that allows comparison of the image formed by the cornea with a second calibrated image projected into the field of view by a prism arrangement. Determine the radius of curvature of the cornea for the case $p = 30.0$ cm and $M = 0.013\,0$.

18. A certain Christmas tree ornament is a silver sphere having a diameter of 8.50 cm. (a) If the size of an image created by reflection in the ornament is three-fourths the reflected object's actual size, determine the object's location. (b) Use a principal-ray diagram to determine whether the image is upright or inverted.

19. W (a) A concave spherical mirror forms an inverted image 4.00 times larger than the object. Assuming the distance between object and image is 0.600 m, find the focal length of the mirror. (b) **What If?** Suppose the mirror is convex. The distance between the image and the object is the same as in part (a), but the image is 0.500 the size of the object. Determine the focal length of the mirror.

20. (a) A concave spherical mirror forms an inverted image different in size from the object by a factor $a > 1$. The distance between object and image is d. Find the focal length of the mirror. (b) **What If?** Suppose the mirror is convex, an upright image is formed, and $a < 1$. Determine the focal length of the mirror.

21. W An object 10.0 cm tall is placed at the zero mark of a meterstick. A spherical mirror located at some point on the meterstick creates an image of the object that is upright, 4.00 cm tall, and located at the 42.0-cm mark of the meterstick. (a) Is the mirror convex or concave? (b) Where is the mirror? (c) What is the mirror's focal length?

22. A concave spherical mirror has a radius of curvature of magnitude 24.0 cm. (a) Determine the object position for which the resulting image is upright and larger than the object by a factor of 3.00. (b) Draw a ray diagram to determine the position of the image. (c) Is the image real or virtual?

23. W A dedicated sports car enthusiast polishes the inside and outside surfaces of a hubcap that is a thin section of a sphere. When she looks into one side of the hubcap, she sees an image of her face 30.0 cm in back of the hubcap. She then flips the hubcap over and sees another image of her face 10.0 cm in back of the hubcap. (a) How far is her face from the hubcap? (b) What is the radius of curvature of the hubcap?

24. A convex spherical mirror has a focal length of magnitude 8.00 cm. (a) What is the location of an object for which the magnitude of the image distance is one-third the magnitude of the object distance? (b) Find the magnification of the image and (c) state whether it is upright or inverted.

25. M A spherical mirror is to be used to form an image 5.00 times the size of an object on a screen located 5.00 m from the object. (a) Is the mirror required concave or convex? (b) What is the required radius of curvature of the mirror? (c) Where should the mirror be positioned relative to the object?

26. AMT **Review.** A ball is dropped at $t = 0$ from rest 3.00 m directly above the vertex of a concave spherical mirror that has a radius of curvature of magnitude 1.00 m and lies in a horizontal plane. (a) Describe the motion of the ball's image in the mirror. (b) At what instant or instants do the ball and its image coincide?

27. You unconsciously estimate the distance to an object from the angle it subtends in your field of view. This angle θ in radians is related to the linear height of the object h and to the distance d by $\theta = h/d$. Assume you are driving a car and another car, 1.50 m high, is 24.0 m behind you. (a) Suppose your car has a flat passenger-side rearview mirror, 1.55 m from your eyes. How far from your eyes is the image of the car following you? (b) What angle does the image subtend in your field of view? (c) **What If?** Now suppose your car has a convex rearview mirror with a radius of curvature of magnitude 2.00 m (as suggested in Fig. 36.15). How far from your eyes is the image of the car behind you? (d) What angle does the image subtend at your eyes? (e) Based on its angular size, how far away does the following car appear to be?

28. M A man standing 1.52 m in front of a shaving mirror produces an inverted image 18.0 cm in front of it. How close to the mirror should he stand if he wants to form an upright image of his chin that is twice the chin's actual size?

Section 36.3 Images Formed by Refraction

29. One end of a long glass rod ($n = 1.50$) is formed into a convex surface with a radius of curvature of magnitude 6.00 cm. An object is located in air along the axis of the rod. Find the image positions corresponding to object distances of (a) 20.0 cm, (b) 10.0 cm, and (c) 3.00 cm from the convex end of the rod.

30. A cubical block of ice 50.0 cm on a side is placed over a speck of dust on a level floor. Find the location of the image of the speck as viewed from above. The index of refraction of ice is 1.309.

31. The top of a swimming pool is at ground level. If the pool is 2.00 m deep, how far below ground level does the bottom of the pool appear to be located when (a) the pool is completely filled with water? (b) When it is filled halfway with water?

32. The magnification of the image formed by a refracting surface is given by

$$M = -\frac{n_1 q}{n_2 p}$$

where n_1, n_2, p, and q are defined as they are for Figure 36.17 and Equation 36.8. A paperweight is made of a solid glass hemisphere with index of refraction 1.50. The radius of the circular cross section is 4.00 cm. The hemisphere is placed on its flat surface, with the center directly over a 2.50-mm-long line drawn on a sheet of paper. What is the length of this line as seen by someone looking vertically down on the hemisphere?

33. A flint glass plate rests on the bottom of an aquarium tank. The plate is 8.00 cm thick (vertical dimension) and is covered with a layer of water 12.0 cm deep. Calculate the apparent thickness of the plate as viewed from straight above the water.

34. Figure P36.34 shows a curved surface separating a material with index of refraction n_1 from a material with index n_2. The surface forms an image I of object O. The ray shown in red passes through the surface along a radial line. Its angles of incidence and refraction are both zero, so its direction does not change at the surface. For the ray shown in blue, the direction changes according to Snell's law, $n_1 \sin\theta_1 = n_2 \sin\theta_2$. For paraxial rays, we assume θ_1 and θ_2 are small, so we may write $n_1 \tan\theta_1 = n_2 \tan\theta_2$. The magnification is defined as $M = h'/h$. Prove that the magnification is given by $M = -n_1 q/n_2 p$.

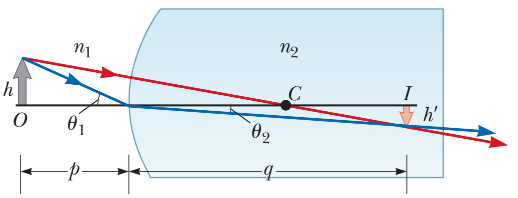


Figure P36.34

35. M A glass sphere ($n = 1.50$) with a radius of 15.0 cm has a tiny air bubble 5.00 cm above its center. The sphere is viewed looking down along the extended radius containing the bubble. What is the apparent depth of the bubble below the surface of the sphere?

36. As shown in Figure P36.36, Ben and Jacob check out an aquarium that has a curved front made of plastic with uniform thickness and a radius of curvature of magnitude $R = 2.25$ m. (a) Locate the images of fish that are located (i) 5.00 cm and (ii) 25.0 cm from the front wall of the aquarium. (b) Find the magnification of images (i) and (ii) from the previous part. (See Problem 32 to find an expression for the magnification of an image formed by a refracting surface.) (c) Explain why you don't need to know the refractive index of the plastic to solve this problem. (d) If this aquarium were very long from front to back, could the image of a fish ever be farther from the front surface than the fish itself is? (e) If not, explain why not. If so, give an example and find the magnification.

Chris Candela

Figure P36.36

37. W A goldfish is swimming at 2.00 cm/s toward the front wall of a rectangular aquarium. What is the apparent speed of the fish measured by an observer looking in from outside the front wall of the tank?

Section 36.4 Images Formed by Thin Lenses

38. A thin lens has a focal length of 25.0 cm. Locate and describe the image when the object is placed (a) 26.0 cm and (b) 24.0 cm in front of the lens.

39. An object located 32.0 cm in front of a lens forms an image on a screen 8.00 cm behind the lens. (a) Find the focal length of the lens. (b) Determine the magnification. (c) Is the lens converging or diverging?

40. M An object is located 20.0 cm to the left of a diverging lens having a focal length $f = -32.0$ cm. Determine (a) the location and (b) the magnification of the image. (c) Construct a ray diagram for this arrangement.

41. W The projection lens in a certain slide projector is a single thin lens. A slide 24.0 mm high is to be projected so that its image fills a screen 1.80 m high. The slide-to-screen distance is 3.00 m. (a) Determine the focal length of the projection lens. (b) How far from the slide should the lens of the projector be placed so as to form the image on the screen?

42. An object's distance from a converging lens is 5.00 times the focal length. (a) Determine the location of the image. Express the answer as a fraction of the focal length. (b) Find the magnification of the image and indicate whether it is (c) upright or inverted and (d) real or virtual.

43. W A contact lens is made of plastic with an index of refraction of 1.50. The lens has an outer radius of curvature of +2.00 cm and an inner radius of curvature of +2.50 cm. What is the focal length of the lens?

44. A converging lens has a focal length of 10.0 cm. Construct accurate ray diagrams for object distances of **(i)** 20.0 cm and **(ii)** 5.00 cm. (a) From your ray diagrams, determine the location of each image. (b) Is the image real or virtual? (c) Is the image upright or inverted? (d) What is the magnification of the image? (e) Compare your results with the values found algebraically. (f) Comment on difficulties in constructing the graph that could lead to differences between the graphical and algebraic answers.

45. A converging lens has a focal length of 10.0 cm. Locate the object if a real image is located at a distance from the lens of (a) 20.0 cm and (b) 50.0 cm. **What If?** Redo the calculations if the images are virtual and located at a distance from the lens of (c) 20.0 cm and (d) 50.0 cm.

46. A diverging lens has a focal length of magnitude 20.0 cm. (a) Locate the image for object distances of **(i)** 40.0 cm, **(ii)** 20.0 cm, and **(iii)** 10.0 cm. For each case, state whether the image is (b) real or virtual and (c) upright or inverted.(d) For each case, find the magnification.

47. The nickel's image in Figure P36.47 has twice the diameter of the nickel and is 2.84 cm from the lens. Determine the focal length of the lens.

Figure P36.47

48. Suppose an object has thickness dp so that it extends from object distance p to $p + dp$. (a) Prove that the thickness dq of its image is given by $(-q^2/p^2)dp$. (b) The longitudinal magnification of the object is $M_{\text{long}} = dq/dp$. How is the longitudinal magnification related to the lateral magnification M?

49. The left face of a biconvex lens has a radius of curvature of magnitude 12.0 cm, and the right face has a radius of curvature of magnitude 18.0 cm. The index of refraction of the glass is 1.44. (a) Calculate the focal length of the lens for light incident from the left. (b) **What If?** After the lens is turned around to interchange the radii of curvature of the two faces, calculate the focal length of the lens for light incident from the left.

50. In Figure P36.50, a thin converging lens of focal length 14.0 cm forms an image of the square $abcd$, which is $h_c = h_b = 10.0$ cm high and lies between distances of $p_d = 20.0$ cm and $p_a = 30.0$ cm from the lens. Let a', b', c', and d' represent the respective corners of the image. Let q_a represent the image distance for points a' and b', q_d represent the image distance for points c' and d',

h'_b represent the distance from point b' to the axis, and h'_c represent the height of c'. (a) Find q_a, q_d, h'_b, and h'_c. (b) Make a sketch of the image. (c) The area of the object is 100 cm^2. By carrying out the following steps, you will evaluate the area of the image. Let q represent the image distance of any point between a' and d', for which the object distance is p. Let h' represent the distance from the axis to the point at the edge of the image between b' and c' at image distance q. Demonstrate that

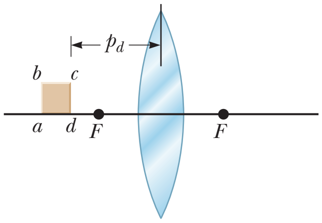


Figure P36.50

$$|h'| = 10.0q\left(\frac{1}{14.0} - \frac{1}{q}\right)$$

where h' and q are in centimeters. (d) Explain why the geometric area of the image is given by

$$\int_{q_a}^{q_d} |h'|\,dq$$

(e) Carry out the integration to find the area of the image.

51. An antelope is at a distance of 20.0 m from a converging lens of focal length 30.0 cm. The lens forms an image of the animal. (a) If the antelope runs away from the lens at a speed of 5.00 m/s, how fast does the image move? (b) Does the image move toward or away from the lens?

52. *Why is the following situation impossible?* An illuminated object is placed a distance d = 2.00 m from a screen. By placing a converging lens of focal length f = 60.0 cm at two locations between the object and the screen, a sharp, real image of the object can be formed on the screen. In one location of the lens, the image is larger than the object, and in the other, the image is smaller.

53. A 1.00-cm-high object is placed 4.00 cm to the left of a converging lens of focal length 8.00 cm. A diverging lens of focal length −16.00 cm is 6.00 cm to the right of the converging lens. Find the position and height of the final image. Is the image inverted or upright? Real or virtual?

Section 36.5 Lens Aberrations

54. The magnitudes of the radii of curvature are 32.5 cm and 42.5 cm for the two faces of a biconcave lens. The glass has index of refraction 1.53 for violet light and 1.51 for red light. For a very distant object, locate (a) the image formed by violet light and (b) the image formed by red light.

55. Two rays traveling parallel to the principal axis strike a large plano-convex lens having a refractive index of 1.60 (Fig. P36.55). If the convex face is spherical, a ray near the edge does not pass through the focal point (spherical aberration occurs). Assume this face has a radius of curvature of R = 20.0 cm and the two rays are at distances h_1 = 0.500 cm and h_2 = 12.0 cm from the principal axis. Find the difference Δx in the positions where each crosses the principal axis.

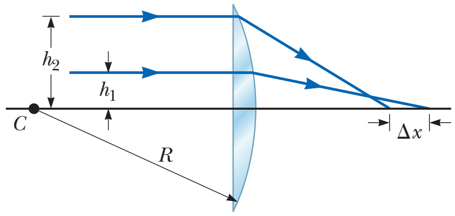


Figure P36.55

Section 36.6 The Camera

56. A camera is being used with a correct exposure at f/4 and a shutter speed of $\frac{1}{15}$ s. In addition to the f-numbers listed in Section 36.6, this camera has f-numbers f/1, f/1.4, and f/2. To photograph a rapidly moving subject, the shutter speed is changed to $\frac{1}{125}$ s. Find the new f-number setting needed on this camera to maintain satisfactory exposure.

57. Figure 36.33 diagrams a cross section of a camera. It has a single lens of focal length 65.0 mm, which is to form an image on the CCD at the back of the camera. Suppose the position of the lens has been adjusted to focus the image of a distant object. How far and in what direction must the lens be moved to form a sharp image of an object that is 2.00 m away?

Section 36.7 The Eye

58. A nearsighted person cannot see objects clearly beyond 25.0 cm (her far point). If she has no astigmatism and contact lenses are prescribed for her, what (a) power and (b) type of lens are required to correct her vision?

59. The near point of a person's eye is 60.0 cm. To see objects clearly at a distance of 25.0 cm, what should be the (a) focal length and (b) power of the appropriate corrective lens? (Neglect the distance from the lens to the eye.)

60. A person sees clearly wearing eyeglasses that have a power of −4.00 diopters when the lenses are 2.00 cm in front of the eyes. (a) What is the focal length of the lens? (b) Is the person nearsighted or farsighted? (c) If the person wants to switch to contact lenses placed directly on the eyes, what lens power should be prescribed?

61. The accommodation limits for a nearsighted person's eyes are 18.0 cm and 80.0 cm. When he wears his glasses, he can see faraway objects clearly. At what minimum distance is he able to see objects clearly?

62. A certain child's near point is 10.0 cm; her far point (with eyes relaxed) is 125 cm. Each eye lens is 2.00 cm from the retina. (a) Between what limits, measured in diopters, does the power of this lens–cornea combination vary? (b) Calculate the power of the eyeglass lens the child should use for relaxed distance vision. Is the lens converging or diverging?

63. A person is to be fitted with bifocals. She can see clearly when the object is between 30 cm and 1.5 m from the eye. (a) The upper portions of the bifocals (Fig. P36.63) should be designed to enable her to see distant objects clearly. What power should they have? (b) The lower portions of the bifocals should enable her to see objects located 25 cm in front of the eye. What power should they have?

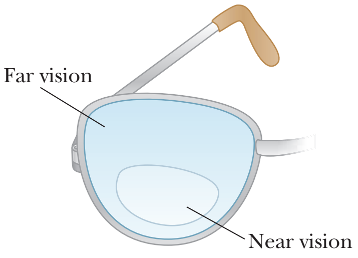


Figure P36.63

64. A simple model of the human eye ignores its lens entirely. Most of what the eye does to light happens at the outer surface of the transparent cornea. Assume that this surface has a radius of curvature of 6.00 mm and that the eyeball contains just one fluid with a refractive index of 1.40. Prove that a very distant object will be imaged on the retina, 21.0 mm behind the cornea. Describe the image.

65. A patient has a near point of 45.0 cm and far point of 85.0 cm. (a) Can a single pair of glasses correct the patient's vision? Explain the patient's options. (b) Calculate the power lens needed to correct the near point so that the patient can see objects 25.0 cm away. Neglect the eye–lens distance. (c) Calculate the power lens needed to correct the patient's far point, again neglecting the eye–lens distance.

Section 36.8 The Simple Magnifier

66. A lens that has a focal length of 5.00 cm is used as a magnifying glass. (a) To obtain maximum magnification and an image that can be seen clearly by a normal eye, where should the object be placed? (b) What is the magnification?

Section 36.9 The Compound Microscope

67. The distance between the eyepiece and the objective lens in a certain compound microscope is 23.0 cm. The focal length of the eyepiece is 2.50 cm and that of the objective is 0.400 cm. What is the overall magnification of the microscope?

Section 36.10 The Telescope

68. M The refracting telescope at the Yerkes Observatory has a 1.00-m diameter objective lens of focal length 20.0 m. Assume it is used with an eyepiece of focal length 2.50 cm. (a) Determine the magnification of Mars as seen through this telescope. (b) Are the Martian polar caps right side up or upside down?

69. A certain telescope has an objective mirror with an aperture diameter of 200 mm and a focal length of 2 000 mm. It captures the image of a nebula on photographic film at its prime focus with an exposure time of 1.50 min. To produce the same light energy per unit area on the film, what is the required exposure time to photograph the same nebula with a smaller telescope that has an objective with a 60.0-mm diameter and a 900-mm focal length?

70. Astronomers often take photographs with the objective lens or mirror of a telescope alone, without an eyepiece. (a) Show that the image size h' for such a telescope is given by $h' = fh/(f - p)$, where f is the objective focal length, h is the object size, and p is the object distance. (b) **What If?** Simplify the expression in part (a) for the case in which the object distance is much greater than objective focal length. (c) The "wingspan" of the International Space Station is 108.6 m, the overall width of its solar panel configuration. When the station is orbiting at an altitude of 407 km, find the width of the image formed by a telescope objective of focal length 4.00 m.

Additional Problems

71. The lens-makers' equation applies to a lens immersed in a liquid if n in the equation is replaced by n_2/n_1. Here n_2 refers to the index of refraction of the lens material and n_1 is that of the medium surrounding the lens. (a) A certain lens has focal length 79.0 cm in air and index of refraction 1.55. Find its focal length in water. (b) A certain mirror has focal length 79.0 cm in air. Find its focal length in water.

72. A real object is located at the zero end of a meterstick. A large concave spherical mirror at the 100-cm end of the meterstick forms an image of the object at the 70.0-cm position. A small convex spherical mirror placed at the 20.0-cm position forms a final image at the 10.0-cm point. What is the radius of curvature of the convex mirror?

73. The distance between an object and its upright image is 20.0 cm. If the magnification is 0.500, what is the focal length of the lens being used to form the image?

74. The distance between an object and its upright image is d. If the magnification is M, what is the focal length of the lens being used to form the image?

75. M A person decides to use an old pair of eyeglasses to make some optical instruments. He knows that the near point in his left eye is 50.0 cm and the near point in his right eye is 100 cm. (a) What is the maximum angular magnification he can produce in a telescope? (b) If he places the lenses 10.0 cm apart, what is the maximum overall magnification he can produce in a microscope? *Hint:* Go back to basics and use the thin lens equation to solve part (b).

76. You are designing an endoscope for use inside an air-filled body cavity. A lens at the end of the endoscope will form an image covering the end of a bundle of optical fibers. This image will then be carried by the optical fibers to an eyepiece lens at the outside end of the fiberscope. The radius of the bundle is 1.00 mm. The scene within the body that is to appear within the image fills a circle of radius 6.00 cm. The lens will be located 5.00 cm from the tissues you wish to observe. (a) How far should the lens be located from the end of an optical fiber bundle? (b) What is the focal length of the lens required?

77. The lens and mirror in Figure P36.77 are separated by $d = 1.00$ m and have focal lengths of +80.0 cm and

−50.0 cm, respectively. An object is placed p = 1.00 m to the left of the lens as shown. (a) Locate the final image, formed by light that has gone through the lens twice. (b) Determine the overall magnification of the image and (c) state whether the image is upright or inverted.

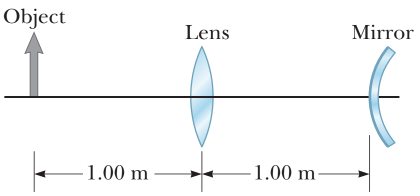


Figure P36.77

78. Two converging lenses having focal lengths of f_1 = 10.0 cm and f_2 = 20.0 cm are placed a distance d = 50.0 cm apart as shown in Figure P36.78. The image due to light passing through both lenses is to be located between the lenses at the position x = 31.0 cm indicated. (a) At what value of p should the object be positioned to the left of the first lens? (b) What is the magnification of the final image? (c) Is the final image upright or inverted? (d) Is the final image real or virtual?

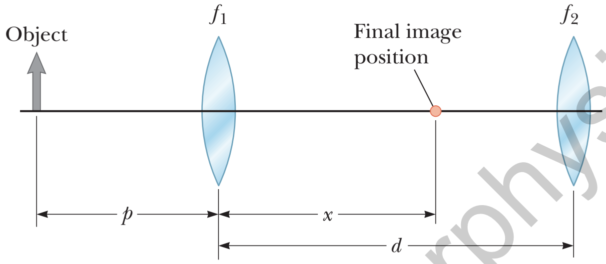


Figure P36.78

79. Figure P36.79 shows a piece of glass with index of refraction n = 1.50 surrounded by air. The ends are hemispheres with radii R_1 = 2.00 cm and R_2 = 4.00 cm, and the centers of the hemispherical ends are separated by a distance of d = 8.00 cm. A point object is in air, a distance p = 1.00 cm from the left end of the glass. (a) Locate the image of the object due to refraction at the two spherical surfaces. (b) Is the final image real or virtual?

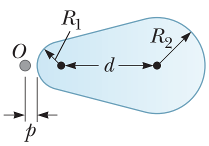


Figure P36.79

80. An object is originally at the x_i = 0 cm position of a meterstick located on the x axis. A converging lens of focal length 26.0 cm is fixed at the position 32.0 cm. Then we gradually slide the object to the position x_f = 12.0 cm. (a) Find the location x' of the object's image as a function of the object position x. (b) Describe the pattern of the image's motion with reference to a graph or a table of values. (c) As the object moves 12.0 cm to the right, how far does the image move? (d) In what direction or directions?

81. The object in Figure P36.81 is midway between the lens and the mirror, which are separated by a distance d = 25.0 cm. The magnitude of the mirror's radius of curvature is 20.0 cm, and the lens has a focal length of −16.7 cm. (a) Considering only the light that leaves the object and travels first toward the mirror, locate the final image formed by this system. (b) Is this image real or virtual? (c) Is it upright or inverted? (d) What is the overall magnification?

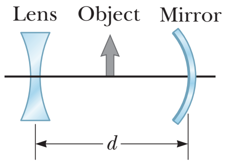


Figure P36.81

82. In many applications, it is necessary to expand or decrease the diameter of a beam of parallel rays of light, which can be accomplished by using a converging lens and a diverging lens in combination. Suppose you have a converging lens of focal length 21.0 cm and a diverging lens of focal length −12.0 cm. (a) How can you arrange these lenses to increase the diameter of a beam of parallel rays? (b) By what factor will the diameter increase?

83. **Review.** A spherical lightbulb of diameter 3.20 cm radiates light equally in all directions, with power 4.50 W. (a) Find the light intensity at the surface of the lightbulb. (b) Find the light intensity 7.20 m away from the center of the lightbulb. (c) At this 7.20-m distance, a lens is set up with its axis pointing toward the lightbulb. The lens has a circular face with a diameter of 15.0 cm and has a focal length of 35.0 cm. Find the diameter of the lightbulb's image. (d) Find the light intensity at the image.

84. A parallel beam of light enters a glass hemisphere perpendicular to the flat face as shown in Figure P36.84. The magnitude of the radius of the hemisphere is R = 6.00 cm, and its index of refraction is n = 1.560. Assuming paraxial rays, determine the point at which the beam is focused.

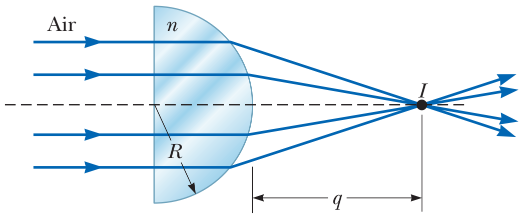


Figure P36.84

85. Two lenses made of kinds of glass having different indices of refraction n_1 and n_2 are cemented together to form an *optical doublet.* Optical doublets are often used to correct chromatic aberrations in optical devices. The first lens of a certain doublet has index of refraction n_1, one flat side, and one concave side with a radius of curvature of magnitude R. The second lens has index of refraction n_2 and two convex sides with radii of curvature also of magnitude R. Show that the doublet can be modeled as a single thin lens with a focal length described by

$$\frac{1}{f} = \frac{2n_2 - n_1 - 1}{R}$$

86. *Why is the following situation impossible?* Consider the lens–mirror combination shown in Figure P36.86 on page 1132. The lens has a focal length of f_L = 0.200 m, and the mirror has a focal length of $f_M = 0.500$ m. The lens and mirror are placed a distance $d = 1.30$ m apart, and an object is placed at $p = 0.300$ m from the lens. By moving a screen to various positions to the left of the lens, a student finds two different positions of the screen that produce a sharp image of the object. One of these positions corresponds to light leaving the object and traveling to the left through the lens. The other position corresponds to light traveling to the right from the object, reflecting from the mirror and then passing through the lens.

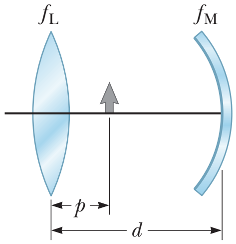


Figure P36.86 Problems 86 and 97.

87. M An object is placed 12.0 cm to the left of a diverging lens of focal length −6.00 cm. A converging lens of focal length 12.0 cm is placed a distance d to the right of the diverging lens. Find the distance d so that the final image is infinitely far away to the right.

88. An object is placed a distance p to the left of a diverging lens of focal length f_1. A converging lens of focal length f_2 is placed a distance d to the right of the diverging lens. Find the distance d so that the final image is infinitely far away to the right.

89. An observer to the right of the mirror–lens combination shown in Figure P36.89 (not to scale) sees two real images that are the same size and in the same location. One image is upright, and the other is inverted. Both images are 1.50 times larger than the object. The lens has a focal length of 10.0 cm. The lens and mirror are separated by 40.0 cm. Determine the focal length of the mirror.

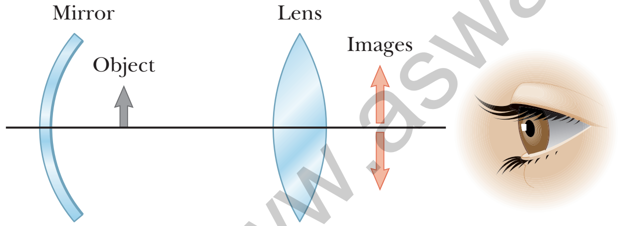


Figure P36.89

90. GP M In a darkened room, a burning candle is placed 1.50 m from a white wall. A lens is placed between the candle and the wall at a location that causes a larger, inverted image to form on the wall. When the lens is in this position, the object distance is p_1. When the lens is moved 90.0 cm toward the wall, another image of the candle is formed on the wall. From this information, we wish to find p_1 and the focal length of the lens. (a) From the lens equation for the first position of the lens, write an equation relating the focal length f of the lens to the object distance p_1, with no other variables in the equation. (b) From the lens equation for the second position of the lens, write another equation relating the focal length f of the lens to the object distance p_1. (c) Solve the equations in parts (a) and (b) simultaneously to find p_1. (d) Use the value in part (c) to find the focal length f of the lens.

91. The disk of the Sun subtends an angle of 0.533° at the Earth. What are (a) the position and (b) the diameter of the solar image formed by a concave spherical mirror with a radius of curvature of magnitude 3.00 m?

92. An object 2.00 cm high is placed 40.0 cm to the left of a converging lens having a focal length of 30.0 cm. A diverging lens with a focal length of −20.0 cm is placed 110 cm to the right of the converging lens. Determine (a) the position and (b) the magnification of the final image. (c) Is the image upright or inverted? (d) **What If?** Repeat parts (a) through (c) for the case in which the second lens is a converging lens having a focal length of 20.0 cm.

Challenge Problems

93. Assume the intensity of sunlight is 1.00 kW/m² at a particular location. A highly reflecting concave mirror is to be pointed toward the Sun to produce a power of at least 350 W at the image point. (a) Assuming the disk of the Sun subtends an angle of 0.533° at the Earth, find the required radius R_a of the circular face area of the mirror. (b) Now suppose the light intensity is to be at least 120 kW/m² at the image. Find the required relationship between R_a and the radius of curvature R of the mirror.

94. A *zoom lens* system is a combination of lenses that produces a variable magnification of a fixed object as it maintains a fixed image position. The magnification is varied by moving one or more lenses along the axis. Multiple lenses are used in practice, but the effect of zooming in on an object can be demonstrated with a simple two-lens system. An object, two converging lenses, and a screen are mounted on an optical bench. Lens 1, which is to the right of the object, has a focal length of $f_1 = 5.00$ cm, and lens 2, which is to the right of the first lens, has a focal length of $f_2 = 10.0$ cm. The screen is to the right of lens 2. Initially, an object is situated at a distance of 7.50 cm to the left of lens 1, and the image formed on the screen has a magnification of +1.00. (a) Find the distance between the object and the screen. (b) Both lenses are now moved along their common axis while the object and the screen maintain fixed positions until the image formed on the screen has a magnification of +3.00. Find the displacement of each lens from its initial position in part (a). (c) Can the lenses be displaced in more than one way?

95. Figure P36.95 shows a thin converging lens for which the radii of curvature of its surfaces have magnitudes of 9.00 cm and 11.0 cm. The lens is in front of a concave spherical mirror with the radius of curvature $R = 8.00$ cm. Assume the focal points F_1 and F_2 of the lens are 5.00 cm from the center of the lens. (a) Determine the index of refraction of the lens material. The lens and mirror are 20.0 cm apart, and an object is placed

8.00 cm to the left of the lens. Determine (b) the posi tion of the final image and (c) its magnification as seen by the eye in the figure. (d) Is the final image inverted or upright? Explain.

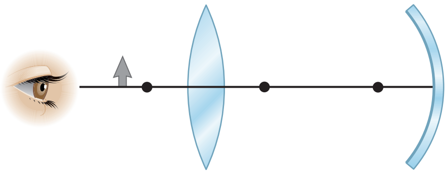


Figure P36.95

96. A floating strawberry illusion is achieved with two par abolic mirrors, each having a focal length 7.50 cm, facing each other as shown in Figure P36.96. If a strawberry is placed on the lower mirror, an image of the strawberry is formed at the small opening at the center of the top mirror, 7.50 cm above the lowest point of the bottom mirror. The position of the eye in Figure P36.96a corresponds to the view of the apparatus in Figure P36.96b. Consider the light path marked Notice that this light path is blocked by the upper mir ror so that the strawberry itself is not directly observ able. The light path marked corresponds to the eye viewing the image of the strawberry that is formed at the opening at the top of the apparatus. (a) Show that the final image is formed at that location and describe its characteristics. (b) A very startling effect is to shine a flashlight beam on this image. Even at a glancing angle, the incoming light beam is seemingly reflected from the image! Explain.

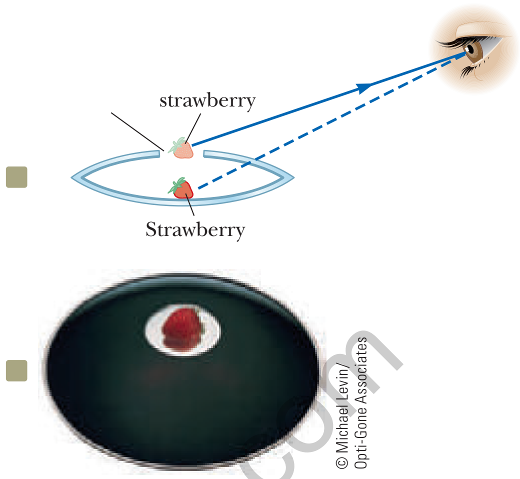


Figure P36.96

97. Consider the lens–mirror arrangement shown in Fig ure P36.86. There are two final image positions to the left of the lens of focal length . One image position is due to light traveling from the object to the left and passing through the lens. The other image position is due to light traveling to the right from the object, reflecting from the mirror of focal length and then passing through the lens. For a given object position between the lens and the mirror and measured with respect to the lens, there are two separation distances between the lens and mirror that will cause the two images described above to be at the same location. Find both positions.

CHAPTER

37

Wave Optics

The colors in many of a hummingbird's feathers are not due to pigment. The *iridescence* that makes the brilliant colors that often appear on the bird's throat and belly is due to an interference effect caused by structures in the feathers. The colors will vary with the viewing angle. *(Dec Hogan/Shutterstock.com)*

In Chapter 36, we studied light rays passing through a lens or reflecting from a mirror to describe the formation of images. This discussion completed our study of *ray optics.* In this chapter and in Chapter 38, we are concerned with *wave optics,* sometimes called *physical optics,* the study of interference, diffraction, and polarization of light. These phenomena cannot be adequately explained with the ray optics used in Chapters 35 and 36. We now learn how treating light as waves rather than as rays leads to a satisfying description of such phenomena.

37.1 Young's Double-Slit Experiment

In Chapter 18, we studied the waves in interference model and found that the superposition of two mechanical waves can be constructive or destructive. In constructive interference, the amplitude of the resultant wave is greater than that of either individual wave, whereas in destructive interference, the resultant amplitude is less than that of the larger wave. Light waves also interfere with one another. Fundamentally, all interference associated with light waves arises when the electromagnetic fields that constitute the individual waves combine.

Interference in light waves from two sources was first demonstrated by Thomas Young in 1801. A schematic diagram of the apparatus Young used is shown in Figure 37.1a. Plane light waves arrive at a barrier that contains two slits S_1 and S_2. The light from S_1 and S_2 produces on a viewing screen a visible pattern of bright and dark parallel bands called **fringes** (Fig. 37.1b). When the light from S_1 and that from S_2 both arrive at a point on the screen such that constructive interference occurs at

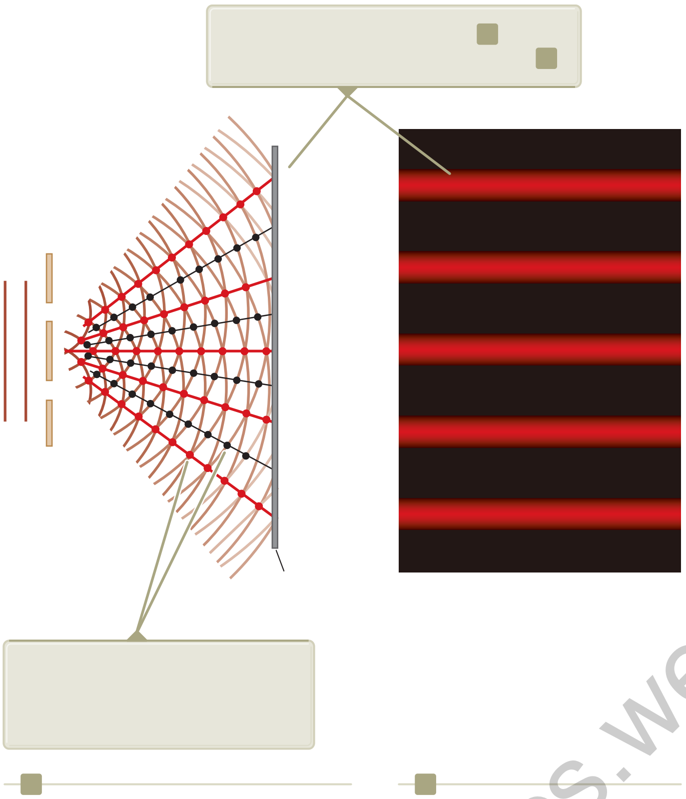


Figure 37.1 (a) Schematic dia gram of Young's double-slit experi ment. Slits S and S behave as coherent sources of light waves that produce an interference pattern on the viewing screen (drawing not to scale). (b) A simulation of an enlargement of the center of a fringe pattern formed on the view ing screen.

that location, a bright fringe appears. When the light from the two slits combines destructively at any location on the screen, a dark fringe results.

Figure 37.2 is a photograph looking down on an interference pattern produced on the surface of a water tank by two vibrating sources. The linear regions of con structive interference, such as at , and destructive interference, such as at , radi ating from the area between the sources are analogous to the red and black lines in Figure 37.1a.

Figure 37.3 on page 1136 shows some of the ways in which two waves can com bine at the screen. In Figure 37.3a, the two waves, which leave the two slits in phase, strike the screen at the central point Because both waves travel the same distance, they arrive at in phase. As a result, constructive interference occurs at this location and a bright fringe is observed. In Figure 37.3b, the two waves also start in phase, but here the lower wave has to travel one wavelength farther than the upper wave to reach point Because the lower wave falls behind

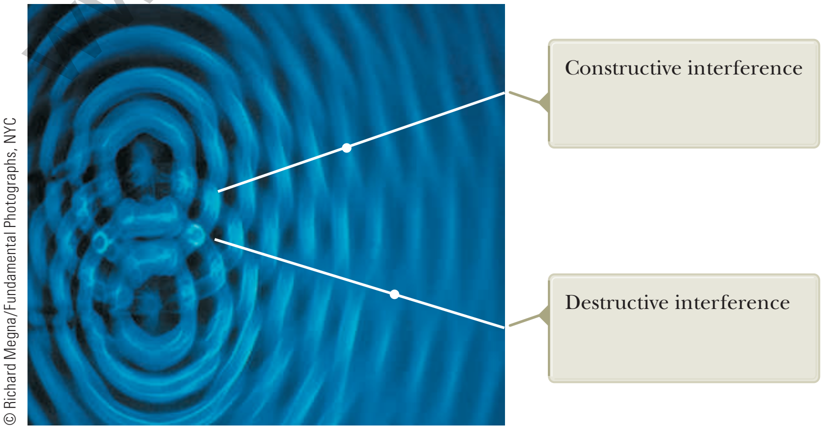


Figure 37.2 An interference pattern involving water waves is produced by two vibrating sources at the water's surface.

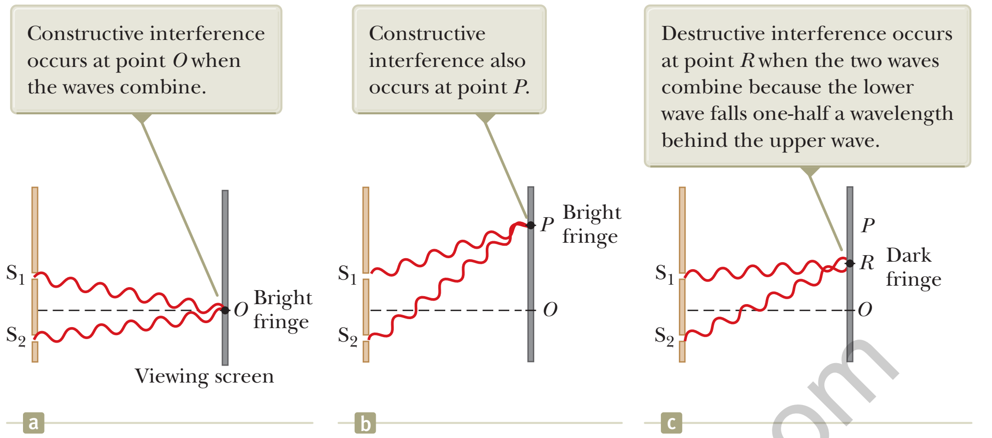


Figure 37.3 Waves leave the slits and combine at various points on the viewing screen. (All figures not to scale.)

the upper one by exactly one wavelength, they still arrive in phase at *P* and a second bright fringe appears at this location. At point *R* in Figure 37.3c, however, between points *O* and *P*, the lower wave has fallen half a wavelength behind the upper wave and a crest of the upper wave overlaps a trough of the lower wave, giving rise to destructive interference at point *R*. A dark fringe is therefore observed at this location.

If two lightbulbs are placed side by side so that light from both bulbs combines, no interference effects are observed because the light waves from one bulb are emitted independently of those from the other bulb. The emissions from the two lightbulbs do not maintain a constant phase relationship with each other over time. Light waves from an ordinary source such as a lightbulb undergo random phase changes in time intervals of less than a nanosecond. Therefore, the conditions for constructive interference, destructive interference, or some intermediate state are maintained only for such short time intervals. Because the eye cannot follow such rapid changes, no interference effects are observed. Such light sources are said to be **incoherent.**

To observe interference of waves from two sources, the following conditions must be met:

Conditions for interference ▶

- The sources must be **coherent;** that is, they must maintain a constant phase with respect to each other.
- The sources should be **monochromatic;** that is, they should be of a single wavelength.

As an example, single-frequency sound waves emitted by two side-by-side loudspeakers driven by a single amplifier can interfere with each other because the two speakers are coherent. In other words, they respond to the amplifier in the same way at the same time.

A common method for producing two coherent light sources is to use a monochromatic source to illuminate a barrier containing two small openings, usually in the shape of slits, as in the case of Young's experiment illustrated in Figure 37.1. The light emerging from the two slits is coherent because a single source produces the original light beam and the two slits serve only to separate the original beam into two parts (which, after all, is what is done to the sound signal from two side-by-side loudspeakers). Any random change in the light emitted by the source occurs in both beams at the same time. As a result, interference effects can be observed when the light from the two slits arrives at a viewing screen.

If the light traveled only in its original direction after passing through the slits as shown in Figure 37.4a, the waves would not overlap and no interference pattern would be seen. Instead, as we have discussed in our treatment of Huygens's principle (Section 35.6), the waves spread out from the slits as shown in Figure 37.4b. In other words, the light deviates from a straight-line path and enters the region that

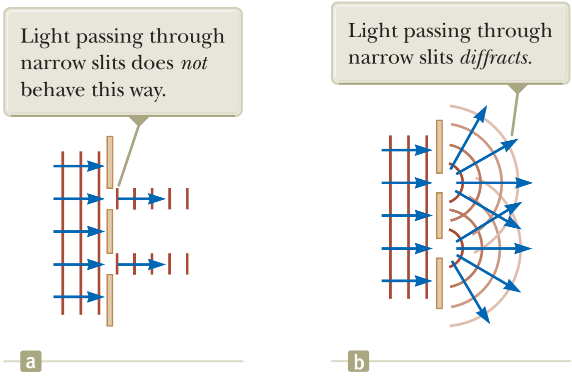


Figure 37.4 (a) If light waves did not spread out after passing through the slits, no interference would occur. (b) The light waves from the two slits overlap as they spread out, filling what we expect to be shadowed regions with light and producing interference fringes on a screen placed to the right of the slits.

would otherwise be shadowed. As noted in Section 35.3, this divergence of light from its initial line of travel is called **diffraction.**

37.2 Analysis Model: Waves in Interference

We discussed the superposition principle for waves on strings in Section 18.1, leading to a one-dimensional version of the waves in interference analysis model. In Example 18.1 on page 537, we briefly discussed a two-dimensional interference phenomenon for sound from two loudspeakers. In walking from point O to point P in Figure 18.5, the listener experienced a maximum in sound intensity at O and a minimum at P. This experience is exactly analogous to an observer looking at point O in Figure 37.3 and seeing a bright fringe and then sweeping his eyes upward to point R, where there is a minimum in light intensity.

Let's look in more detail at the two-dimensional nature of Young's experiment with the help of Figure 37.5. The viewing screen is located a perpendicular distance L from the barrier containing two slits, S_1 and S_2 (Fig. 37.5a). These slits are separated by a distance d, and the source is monochromatic. To reach any arbitrary point P in the upper half of the screen, a wave from the lower slit must travel farther than a wave from the upper slit. The extra distance traveled from the lower slit is the **path difference** δ (Greek letter delta). If we assume the rays labeled r_1 and r_2 are parallel (Fig. 37.5b), which is approximately true if L is much greater than d, then δ is given by

$$\delta = r_2 - r_1 = d \sin \theta \tag{37.1}$$

The value of δ determines whether the two waves are in phase when they arrive at point P. If δ is either zero or some integer multiple of the wavelength, the two waves

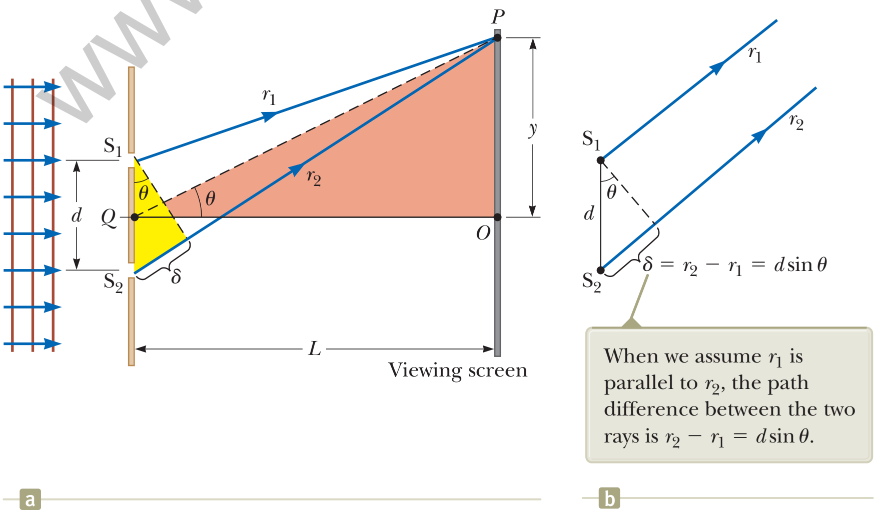


Figure 37.5 (a) Geometric construction for describing Young's double-slit experiment (not to scale). (b) The slits are represented as sources, and the outgoing light rays are assumed to be parallel as they travel to P. To achieve that in practice, it is essential that $L >> d$.

are in phase at point P and constructive interference results. Therefore, the condition for bright fringes, or **constructive interference,** at point P is

Condition for constructive interference

$$d \sin \theta_{\text{bright}} = m\lambda \qquad m = 0, \pm 1, \pm 2, \ldots \tag{37.2}$$

The number m is called the **order number.** For constructive interference, the order number is the same as the number of wavelengths that represents the path difference between the waves from the two slits. The central bright fringe at $\theta_{\text{bright}} = 0$ is called the *zeroth-order maximum*. The first maximum on either side, where $m = \pm 1$, is called the *first-order maximum,* and so forth.

When δ is an odd multiple of $\lambda/2$, the two waves arriving at point P are 180° out of phase and give rise to destructive interference. Therefore, the condition for dark fringes, or **destructive interference,** at point P is

Condition for destructive interference

$$d \sin \theta_{\text{dark}} = (m + \tfrac{1}{2})\lambda \qquad m = 0, \pm 1, \pm 2, \ldots \tag{37.3}$$

These equations provide the *angular* positions of the fringes. It is also useful to obtain expressions for the *linear* positions measured along the screen from O to P. From the triangle OPQ in Figure 37.5a, we see that

$$\tan \theta = \frac{y}{L} \tag{37.4}$$

Using this result, the linear positions of bright and dark fringes are given by

$$y_{\text{bright}} = L \tan \theta_{\text{bright}} \tag{37.5}$$

$$y_{\text{dark}} = L \tan \theta_{\text{dark}} \tag{37.6}$$

where θ_{bright} and θ_{dark} are given by Equations 37.2 and 37.3.

When the angles to the fringes are small, the positions of the fringes are linear near the center of the pattern. That can be verified by noting that for small angles, $\tan \theta \approx \sin \theta$, so Equation 37.5 gives the positions of the bright fringes as $y_{\text{bright}} = L \sin \theta_{\text{bright}}$. Incorporating Equation 37.2 gives

$$y_{\text{bright}} = L \frac{m\lambda}{d} \quad \text{(small angles)} \tag{37.7}$$

This result shows that y_{bright} is linear in the order number m, so the fringes are equally spaced for small angles. Similarly, for dark fringes,

$$y_{\text{dark}} = L \frac{(m + \frac{1}{2})\lambda}{d} \quad \text{(small angles)} \tag{37.8}$$

As demonstrated in Example 37.1, Young's double-slit experiment provides a method for measuring the wavelength of light. In fact, Young used this technique to do precisely that. In addition, his experiment gave the wave model of light a great deal of credibility. It was inconceivable that particles of light coming through the slits could cancel one another in a way that would explain the dark fringes.

The principles discussed in this section are the basis of the **waves in interference** analysis model. This model was applied to mechanical waves in one dimension in Chapter 18. Here we see the details of applying this model in three dimensions to light.

Quick Quiz 37.1 Which of the following causes the fringes in a two-slit interference pattern to move farther apart? **(a)** decreasing the wavelength of the light **(b)** decreasing the screen distance L **(c)** decreasing the slit spacing d **(d)** immersing the entire apparatus in water

Analysis Model Waves in Interference

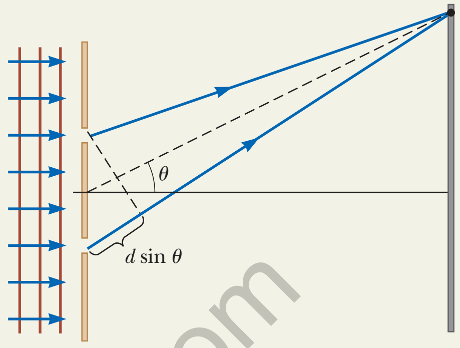
Imagine a broad beam of light that illuminates a double slit in an otherwise opaque material. An interference pattern of bright and dark fringes is created on a distant screen. The condition for bright fringes **(constructive interference)** is

$$d \sin \theta_{\text{bright}} = m\lambda \quad m = 0, \pm 1, \pm 2, \ldots \tag{37.2}$$

The condition for dark fringes **(destructive interference)** is

$$d \sin \theta_{\text{dark}} = \left(m + \tfrac{1}{2}\right)\lambda \quad m = 0, \pm 1, \pm 2, \ldots \tag{37.3}$$

The number m is called the **order number** of the fringe.



Examples:

- a thin film of oil on top of water shows swirls of color (Section 37.5)
- x-rays passing through a crystalline solid combine to form a Laue pattern (Chapter 38)
- a Michelson interferometer (Section 37.6) is used to search for the ether representing the medium through which light travels (Chapter 39)
- electrons exhibit interference just like light waves when they pass through a double slit (Chapter 40)

Example 37.1 Measuring the Wavelength of a Light Source AM

A viewing screen is separated from a double slit by 4.80 m. The distance between the two slits is 0.030 0 mm. Monochromatic light is directed toward the double slit and forms an interference pattern on the screen. The first dark fringe is 4.50 cm from the center line on the screen.

(A) Determine the wavelength of the light.

SOLUTION

Conceptualize Study Figure 37.5 to be sure you understand the phenomenon of interference of light waves. The distance of 4.50 cm is y in Figure 37.5. Because $L >> y$, the angles for the fringes are small.

Categorize This problem is a simple application of the *waves in interference* model.

Analyze

Solve Equation 37.8 for the wavelength and substitute numerical values, taking $m = 0$ for the first dark fringe:

$$\lambda = \frac{y_{\text{dark}}d}{\left(m + \frac{1}{2}\right)L} = \frac{(4.50 \times 10^{-2}\text{ m})(3.00 \times 10^{-5}\text{ m})}{\left(0 + \frac{1}{2}\right)(4.80\text{ m})}$$

$$= 5.62 \times 10^{-7}\text{ m} = \boxed{562\text{ nm}}$$

(B) Calculate the distance between adjacent bright fringes.

SOLUTION

Find the distance between adjacent bright fringes from Equation 37.7 and the results of part (A):

$$y_{m+1} - y_m = L\frac{(m+1)\lambda}{d} - L\frac{m\lambda}{d}$$

$$= L\frac{\lambda}{d} = 4.80\text{ m}\left(\frac{5.62 \times 10^{-7}\text{ m}}{3.00 \times 10^{-5}\text{ m}}\right)$$

$$= 9.00 \times 10^{-2}\text{ m} = \boxed{9.00\text{ cm}}$$

Finalize For practice, find the wavelength of the sound in Example 18.1 using the procedure in part (A) of this example.

Example 37.2 Separating Double-Slit Fringes of Two Wavelengths AM

A light source emits visible light of two wavelengths: $\lambda = 430$ nm and $\lambda' = 510$ nm. The source is used in a double-slit interference experiment in which $L = 1.50$ m and $d = 0.025\,0$ mm. Find the separation distance between the third-order bright fringes for the two wavelengths.

SOLUTION

Conceptualize In Figure 37.5a, imagine light of two wavelengths incident on the slits and forming two interference patterns on the screen. At some points, the fringes of the two colors might overlap, but at most points, they will not.

Categorize This problem is an application of the mathematical representation of the *waves in interference* analysis model.

Analyze

Use Equation 37.7 to find the fringe positions corresponding to these two wavelengths and subtract them:

$$\Delta y = y'_{\text{bright}} - y_{\text{bright}} = L\frac{m\lambda'}{d} - L\frac{m\lambda}{d} = \frac{Lm}{d}(\lambda' - \lambda)$$

Substitute numerical values:

$$\Delta y = \frac{(1.50\text{ m})(3)}{0.025\,0 \times 10^{-3}\text{ m}}(510 \times 10^{-9}\text{ m} - 430 \times 10^{-9}\text{ m})$$

$$= 0.014\,4\text{ m} = \boxed{1.44\text{ cm}}$$

Finalize Let's explore further details of the interference pattern in the following **What If?**

WHAT IF? What if we examine the entire interference pattern due to the two wavelengths and look for overlapping fringes? Are there any locations on the screen where the bright fringes from the two wavelengths overlap exactly?

Answer Find such a location by setting the location of any bright fringe due to λ equal to one due to λ', using Equation 37.7:

$$L\frac{m\lambda}{d} = L\frac{m'\lambda'}{d} \rightarrow \frac{m'}{m} = \frac{\lambda}{\lambda'}$$

Substitute the wavelengths:

$$\frac{m'}{m} = \frac{430\text{ nm}}{510\text{ nm}} = \frac{43}{51}$$

Therefore, the 51st fringe of the 430-nm light overlaps with the 43rd fringe of the 510-nm light.

Use Equation 37.7 to find the value of y for these fringes:

$$y = (1.50\text{ m})\left[\frac{51(430 \times 10^{-9}\text{ m})}{0.025\,0 \times 10^{-3}\text{ m}}\right] = 1.32\text{ m}$$

This value of y is comparable to L, so the small-angle approximation used for Equation 37.7 is *not* valid. This conclusion suggests we should not expect Equation 37.7 to give us the correct result. If you use Equation 37.5, you can show that the bright fringes do indeed overlap when the same condition, $m'/m = \lambda/\lambda'$, is met (see Problem 48). Therefore, the 51st fringe of the 430-nm light does overlap with the 43rd fringe of the 510-nm light, but not at the location of 1.32 m. You are asked to find the correct location as part of Problem 48.

37.3 Intensity Distribution of the Double-Slit Interference Pattern

Notice that the edges of the bright fringes in Figure 37.1b are not sharp; rather, there is a gradual change from bright to dark. So far, we have discussed the locations of only the centers of the bright and dark fringes on a distant screen. Let's now direct our attention to the intensity of the light at other points between the positions of maximum constructive and destructive interference. In other words, we now calculate the distribution of light intensity associated with the double-slit interference pattern.

Again, suppose the two slits represent coherent sources of sinusoidal waves such that the two waves from the slits have the same angular frequency ω and are in phase. The total magnitude of the electric field at point P on the screen in Figure 37.5 is the superposition of the two waves. Assuming the two waves have the same amplitude E_0, we can write the magnitude of the electric field at point P due to each wave separately as

$$E_1 = E_0 \sin \omega t \quad \text{and} \quad E_2 = E_0 \sin (\omega t + \phi) \tag{37.9}$$

Although the waves are in phase at the slits, their phase difference ϕ at P depends on the path difference $\delta = r_2 - r_1 = d \sin \theta$. A path difference of λ (for constructive interference) corresponds to a phase difference of 2π rad. A path difference of δ is the same fraction of λ as the phase difference ϕ is of 2π. We can describe this fraction mathematically with the ratio

$$\frac{\delta}{\lambda} = \frac{\phi}{2\pi}$$

which gives

$$\phi = \frac{2\pi}{\lambda}\delta = \frac{2\pi}{\lambda} d \sin \theta \tag{37.10}$$

◀ **Phase difference**

This equation shows how the phase difference ϕ depends on the angle θ in Figure 37.5.

Using the superposition principle and Equation 37.9, we obtain the following expression for the magnitude of the resultant electric field at point P:

$$E_P = E_1 + E_2 = E_0[\sin \omega t + \sin (\omega t + \phi)] \tag{37.11}$$

We can simplify this expression by using the trigonometric identity

$$\sin A + \sin B = 2 \sin \left(\frac{A + B}{2}\right) \cos \left(\frac{A - B}{2}\right)$$

Taking $A = \omega t + \phi$ and $B = \omega t$, Equation 37.11 becomes

$$E_P = 2E_0 \cos \left(\frac{\phi}{2}\right) \sin \left(\omega t + \frac{\phi}{2}\right) \tag{37.12}$$

This result indicates that the electric field at point P has the same frequency ω as the light at the slits but that the amplitude of the field is multiplied by the factor $2 \cos (\phi/2)$. To check the consistency of this result, note that if $\phi = 0, 2\pi, 4\pi, \ldots$, the magnitude of the electric field at point P is $2E_0$, corresponding to the condition for maximum constructive interference. These values of ϕ are consistent with Equation 37.2 for constructive interference. Likewise, if $\phi = \pi, 3\pi, 5\pi, \ldots$, the magnitude of the electric field at point P is zero, which is consistent with Equation 37.3 for total destructive interference.

Finally, to obtain an expression for the light intensity at point P, recall from Section 34.4 that the intensity of a wave is proportional to the square of the resultant electric field magnitude at that point (Eq. 34.24). Using Equation 37.12, we can therefore express the light intensity at point P as

$$I \propto E_P^2 = 4E_0^2 \cos^2 \left(\frac{\phi}{2}\right) \sin^2 \left(\omega t + \frac{\phi}{2}\right)$$

Most light-detecting instruments measure time-averaged light intensity, and the time-averaged value of $\sin^2 (\omega t + \phi/2)$ over one cycle is $\frac{1}{2}$. (See Fig. 33.5.) Therefore, we can write the average light intensity at point P as

$$I = I_{\max} \cos^2 \left(\frac{\phi}{2}\right) \tag{37.13}$$

where I_{max} is the maximum intensity on the screen and the expression represents the time average. Substituting the value for ϕ given by Equation 37.10 into this expression gives

$$I = I_{max} \cos^2\left(\frac{\pi d \sin\theta}{\lambda}\right) \quad \textbf{(37.14)}$$

Alternatively, because $\sin\theta \approx y/L$ for small values of θ in Figure 37.5, we can write Equation 37.14 in the form

$$I = I_{max} \cos^2\left(\frac{\pi d}{\lambda L} y\right) \quad \text{(small angles)} \quad \textbf{(37.15)}$$

Constructive interference, which produces light intensity maxima, occurs when the quantity $\pi dy/\lambda L$ is an integral multiple of π, corresponding to $y = (\lambda L/d)m$, where m is the order number. This result is consistent with Equation 37.7.

A plot of light intensity versus $d \sin\theta$ is given in Figure 37.6. The interference pattern consists of equally spaced fringes of equal intensity.

Figure 37.7 shows similar plots of light intensity versus $d \sin\theta$ for light passing through multiple slits. For more than two slits, we would add together more electric field magnitudes than the two in Equation 37.9. In this case, the pattern contains primary and secondary maxima. For three slits, notice that the primary maxima are nine times more intense than the secondary maxima as measured by the height of the curve because the intensity varies as E^2. For N slits, the intensity of the primary maxima is N^2 times greater than that for the secondary maxima. As the number of slits increases, the primary maxima increase in intensity and become narrower, while the secondary maxima decrease in intensity relative to the primary maxima. Figure 37.7 also shows that as the number of slits increases, the number of secondary maxima also increases. In fact, the number of secondary maxima is always $N - 2$, where N is the number of slits. In Section 38.4, we shall investigate the pattern for a very large number of slits in a device called a *diffraction grating*.

Quick Quiz 37.2 Using Figure 37.7 as a model, sketch the interference pattern from six slits.

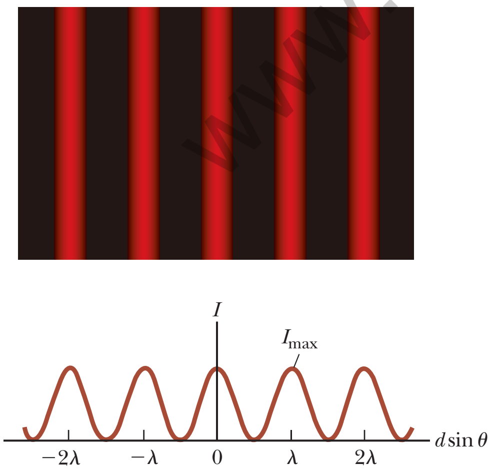


Figure 37.6 Light intensity versus $d \sin\theta$ for a double-slit interference pattern when the screen is far from the two slits ($L >> d$).

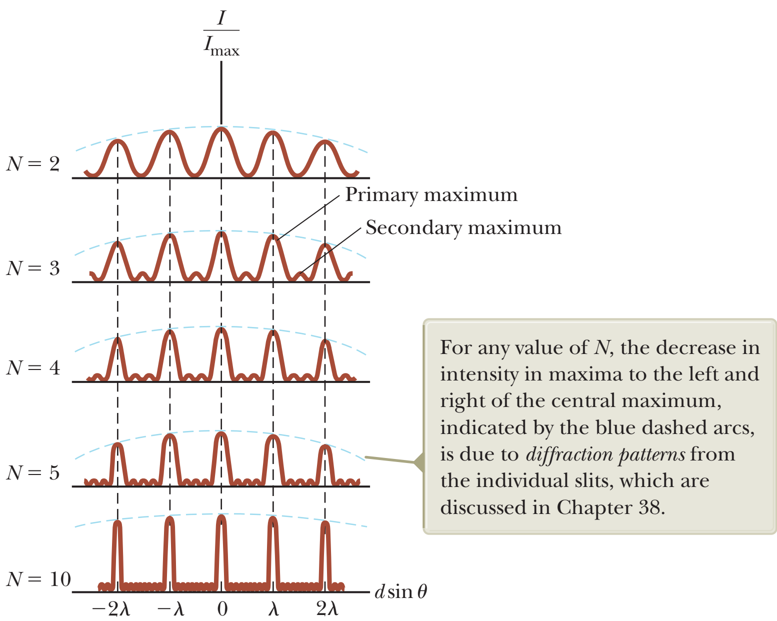


Figure 37.7 Multiple-slit interference patterns. As N, the number of slits, is increased, the primary maxima (the tallest peaks in each graph) become narrower but remain fixed in position and the number of secondary maxima increases.

37.4 Change of Phase Due to Reflection

Young's method for producing two coherent light sources involves illuminating a pair of slits with a single source. Another simple, yet ingenious, arrangement for producing an interference pattern with a single light source is known as *Lloyd's mirror*[1] (Fig. 37.8). A point light source S is placed close to a mirror, and a viewing screen is positioned some distance away and perpendicular to the mirror. Light waves can reach point P on the screen either directly from S to P or by the path involving reflection from the mirror. The reflected ray can be treated as a ray originating from a virtual source S′. As a result, we can think of this arrangement as a double-slit source where the distance d between sources S and S′ in Figure 37.8 is analogous to length d in Figure 37.5. Hence, at observation points far from the source ($L >> d$), we expect waves from S and S′ to form an interference pattern exactly like the one formed by two real coherent sources. An interference pattern is indeed observed. The positions of the dark and bright fringes, however, are reversed relative to the pattern created by two real coherent sources (Young's experiment). Such a reversal can only occur if the coherent sources S and S′ differ in phase by 180°.

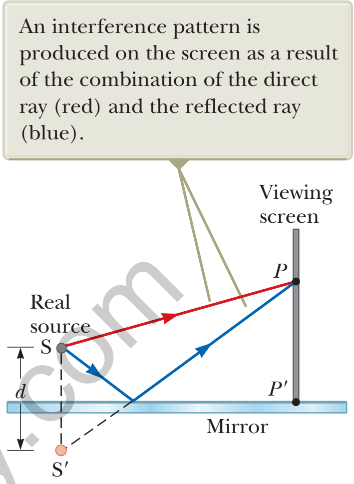


Figure 37.8 Lloyd's mirror. The reflected ray undergoes a phase change of 180°.

To illustrate further, consider point P', the point where the mirror intersects the screen. This point is equidistant from sources S and S′. If path difference alone were responsible for the phase difference, we would see a bright fringe at P' (because the path difference is zero for this point), corresponding to the central bright fringe of the two-slit interference pattern. Instead, a dark fringe is observed at P'. We therefore conclude that a 180° phase change must be produced by reflection from the mirror. In general, an electromagnetic wave undergoes a phase change of 180° upon reflection from a medium that has a higher index of refraction than the one in which the wave is traveling.

It is useful to draw an analogy between reflected light waves and the reflections of a transverse pulse on a stretched string (Section 16.4). The reflected pulse on a string undergoes a phase change of 180° when reflected from the boundary of a denser string or a rigid support, but no phase change occurs when the pulse is reflected from the boundary of a less dense string or a freely-supported end. Similarly, an electromagnetic wave undergoes a 180° phase change when reflected from a boundary leading to an optically denser medium (defined as a medium with a higher index of refraction), but no phase change occurs when the wave is reflected from a boundary leading to a less dense medium. These rules, summarized in Figure 37.9, can be deduced from Maxwell's equations, but the treatment is beyond the scope of this text.

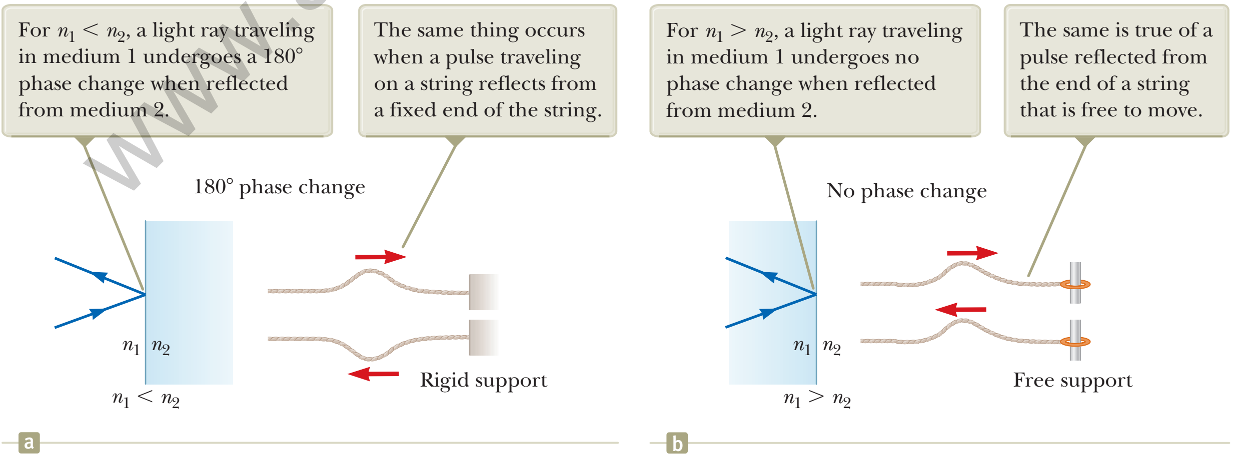


Figure 37.9 Comparisons of reflections of light waves and waves on strings.

[1]Developed in 1834 by Humphrey Lloyd (1800–1881), Professor of Natural and Experimental Philosophy, Trinity College, Dublin.

37.5 Interference in Thin Films

Interference effects are commonly observed in thin films, such as thin layers of oil on water or the thin surface of a soap bubble. The varied colors observed when white light is incident on such films result from the interference of waves reflected from the two surfaces of the film.

Consider a film of uniform thickness t and index of refraction n. The wavelength of light λ_n in the film (see Section 35.5) is

$$\lambda_n = \frac{\lambda}{n}$$

where λ is the wavelength of the light in free space and n is the index of refraction of the film material. Let's assume light rays traveling in air are nearly normal to the two surfaces of the film as shown in Figure 37.10.

Reflected ray 1, which is reflected from the upper surface (A) in Figure 37.10, undergoes a phase change of 180° with respect to the incident wave. Reflected ray 2, which is reflected from the lower film surface (B), undergoes no phase change because it is reflected from a medium (air) that has a lower index of refraction. Therefore, ray 1 is 180° out of phase with ray 2, which is equivalent to a path difference of $\lambda_n/2$. We must also consider, however, that ray 2 travels an extra distance $2t$ before the waves recombine in the air above surface A. (Remember that we are considering light rays that are close to normal to the surface. If the rays are not close to normal, the path difference is larger than $2t$.) If $2t = \lambda_n/2$, rays 1 and 2 recombine in phase and the result is constructive interference. In general, the condition for *constructive* interference in thin films is[2]

$$2t = (m + \tfrac{1}{2})\lambda_n \quad m = 0, 1, 2, \ldots \tag{37.16}$$

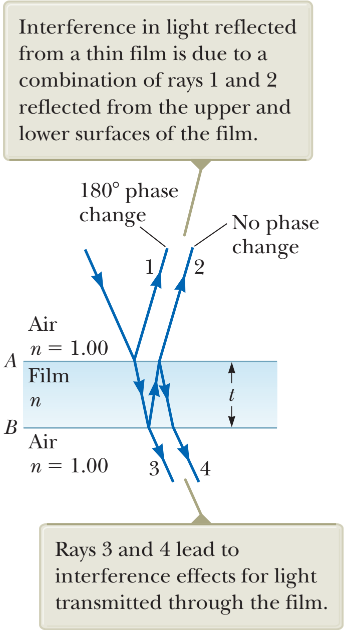


Figure 37.10 Light paths through a thin film.

This condition takes into account two factors: (1) the difference in path length for the two rays (the term $m\lambda_n$) and (2) the 180° phase change upon reflection (the term $\frac{1}{2}\lambda_n$). Because $\lambda_n = \lambda/n$, we can write Equation 37.16 as

$$2nt = (m + \tfrac{1}{2})\lambda \quad m = 0, 1, 2, \ldots \tag{37.17}$$

If the extra distance $2t$ traveled by ray 2 corresponds to a multiple of λ_n, the two waves combine out of phase and the result is destructive interference. The general equation for *destructive* interference in thin films is

$$2nt = m\lambda \quad m = 0, 1, 2, \ldots \tag{37.18}$$

The foregoing conditions for constructive and destructive interference are valid when the medium above the top surface of the film is the same as the medium below the bottom surface or, if there are different media above and below the film, the index of refraction of both is less than n. If the film is placed between two different media, one with $n < n_{\text{film}}$ and the other with $n > n_{\text{film}}$, the conditions for constructive and destructive interference are reversed. In that case, either there is a phase change of 180° for both ray 1 reflecting from surface A and ray 2 reflecting from surface B or there is no phase change for either ray; hence, the net change in relative phase due to the reflections is zero.

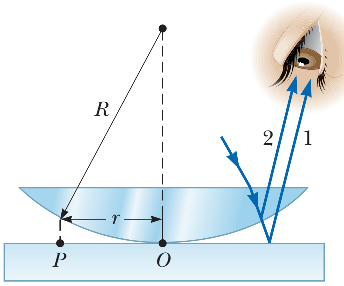
Rays 3 and 4 in Figure 37.10 lead to interference effects in the light transmitted through the thin film. The analysis of these effects is similar to that of the reflected light. You are asked to explore the transmitted light in Problems 35, 36, and 38.

Pitfall Prevention 37.1

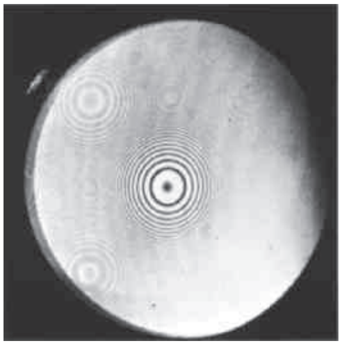
Be Careful with Thin Films Be sure to include *both* effects—path length and phase change—when analyzing an interference pattern resulting from a thin film. The possible phase change is a new feature we did not need to consider for double-slit interference. Also think carefully about the material on either side of the film. If there are different materials on either side of the film, you may have a situation in which there is a 180° phase change at *both* surfaces or at *neither* surface.

Quick Quiz 37.3 One microscope slide is placed on top of another with their left edges in contact and a human hair under the right edge of the upper slide. As a result, a wedge of air exists between the slides. An interference pattern results when monochromatic light is incident on the wedge. What is at the left edges of the slides? **(a)** a dark fringe **(b)** a bright fringe **(c)** impossible to determine

[2]The full interference effect in a thin film requires an analysis of an infinite number of reflections back and forth between the top and bottom surfaces of the film. We focus here only on a single reflection from the bottom of the film, which provides the largest contribution to the interference effect.



a



b

Courtesy of Bausch and Lomb

Figure 37.11 (a) The combination of rays reflected from the flat plate and the curved lens surface gives rise to an interference pattern known as Newton's rings. (b) Photograph of Newton's rings.

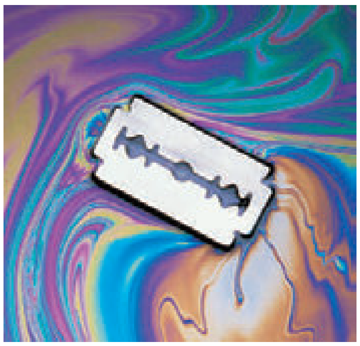
Newton's Rings

Another method for observing interference in light waves is to place a plano-convex lens on top of a flat glass surface as shown in Figure 37.11a. With this arrangement, the air film between the glass surfaces varies in thickness from zero at the point of contact to some nonzero value at point P. If the radius of curvature R of the lens is much greater than the distance r and the system is viewed from above, a pattern of light and dark rings is observed as shown in Figure 37.11b. These circular fringes, discovered by Newton, are called **Newton's rings.**

The interference effect is due to the combination of ray 1, reflected from the flat plate, with ray 2, reflected from the curved surface of the lens. Ray 1 undergoes a phase change of 180° upon reflection (because it is reflected from a medium of higher index of refraction), whereas ray 2 undergoes no phase change (because it is reflected from a medium of lower index of refraction). Hence, the conditions for constructive and destructive interference are given by Equations 37.17 and 37.18, respectively, with $n = 1$ because the film is air. Because there is no path difference and the total phase change is due only to the 180° phase change upon reflection, the contact point at O is dark as seen in Figure 37.11b.

Using the geometry shown in Figure 37.11a, we can obtain expressions for the radii of the bright and dark bands in terms of the radius of curvature R and wavelength λ. For example, the dark rings have radii given by the expression $r \approx \sqrt{m\lambda R/n}$. The details are left as a problem (see Problem 66). We can obtain the wavelength of the light causing the interference pattern by measuring the radii of the rings, provided R is known. Conversely, we can use a known wavelength to obtain R.

One important use of Newton's rings is in the testing of optical lenses. A circular pattern like that pictured in Figure 37.11b is obtained only when the lens is ground to a perfectly symmetric curvature. Variations from such symmetry produce a pattern with fringes that vary from a smooth, circular shape. These variations indicate how the lens must be reground and repolished to remove imperfections.



a

Peter Aprahamian/Photo Researchers, Inc.

b

Dr. Jeremy Burgess/Science Photo Library/Photo Researchers, Inc.

(a) A thin film of oil floating on water displays interference, shown by the pattern of colors when white light is incident on the film. Variations in film thickness produce the interesting color pattern. The razor blade gives you an idea of the size of the colored bands. (b) Interference in soap bubbles. The colors are due to interference between light rays reflected from the inner and outer surfaces of the thin film of soap making up the bubble. The color depends on the thickness of the film, ranging from black, where the film is thinnest, to magenta, where it is thickest.

Problem-Solving Strategy Thin-Film Interference

The following features should be kept in mind when working thin-film interference problems.

1. Conceptualize. Think about what is going on physically in the problem. Identify the light source and the location of the observer.

2. Categorize. Confirm that you should use the techniques for thin-film interference by identifying the thin film causing the interference.

3. Analyze. The type of interference that occurs is determined by the phase relationship between the portion of the wave reflected at the upper surface of the film and the portion reflected at the lower surface. Phase differences between the two portions of the wave have two causes: differences in the distances traveled by the two portions and phase changes occurring on reflection. *Both* causes must be considered when determining which type of interference occurs. If the media above and below the film both have index of refraction larger than that of the film or if both indices are smaller, use Equation 37.17 for constructive interference and Equation 37.18 for destructive interference. If the film is located between two different media, one with $n < n_{\text{film}}$ and the other with $n > n_{\text{film}}$, reverse these two equations for constructive and destructive interference.

4. Finalize. Inspect your final results to see if they make sense physically and are of an appropriate size.

Example 37.3 Interference in a Soap Film

Calculate the minimum thickness of a soap-bubble film that results in constructive interference in the reflected light if the film is illuminated with light whose wavelength in free space is $\lambda = 600$ nm. The index of refraction of the soap film is 1.33.

SOLUTION

Conceptualize Imagine that the film in Figure 37.10 is soap, with air on both sides.

Categorize We determine the result using an equation from this section, so we categorize this example as a substitution problem.

The minimum film thickness for constructive interference in the reflected light corresponds to $m = 0$ in Equation 37.17. Solve this equation for t and substitute numerical values:

$$t = \frac{(0 + \frac{1}{2})\lambda}{2n} = \frac{\lambda}{4n} = \frac{(600 \text{ nm})}{4(1.33)} = \boxed{113 \text{ nm}}$$

WHAT IF? What if the film is twice as thick? Does this situation produce constructive interference?

Answer Using Equation 37.17, we can solve for the thicknesses at which constructive interference occurs:

$$t = (m + \tfrac{1}{2})\frac{\lambda}{2n} = (2m + 1)\frac{\lambda}{4n} \quad m = 0, 1, 2, \ldots$$

The allowed values of m show that constructive interference occurs for *odd* multiples of the thickness corresponding to $m = 0$, $t = 113$ nm. Therefore, constructive interference does *not* occur for a film that is twice as thick.

Example 37.4 Nonreflective Coatings for Solar Cells

Solar cells—devices that generate electricity when exposed to sunlight—are often coated with a transparent, thin film of silicon monoxide (SiO, $n = 1.45$) to minimize reflective losses from the surface. Suppose a silicon solar cell ($n = 3.5$) is coated with a thin film of silicon monoxide for this purpose (Fig. 37.12a). Determine the minimum film thickness that produces the least reflection at a wavelength of 550 nm, near the center of the visible spectrum.

37.4 continued

SOLUTION

Conceptualize Figure 37.12a helps us visualize the path of the rays in the SiO film that result in interference in the reflected light.

Categorize Based on the geometry of the SiO layer, we categorize this example as a thin-film interference problem.

Analyze The reflected light is a minimum when rays 1 and 2 in Figure 37.12a meet the condition of destructive interference. In this situation, *both* rays undergo a 180° phase change upon reflection: ray 1 from the upper SiO surface and ray 2 from the lower SiO surface. The net change in phase due to reflection is therefore zero, and the condition for a reflection minimum requires a path difference of $\lambda_n/2$, where λ_n is the wavelength of the light in SiO. Hence, $2nt = \lambda/2$, where λ is the wavelength in air and n is the index of refraction of SiO.

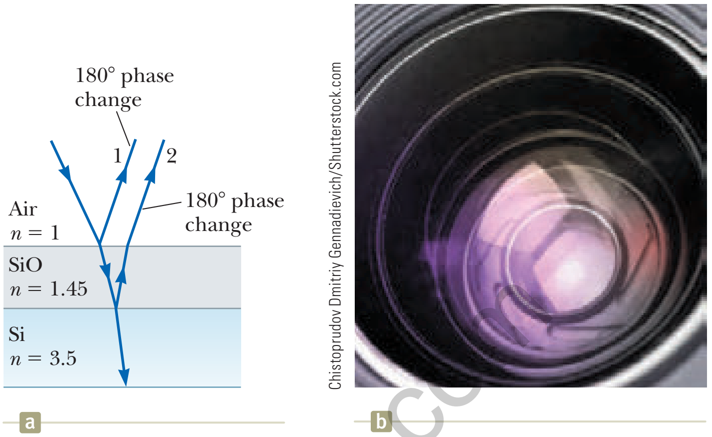


Figure 37.12 (Example 37.4) (a) Reflective losses from a silicon solar cell are minimized by coating the surface of the cell with a thin film of silicon monoxide. (b) The reflected light from a coated camera lens often has a reddish-violet appearance.

Solve the equation $2nt = \lambda/2$ for t and substitute numerical values:

$$t = \frac{\lambda}{4n} = \frac{550 \text{ nm}}{4(1.45)} = \boxed{94.8 \text{ nm}}$$

Finalize A typical uncoated solar cell has reflective losses as high as 30%, but a coating of SiO can reduce this value to about 10%. This significant decrease in reflective losses increases the cell's efficiency because less reflection means that more sunlight enters the silicon to create charge carriers in the cell. No coating can ever be made perfectly nonreflecting because the required thickness is wavelength-dependent and the incident light covers a wide range of wavelengths.

Glass lenses used in cameras and other optical instruments are usually coated with a transparent thin film to reduce or eliminate unwanted reflection and to enhance the transmission of light through the lenses. The camera lens in Figure 37.12b has several coatings (of different thicknesses) to minimize reflection of light waves having wavelengths near the center of the visible spectrum. As a result, the small amount of light that is reflected by the lens has a greater proportion of the far ends of the spectrum and often appears reddish violet.

37.6 The Michelson Interferometer

The **interferometer,** invented by American physicist A. A. Michelson (1852–1931), splits a light beam into two parts and then recombines the parts to form an interference pattern. The device can be used to measure wavelengths or other lengths with great precision because a large and precisely measurable displacement of one of the mirrors is related to an exactly countable number of wavelengths of light.

A schematic diagram of the interferometer is shown in Figure 37.13 (page 1148). A ray of light from a monochromatic source is split into two rays by mirror M_0, which is inclined at 45° to the incident light beam. Mirror M_0, called a *beam splitter,* transmits half the light incident on it and reflects the rest. One ray is reflected from M_0 to the right toward mirror M_1, and the second ray is transmitted vertically through M_0 toward mirror M_2. Hence, the two rays travel separate paths L_1 and L_2. After reflecting from M_1 and M_2, the two rays eventually recombine at M_0 to produce an interference pattern, which can be viewed through a telescope.

The interference condition for the two rays is determined by the difference in their path length. When the two mirrors are exactly perpendicular to each other, the interference pattern is a target pattern of bright and dark circular fringes. As M_1 is moved, the fringe pattern collapses or expands, depending on the direction in which M_1 is moved. For example, if a dark circle appears at the center of the

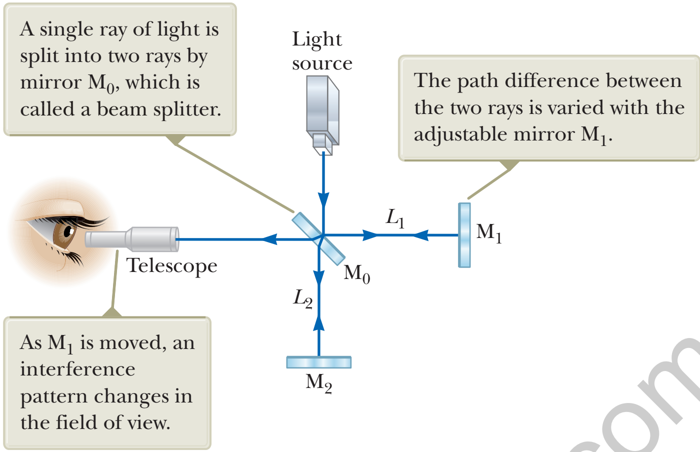


Figure 37.13 Diagram of the Michelson interferometer.

target pattern (corresponding to destructive interference) and M_1 is then moved a distance $\lambda/4$ toward M_0, the path difference changes by $\lambda/2$. What was a dark circle at the center now becomes a bright circle. As M_1 is moved an additional distance $\lambda/4$ toward M_0, the bright circle becomes a dark circle again. Therefore, the fringe pattern shifts by one-half fringe each time M_1 is moved a distance $\lambda/4$. The wavelength of light is then measured by counting the number of fringe shifts for a given displacement of M_1. If the wavelength is accurately known, mirror displacements can be measured to within a fraction of the wavelength.

We will see an important historical use of the Michelson interferometer in our discussion of relativity in Chapter 39. Modern uses include the following two applications, Fourier transform infrared spectroscopy and the laser interferometer gravitational-wave observatory.

Fourier Transform Infrared Spectroscopy

Spectroscopy is the study of the wavelength distribution of radiation from a sample that can be used to identify the characteristics of atoms or molecules in the sample. Infrared spectroscopy is particularly important to organic chemists when analyzing organic molecules. Traditional spectroscopy involves the use of an optical element, such as a prism (Section 35.5) or a diffraction grating (Section 38.4), which spreads out various wavelengths in a complex optical signal from the sample into different angles. In this way, the various wavelengths of radiation and their intensities in the signal can be determined. These types of devices are limited in their resolution and effectiveness because they must be scanned through the various angular deviations of the radiation.

The technique of *Fourier transform infrared* (FTIR) *spectroscopy* is used to create a higher-resolution spectrum in a time interval of 1 second that may have required 30 minutes with a standard spectrometer. In this technique, the radiation from a sample enters a Michelson interferometer. The movable mirror is swept through the zero-path-difference condition, and the intensity of radiation at the viewing position is recorded. The result is a complex set of data relating light intensity as a function of mirror position, called an *interferogram.* Because there is a relationship between mirror position and light intensity for a given wavelength, the interferogram contains information about all wavelengths in the signal.

In Section 18.8, we discussed Fourier analysis of a waveform. The waveform is a function that contains information about all the individual frequency components that make up the waveform.[3] Equation 18.13 shows how the waveform is generated from the individual frequency components. Similarly, the interferogram can be

[3]In acoustics, it is common to talk about the components of a complex signal in terms of frequency. In optics, it is more common to identify the components by wavelength.

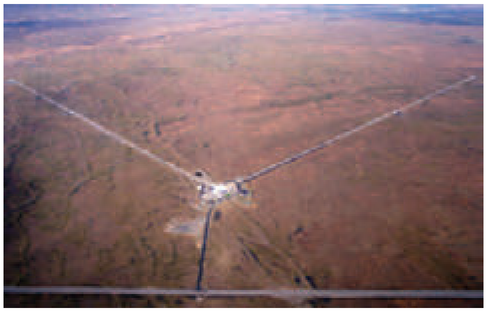


Figure 37.14 The Laser Interferometer Gravitational-Wave Observatory (LIGO) near Richland, Washington. Notice the two perpendicular arms of the Michelson interferometer.

analyzed by computer, in a process called a *Fourier transform,* to provide all the wavelength components. This information is the same as that generated by traditional spectroscopy, but the resolution of FTIR spectroscopy is much higher.

Laser Interferometer Gravitational-Wave Observatory

Einstein's general theory of relativity (Section 39.9) predicts the existence of *gravitational waves.* These waves propagate from the site of any gravitational disturbance, which could be periodic and predictable, such as the rotation of a double star around a center of mass, or unpredictable, such as the supernova explosion of a massive star.

In Einstein's theory, gravitation is equivalent to a distortion of space. Therefore, a gravitational disturbance causes an additional distortion that propagates through space in a manner similar to mechanical or electromagnetic waves. When gravitational waves from a disturbance pass by the Earth, they create a distortion of the local space. The laser interferometer gravitational-wave observatory (LIGO) apparatus is designed to detect this distortion. The apparatus employs a Michelson interferometer that uses laser beams with an effective path length of several kilometers. At the end of an arm of the interferometer, a mirror is mounted on a massive pendulum. When a gravitational wave passes by, the pendulum and the attached mirror move and the interference pattern due to the laser beams from the two arms changes.

Two sites for interferometers have been developed in the United States—in Richland, Washington, and in Livingston, Louisiana—to allow coincidence studies of gravitational waves. Figure 37.14 shows the Washington site. The two arms of the Michelson interferometer are evident in the photograph. Six data runs have been performed as of 2010. These runs have been coordinated with other gravitational wave detectors, such as GEO in Hannover, Germany, TAMA in Mitaka, Japan, and VIRGO in Cascina, Italy. So far, gravitational waves have not yet been detected, but the data runs have provided critical information for modifications and design features for the next generation of detectors. The original detectors are currently being dismantled, in preparation for the installation of Advanced LIGO, an upgrade that should increase the sensitivity of the observatory by a factor of 10. The target date for the beginning of scientific operation of Advanced LIGO is 2014.

Summary

Concepts and Principles

Interference in light waves occurs whenever two or more waves overlap at a given point. An interference pattern is observed if (1) the sources are coherent and (2) the sources have identical wavelengths.

The **intensity** at a point in a double-slit interference pattern is

$$I = I_{\max} \cos^2\left(\frac{\pi d \sin\theta}{\lambda}\right) \qquad (37.14)$$

where $I_{\max}$ is the maximum intensity on the screen and the expression represents the time average.

continued

A wave traveling from a medium of index of refraction toward a medium of index of refraction undergoes a 180° phase change upon reflection when and undergoes no phase change when

The condition for constructive interference in a film of thickness and index of refraction surrounded by air is

$$nt \quad - \quad 0, 1, 2, \tag{37.17}$$

where is the wavelength of the light in free space.

Similarly, the condition for destructive interference in a thin film surrounded by air is

$$0, 1, 2, \tag{37.18}$$

Analysis Models for Problem Solving

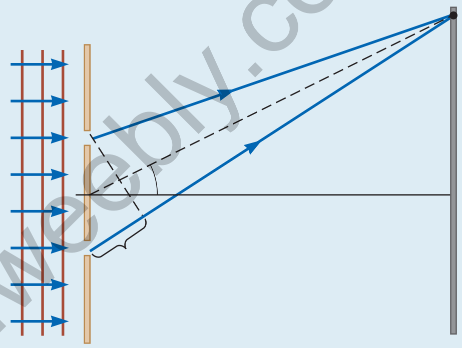
Waves in Interference. Young's double-slit experiment serves as a prototype for interference phenomena involving electromagnetic radiation. In this experiment, two slits separated by a distance are illuminated by a single-wavelength light source. The condition for bright fringes **(constructive interference)**

$$\sin_{\text{bright}} \quad 0, \quad 1, \quad 2, \tag{37.2}$$

The condition for dark fringes **(destructive interference)**

$$\sin_{\text{dark}} \quad - \quad 0, \quad 1, \quad 2, \tag{37.3}$$

The number is called the **order number** of the fringe.



Objective Questions

While using a Michelson interferometer (shown in Fig. 37.13), you see a dark circle at the center of the interference pattern. **(i)** As you gradually move the light source toward the central mirror M , through a distance /2, what do you see? (a) There is no change in the pattern. (b) The dark circle changes into a bright circle. (c) The dark circle changes into a bright circle and then back into a dark circle. (d) The dark circle changes into a bright circle, then into a dark circle, and then into a bright circle. **(ii)** As you gradually move the moving mirror toward the central mirror M , through a distance /2, what do you see? Choose from the same possibilities.

2. Four trials of Young's double-slit experiment are conducted. (a) In the first trial, blue light passes through two fine slits 400 m apart and forms an interference pattern on a screen 4 m away. (b) In a second trial, red light passes through the same slits and falls on the same screen. (c) A third trial is performed with red light and the same screen, but with slits 800 apart. (d) A final trial is performed with red light, slits 800 m apart, and a screen 8 m away. **(i)** Rank the trials (a) through (d) from the largest to the smallest value of the angle between the central maximum and the first-order side maximum. In your ranking, note any cases of equality. **(ii)** Rank the same trials according to the distance between the central maximum and the first-order side maximum on the screen.

3. Suppose Young's double-slit experiment is performed in air using red light and then the apparatus is immersed in water. What happens to the interference pattern on the screen? (a) It disappears. (b) The bright and dark fringes stay in the same locations, but the contrast is reduced. (c) The bright fringes are closer together. (d) The bright fringes are farther apart. (e) No change happens in the interference pattern.

4. Green light has a wavelength of 500 nm in air. **(i)** Assume green light is reflected from a mirror with angle of incidence 0°. The incident and reflected waves together constitute a standing wave with what distance from one node to the next node? (a) 1 000 nm (b) 500 nm (c) 250 nm (d) 125 nm (e) 62.5 nm **(ii).** The green light is sent into a Michelson interferometer that is adjusted to produce a central bright circle. How far must the interferometer's moving mirror be shifted to change the center of the pattern into a dark circle? Choose from the same possibilities as in part (i). **(iii).** The green light is reflected perpendicularly from a thin film of a plastic with an index of refraction 2.00. The film appears bright in the reflected light. How much additional thickness would make the film appear dark?

5. A thin layer of oil (1.25) is floating on water (1.33). What is the minimum nonzero thickness of the oil in the region that strongly reflects green light (530 nm)? (a) 500 nm (b) 313 nm (c) 404 nm (d) 212 nm (e) 285 nm

6. A monochromatic beam of light of wavelength 500 nm illuminates a double slit having a slit separation of 2.00 m. What is the angle of the second-order bright fringe? (a) 0.050 0 rad (b) 0.025 0 rad (c) 0.100 rad (d) 0.250 rad (e) 0.010 0 rad

7. According to Table 35.1, the index of refraction of flint glass is 1.66 and the index of refraction of crown glass is 1.52. **(i)** A film formed by one drop of sassafras oil, on a horizontal surface of a flint glass block, is viewed by reflected light. The film appears brightest at its outer margin, where it is thinnest. A film of the same oil on crown glass appears dark at its outer margin. What can you say about the index of refraction of the oil? (a) It must be less than 1.52. (b) It must be between 1.52 and 1.66. (c) It must be greater than 1.66. (d) None of those statements is necessarily true. **(ii)** Could a very thin film of some other liquid appear bright by reflected light on both of the glass blocks? **(iii)** Could it appear dark on both? **(iv)** Could it appear dark on crown glass and bright on flint glass? Experiments described by Thomas Young suggested this question.

8. Suppose you perform Young's double-slit experiment with the slit separation slightly smaller than the wavelength of the light. As a screen, you use a large half-cylinder with its axis along the midline between the slits. What interference pattern will you see on the interior surface of the cylinder? (a) bright and dark fringes so closely spaced as to be indistinguishable (b) one central bright fringe and two dark fringes only (c) a completely bright screen with no dark fringes (d) one central dark fringe and two bright fringes only (e) a completely dark screen with no bright fringes

A plane monochromatic light wave is incident on a double slit as illustrated in Figure 37.1. **(i)** As the viewing screen is moved away from the double slit, what happens to the separation between the interference fringes on the screen? (a) It increases. (b) It decreases. (c) It remains the same. (d) It may increase or decrease, depending on the wavelength of the light. (e) More information is required. **(ii)** As the slit separation increases, what happens to the separation between the interference fringes on the screen? Select from the same choices.

10. A film of oil on a puddle in a parking lot shows a variety of bright colors in swirled patches. What can you say about the thickness of the oil film? (a) It is much less than the wavelength of visible light. (b) It is on the same order of magnitude as the wavelength of visible light. (c) It is much greater than the wavelength of visible light. (d) It might have any relationship to the wavelength of visible light.

Conceptual Questions

Why is the lens on a good-quality camera coated with a thin film?

2. A soap film is held vertically in air and is viewed in reflected light as in Figure CQ37.2. Explain why the film appears to be dark at the top.

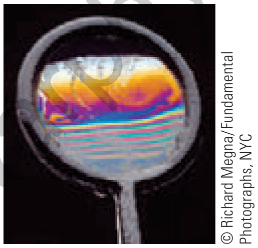


Figure CQ37.2 Conceptual Question 2 and Problem 70.

3. Explain why two flashlights held close together do not produce an interference pattern on a distant screen.

4. A lens with outer radius of curvature and index of refraction rests on a flat glass plate. The combination is illuminated with white light from above and observed from above. (a) Is there a dark spot or a light spot at the center of the lens? (b) What does it mean if the observed rings are noncircular?

5. Consider a dark fringe in a double-slit interference pattern at which almost no light energy is arriving. Light from both slits is arriving at the location of the dark fringe, but the waves cancel. Where does the energy at the positions of dark fringes go?

6. (a) In Young's double-slit experiment, why do we use monochromatic light? (b) If white light is used, how would the pattern change?

7. What is the necessary condition on the path length difference between two waves that interfere (a) constructively and (b) destructively?

8. In a laboratory accident, you spill two liquids onto different parts of a water surface. Neither of the liquids mixes with the water. Both liquids form thin films on the water surface. As the films spread and become very thin, you notice that one film becomes brighter and the other darker in reflected light. Why?

9. A theatrical smoke machine fills the space between the barrier and the viewing screen in the Young's double-slit experiment shown in Figure CQ37.9. Would the smoke show evidence of interference within this space? Explain your answer.

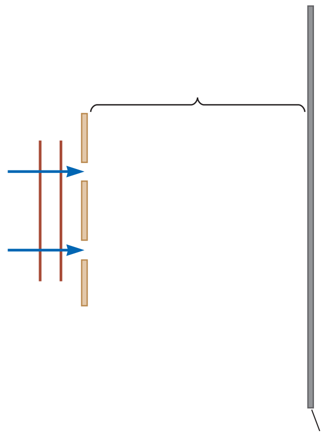


Figure CQ37.9

Problems

1. straightforward; **2.** intermediate; **3.** challenging

1. full solution available in the *Student Solutions Manual/Study Guide*

AMT Analysis Model tutorial available in Enhanced WebAssign

GP Guided Problem

M Master It tutorial available in Enhanced WebAssign

W Watch It video solution available in Enhanced WebAssign

Section 37.1 Young's Double-Slit Experiment

Section 37.2 Analysis Model: Waves in Interference

Problems 3, 5, 8, 10, and 13 in Chapter 18 can be assigned with this section.

1. Two slits are separated by 0.320 mm. A beam of 500-nm light strikes the slits, producing an interference pattern. Determine the number of maxima observed in the angular range $-30.0° < \theta < 30.0°$.

2. Light of wavelength 530 nm illuminates a pair of slits separated by 0.300 mm. If a screen is placed 2.00 m from the slits, determine the distance between the first and second dark fringes.

3. A laser beam is incident on two slits with a separation of 0.200 mm, and a screen is placed 5.00 m from the slits. An interference pattern appears on the screen. If the angle from the center fringe to the first bright fringe to the side is 0.181°, what is the wavelength of the laser light?

4. W A Young's interference experiment is performed with blue-green argon laser light. The separation between the slits is 0.500 mm, and the screen is located 3.30 m from the slits. The first bright fringe is located 3.40 mm from the center of the interference pattern. What is the wavelength of the argon laser light?

5. W Young's double-slit experiment is performed with 589-nm light and a distance of 2.00 m between the slits and the screen. The tenth interference minimum is observed 7.26 mm from the central maximum. Determine the spacing of the slits.

6. *Why is the following situation impossible?* Two narrow slits are separated by 8.00 mm in a piece of metal. A beam of microwaves strikes the metal perpendicularly, passes through the two slits, and then proceeds toward a wall some distance away. You know that the wavelength of the radiation is 1.00 cm ±5%, but you wish to measure it more precisely. Moving a microwave detector along the wall to study the interference pattern, you measure the position of the $m = 1$ bright fringe, which leads to a successful measurement of the wavelength of the radiation.

7. Light of wavelength 620 nm falls on a double slit, and the first bright fringe of the interference pattern is seen at an angle of 15.0° with the horizontal. Find the separation between the slits.

8. In a Young's double-slit experiment, two parallel slits with a slit separation of 0.100 mm are illuminated by light of wavelength 589 nm, and the interference pattern is observed on a screen located 4.00 m from the slits. (a) What is the difference in path lengths from each of the slits to the location of the center of a third-order bright fringe on the screen? (b) What is the difference in path lengths from the two slits to the location of the center of the third dark fringe away from the center of the pattern?

9. AMT M A pair of narrow, parallel slits separated by 0.250 mm is illuminated by green light ($\lambda = 546.1$ nm). The interference pattern is observed on a screen 1.20 m away from the plane of the parallel slits. Calculate the distance (a) from the central maximum to the first bright region on either side of the central maximum and (b) between the first and second dark bands in the interference pattern.

10. Light with wavelength 442 nm passes through a double-slit system that has a slit separation $d = 0.400$ mm. Determine how far away a screen must be placed so that dark fringes appear directly opposite both slits, with only one bright fringe between them.

11. AMT M The two speakers of a boom box are 35.0 cm apart. A single oscillator makes the speakers vibrate in phase at a frequency of 2.00 kHz. At what angles, measured from the perpendicular bisector of the line joining the speakers, would a distant observer hear maximum sound intensity? Minimum sound intensity? (Take the speed of sound as 340 m/s.)

12. In a location where the speed of sound is 343 m/s, a 2 000-Hz sound wave impinges on two slits 30.0 cm apart. (a) At what angle is the first maximum of sound intensity located? (b) **What If?** If the sound wave is replaced by 3.00-cm microwaves, what slit separation gives the same angle for the first maximum of microwave intensity? (c) **What If?** If the slit separation is 1.00 μm, what frequency of light gives the same angle to the first maximum of light intensity?

13. AMT M Two radio antennas separated by $d = 300$ m as shown in Figure P37.13 simultaneously broadcast identical signals at the same wavelength. A car travels due north along a straight line at position $x = 1\,000$ m from the center point between the antennas, and its radio receives the signals. (a) If the car is at the position of the second maximum after that at point O when it has traveled a distance y = 400 m northward, what is the wavelength of the signals? (b) How much farther must the car travel from this position to encounter the next minimum in reception? *Note:* Do not use the small-angle approximation in this problem.

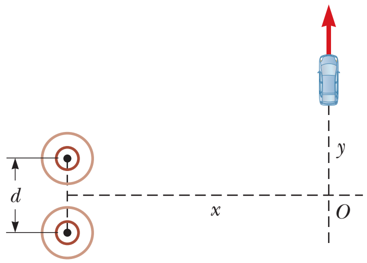


Figure P37.13

14. A riverside warehouse has several small doors facing **W** the river. Two of these doors are open as shown in Figure P37.14. The walls of the warehouse are lined with sound-absorbing material. Two people stand at a distance L = 150 m from the wall with the open doors. Person A stands along a line passing through the midpoint between the open doors, and person B stands a distance y = 20 m to his side. A boat on the river sounds its horn. To person A, the sound is loud and clear. To person B, the sound is barely audible. The principal wavelength of the sound waves is 3.00 m. Assuming person B is at the position of the first minimum, determine the distance d between the doors, center to center.

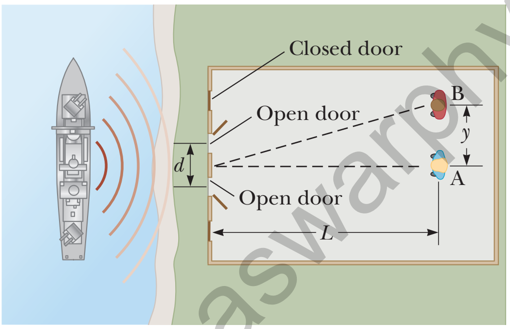


Figure P37.14

15. A student holds a laser that emits light of wavelength 632.8 nm. The laser beam passes though a pair of slits separated by 0.300 mm, in a glass plate attached to the front of the laser. The beam then falls perpendicularly on a screen, creating an interference pattern on it. The student begins to walk directly toward the screen at 3.00 m/s. The central maximum on the screen is stationary. Find the speed of the 50th-order maxima on the screen.

16. A student holds a laser that emits light of wavelength λ. The laser beam passes though a pair of slits separated by a distance d, in a glass plate attached to the front of the laser. The beam then falls perpendicularly on a screen, creating an interference pattern on it. The student begins to walk directly toward the screen at speed v. The central maximum on the screen is stationary. Find the speed of the mth-order maxima on the screen, where m can be very large.

17. Radio waves of wavelength 125 m from a galaxy reach a radio telescope by two separate paths as shown in Figure P37.17. One is a direct path to the receiver, which is situated on the edge of a tall cliff by the ocean, and the second is by reflection off the water. As the galaxy rises in the east over the water, the first minimum of destructive interference occurs when the galaxy is θ = 25.0° above the horizon. Find the height of the radio telescope dish above the water.

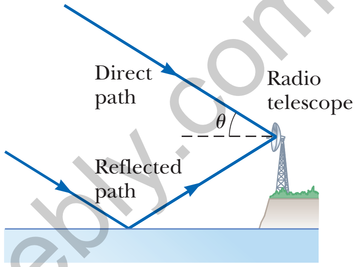


Figure P37.17 Problems 17 and 69.

18. **M** In Figure P37.18 (not to scale), let L = 1.20 m and d = 0.120 mm and assume the slit system is illuminated with monochromatic 500-nm light. Calculate the phase difference between the two wave fronts arriving at P when (a) θ = 0.500° and (b) y = 5.00 mm. (c) What is the value of θ for which the phase difference is 0.333 rad? (d) What is the value of θ for which the path difference is $\lambda/4$?

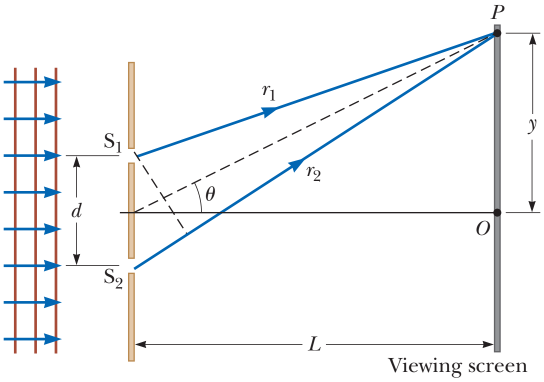


Figure P37.18 Problems 18 and 25.

19. Coherent light rays of wavelength λ strike a pair of slits separated by distance d at an angle θ_1 with respect to the normal to the plane containing the slits as shown in Figure P37.19. The rays leaving the slits make an

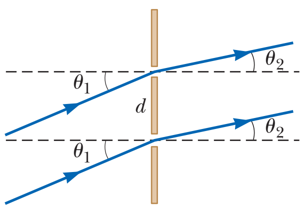


Figure P37.19

angle θ_2 with respect to the normal, and an interference maximum is formed by those rays on a screen that is a great distance from the slits. Show that the angle θ_2 is given by

$$\theta_2 = \sin^{-1}\left(\sin\theta_1 - \frac{m\lambda}{d}\right)$$

where m is an integer.

20. GP Monochromatic light of wavelength λ is incident on a pair of slits separated by 2.40×10^{-4} m and forms an interference pattern on a screen placed 1.80 m from the slits. The first-order bright fringe is at a position $y_{\text{bright}} = 4.52$ mm measured from the center of the central maximum. From this information, we wish to predict where the fringe for $n = 50$ would be located. (a) Assuming the fringes are laid out linearly along the screen, find the position of the $n = 50$ fringe by multiplying the position of the $n = 1$ fringe by 50.0. (b) Find the tangent of the angle the first-order bright fringe makes with respect to the line extending from the point midway between the slits to the center of the central maximum. (c) Using the result of part (b) and Equation 37.2, calculate the wavelength of the light. (d) Compute the angle for the 50th-order bright fringe from Equation 37.2. (e) Find the position of the 50th-order bright fringe on the screen from Equation 37.5. (f) Comment on the agreement between the answers to parts (a) and (e).

21. W In the double-slit arrangement of Figure P37.21, $d = 0.150$ mm, $L = 140$ cm, $\lambda = 643$ nm, and $y = 1.80$ cm. (a) What is the path difference δ for the rays from the two slits arriving at P? (b) Express this path difference in terms of λ. (c) Does P correspond to a maximum, a minimum, or an intermediate condition? Give evidence for your answer.

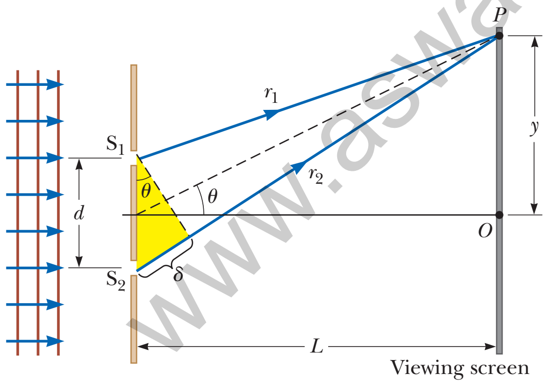


Figure P37.21

22. Young's double-slit experiment underlies the *instrument landing system* used to guide aircraft to safe landings at some airports when the visibility is poor. Although real systems are more complicated than the example described here, they operate on the same principles. A pilot is trying to align her plane with a runway as suggested in Figure P37.22. Two radio antennas (the black dots in the figure) are positioned adjacent to the runway, separated by $d = 40.0$ m. The antennas broadcast unmodulated coherent radio waves at 30.0 MHz. The red lines in Figure P37.22 represent paths along which maxima in the interference pattern of the radio waves exist. (a) Find the wavelength of the waves. The pilot "locks onto" the strong signal radiated along an interference maximum and steers the plane to keep the received signal strong. If she has found the central maximum, the plane will have precisely the correct heading to land when it reaches the runway as exhibited by plane A. (b) **What If?** Suppose the plane is flying along the first side maximum instead as is the case for plane B. How far to the side of the runway centerline will the plane be when it is 2.00 km from the antennas, measured along its direction of travel? (c) It is possible to tell the pilot that she is on the wrong maximum by sending out two signals from each antenna and equipping the aircraft with a two-channel receiver. The ratio of the two frequencies must not be the ratio of small integers (such as $\frac{3}{4}$). Explain how this two-frequency system would work and why it would not necessarily work if the frequencies were related by an integer ratio.

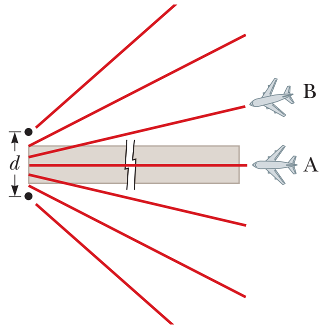


Figure P37.22

Section 37.3 Intensity Distribution of the Double-Slit Interference Pattern

23. Two slits are separated by 0.180 mm. An interference pattern is formed on a screen 80.0 cm away by 656.3-nm light. Calculate the fraction of the maximum intensity a distance $y = 0.600$ cm away from the central maximum.

24. Show that the two waves with wave functions given by $E_1 = 6.00 \sin(100\pi t)$ and $E_2 = 8.00 \sin(100\pi t + \pi/2)$ add to give a wave with the wave function $E_R \sin(100\pi t + \phi)$. Find the required values for E_R and ϕ.

25. M In Figure P37.18, let $L = 120$ cm and $d = 0.250$ cm. The slits are illuminated with coherent 600-nm light. Calculate the distance y from the central maximum for which the average intensity on the screen is 75.0% of the maximum.

26. Monochromatic coherent light of amplitude E_0 and angular frequency ω passes through three parallel slits, each separated by a distance d from its neighbor. (a) Show that the time-averaged intensity as a function of the angle θ is

$$I(\theta) = I_{\max}\left[1 + 2\cos\left(\frac{2\pi d \sin\theta}{\lambda}\right)\right]^2$$

(b) Explain how this expression describes both the primary and the secondary maxima. (c) Determine the ratio of the intensities of the primary and secondary maxima.

27. The intensity on the screen at a certain point in a double-slit interference pattern is 64.0% of the maximum value. (a) What minimum phase difference (in radians) between sources produces this result? (b) Express this phase difference as a path difference for 486.1-nm light.

28. Green light (λ = 546 nm) illuminates a pair of narrow, parallel slits separated by 0.250 mm. Make a graph of I/I_{max} as a function of θ for the interference pattern observed on a screen 1.20 m away from the plane of the parallel slits. Let θ range over the interval from $-0.3°$ to $+0.3°$.

29. Two narrow, parallel slits separated by 0.850 mm are illuminated by 600-nm light, and the viewing screen is 2.80 m away from the slits. (a) What is the phase difference between the two interfering waves on a screen at a point 2.50 mm from the central bright fringe? (b) What is the ratio of the intensity at this point to the intensity at the center of a bright fringe?

Section 37.4 Change of Phase Due to Reflection

Section 37.5 Interference in Thin Films

30. A soap bubble (n = 1.33) floating in air has the shape of a spherical shell with a wall thickness of 120 nm. (a) What is the wavelength of the visible light that is most strongly reflected? (b) Explain how a bubble of different thickness could also strongly reflect light of this same wavelength. (c) Find the two smallest film thicknesses larger than 120 nm that can produce strongly reflected light of the same wavelength.

31. A thin film of oil (n = 1.25) is located on smooth, wet pavement. When viewed perpendicular to the pavement, the film reflects most strongly red light at 640 nm and reflects no green light at 512 nm. How thick is the oil film?

32. A material having an index of refraction of 1.30 is used as an antireflective coating on a piece of glass (n = 1.50). What should the minimum thickness of this film be to minimize reflection of 500-nm light?

33. A possible means for making an airplane invisible to radar is to coat the plane with an antireflective polymer. If radar waves have a wavelength of 3.00 cm and the index of refraction of the polymer is n = 1.50, how thick would you make the coating?

34. A film of MgF_2 (n = 1.38) having thickness 1.00×10^{-5} cm is used to coat a camera lens. (a) What are the three longest wavelengths that are intensified in the reflected light? (b) Are any of these wavelengths in the visible spectrum?

35. A beam of 580-nm light passes through two closely spaced glass plates at close to normal incidence as shown in Figure P37.35. For what minimum nonzero value of the plate separation d is the transmitted light bright?

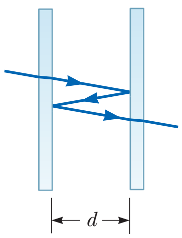


Figure P37.35

36. An oil film (n = 1.45) floating on water is illuminated by white light at normal incidence. The film is 280 nm thick. Find (a) the wavelength and color of the light in the visible spectrum most strongly reflected and (b) the wavelength and color of the light in the spectrum most strongly transmitted. Explain your reasoning.

37. An air wedge is formed between two glass plates separated at one edge by a very fine wire of circular cross section as shown in Figure P37.37. When the wedge is illuminated from above by 600-nm light and viewed from above, 30 dark fringes are observed. Calculate the diameter d of the wire.

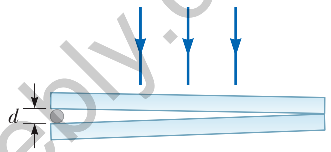


Figure P37.37 Problems 37, 41, 49, and 59.

38. Astronomers observe the chromosphere of the Sun with a filter that passes the red hydrogen spectral line of wavelength 656.3 nm, called the H_α line. The filter consists of a transparent dielectric of thickness d held between two partially aluminized glass plates. The filter is held at a constant temperature. (a) Find the minimum value of d that produces maximum transmission of perpendicular H_α light if the dielectric has an index of refraction of 1.378. (b) **What If?** If the temperature of the filter increases above the normal value, increasing its thickness, what happens to the transmitted wavelength? (c) The dielectric will also pass what near-visible wavelength? One of the glass plates is colored red to absorb this light.

39. When a liquid is introduced into the air space between the lens and the plate in a Newton's-rings apparatus, the diameter of the tenth ring changes from 1.50 to 1.31 cm. Find the index of refraction of the liquid.

40. A lens made of glass (n_g = 1.52) is coated with a thin film of MgF_2 (n_s = 1.38) of thickness t. Visible light is incident normally on the coated lens as in Figure P37.40. (a) For what minimum value of t will the

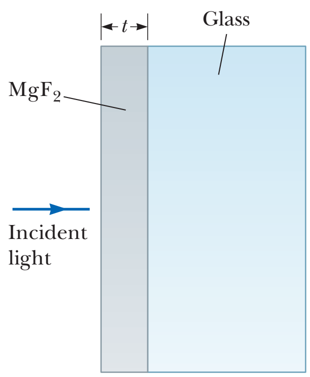


Figure P37.40

reflected light of wavelength 540 nm (in air) be missing? (b) Are there other values of t that will minimize the reflected light at this wavelength? Explain.

41. Two glass plates 10.0 cm long are in contact at one end and separated at the other end by a thread with a diameter $d = 0.050\,0$ mm (Fig. P37.37). Light containing the two wavelengths 400 nm and 600 nm is incident perpendicularly and viewed by reflection. At what distance from the contact point is the next dark fringe?

Section 37.6 The Michelson Interferometer

42. Mirror M_1 in Figure 37.13 is moved through a displacement ΔL. During this displacement, 250 fringe reversals (formation of successive dark or bright bands) are counted. The light being used has a wavelength of 632.8 nm. Calculate the displacement ΔL.

43. The Michelson interferometer can be used to measure the index of refraction of a gas by placing an evacuated transparent tube in the light path along one arm of the device. Fringe shifts occur as the gas is slowly added to the tube. Assume 600-nm light is used, the tube is 5.00 cm long, and 160 bright fringes pass on the screen as the pressure of the gas in the tube increases to atmospheric pressure. What is the index of refraction of the gas? *Hint:* The fringe shifts occur because the wavelength of the light changes inside the gas-filled tube.

44. One leg of a Michelson interferometer contains an evacuated cylinder of length L, having glass plates on each end. A gas is slowly leaked into the cylinder until a pressure of 1 atm is reached. If N bright fringes pass on the screen during this process when light of wavelength λ is used, what is the index of refraction of the gas? *Hint:* The fringe shifts occur because the wavelength of the light changes inside the gas-filled tube.

Additional Problems

45. Radio transmitter A operating at 60.0 MHz is 10.0 m from another similar transmitter B that is 180° out of phase with A. How far must an observer move from A toward B along the line connecting the two transmitters to reach the nearest point where the two beams are in phase?

46. A room is 6.0 m long and 3.0 m wide. At the front of the room, along one of the 3.0-m-wide walls, two loudspeakers are set 1.0 m apart, with the center point between them coinciding with the center point of the wall. The speakers emit a sound wave of a single frequency and a maximum in sound intensity is heard at the center of the back wall, 6.0 m from the speakers. What is the highest possible frequency of the sound from the speakers if no other maxima are heard anywhere along the back wall?

47. In an experiment similar to that of Example 37.1, green light with wavelength 560 nm, sent through a pair of slits 30.0 μm apart, produces bright fringes 2.24 cm apart on a screen 1.20 m away. If the apparatus is now submerged in a tank containing a sugar solution with index of refraction 1.38, calculate the fringe separation for this same arrangement.

48. In the What If? section of Example 37.2, it was claimed that overlapping fringes in a two-slit interference pattern for two different wavelengths obey the following relationship even for large values of the angle θ:

$$\frac{m'}{m} = \frac{\lambda}{\lambda'}$$

(a) Prove this assertion. (b) Using the data in Example 37.2, find the nonzero value of y on the screen at which the fringes from the two wavelengths first coincide.

49. An investigator finds a fiber at a crime scene that he wishes to use as evidence against a suspect. He gives the fiber to a technician to test the properties of the fiber. To measure the diameter d of the fiber, the technician places it between two flat glass plates at their ends as in Figure P37.37. When the plates, of length 14.0 cm, are illuminated from above with light of wavelength 650 nm, she observes interference bands separated by 0.580 mm. What is the diameter of the fiber?

50. Raise your hand and hold it flat. Think of the space between your index finger and your middle finger as one slit and think of the space between middle finger and ring finger as a second slit. (a) Consider the interference resulting from sending coherent visible light perpendicularly through this pair of openings. Compute an order-of-magnitude estimate for the angle between adjacent zones of constructive interference. (b) To make the angles in the interference pattern easy to measure with a plastic protractor, you should use an electromagnetic wave with frequency of what order of magnitude? (c) How is this wave classified on the electromagnetic spectrum?

51. Two coherent waves, coming from sources at different locations, move along the x axis. Their wave functions are

$$E_1 = 860 \sin\left[\frac{2\pi x_1}{650} - 924\pi t + \frac{\pi}{6}\right]$$

and

$$E_2 = 860 \sin\left[\frac{2\pi x_2}{650} - 924\pi t + \frac{\pi}{8}\right]$$

where E_1 and E_2 are in volts per meter, x_1 and x_2 are in nanometers, and t is in picoseconds. When the two waves are superposed, determine the relationship between x_1 and x_2 that produces constructive interference.

52. In a Young's interference experiment, the two slits are separated by 0.150 mm and the incident light includes two wavelengths: $\lambda_1 = 540$ nm (green) and $\lambda_2 = 450$ nm (blue). The overlapping interference patterns are observed on a screen 1.40 m from the slits. Calculate the minimum distance from the center of the screen to a point where a bright fringe of the green light coincides with a bright fringe of the blue light.

53. In a Young's double-slit experiment using light of wavelength λ, a thin piece of Plexiglas having index of refraction n covers one of the slits. If the center point on the screen is a dark spot instead of a bright spot, what is the minimum thickness of the Plexiglas?

54. AMT **Review.** A flat piece of glass is held stationary and horizontal above the highly polished, flat top end of a 10.0-cm-long vertical metal rod that has its lower end rigidly fixed. The thin film of air between the rod and glass is observed to be bright by reflected light when it is illuminated by light of wavelength 500 nm. As the temperature is slowly increased by 25.0°C, the film changes from bright to dark and back to bright 200 times. What is the coefficient of linear expansion of the metal?

55. A certain grade of crude oil has an index of refraction of 1.25. A ship accidentally spills 1.00 m^3 of this oil into the ocean, and the oil spreads into a thin, uniform slick. If the film produces a first-order maximum of light of wavelength 500 nm normally incident on it, how much surface area of the ocean does the oil slick cover? Assume the index of refraction of the ocean water is 1.34.

56. The waves from a radio station can reach a home receiver by two paths. One is a straight-line path from transmitter to home, a distance of 30.0 km. The second is by reflection from the ionosphere (a layer of ionized air molecules high in the atmosphere). Assume this reflection takes place at a point midway between receiver and transmitter, the wavelength broadcast by the radio station is 350 m, and no phase change occurs on reflection. Find the minimum height of the ionospheric layer that could produce destructive interference between the direct and reflected beams.

57. Interference effects are produced at point P on a screen as a result of direct rays from a 500-nm source and reflected rays from the mirror as shown in Figure P37.57. Assume the source is 100 m to the left of the screen and 1.00 cm above the mirror. Find the distance y to the first dark band above the mirror.

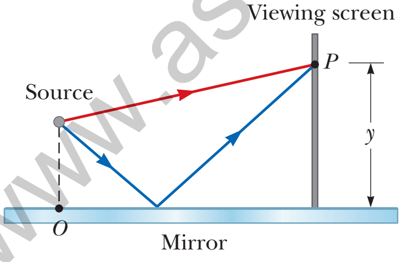


Figure P37.57

58. Measurements are made of the intensity distribution within the central bright fringe in a Young's interference pattern (see Fig. 37.6). At a particular value of y, it is found that $I/I_{max} = 0.810$ when 600-nm light is used. What wavelength of light should be used to reduce the relative intensity at the same location to 64.0% of the maximum intensity?

59. Many cells are transparent and colorless. Structures of great interest in biology and medicine can be practically invisible to ordinary microscopy. To indicate the size and shape of cell structures, an *interference microscope* reveals a difference in index of refraction as a shift in interference fringes. The idea is exemplified in the following problem. An air wedge is formed between two glass plates in contact along one edge and slightly separated at the opposite edge as in Figure P37.37. When the plates are illuminated with monochromatic light from above, the reflected light has 85 dark fringes. Calculate the number of dark fringes that appear if water ($n = 1.33$) replaces the air between the plates.

60. W Consider the double-slit arrangement shown in Figure P37.60, where the slit separation is d and the distance from the slit to the screen is L. A sheet of transparent plastic having an index of refraction n and thickness t is placed over the upper slit. As a result, the central maximum of the interference pattern moves upward a distance y'. Find y'.

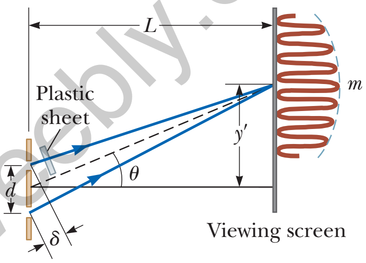


Figure P37.60

61. Figure P37.61 shows a radio-wave transmitter and a receiver separated by a distance $d = 50.0$ m and both a distance $h = 35.0$ m above the ground. The receiver can receive signals both directly from the transmitter and indirectly from signals that reflect from the ground. Assume the ground is level between the transmitter and receiver and a 180° phase shift occurs upon reflection. Determine the longest wavelengths that interfere (a) constructively and (b) destructively.

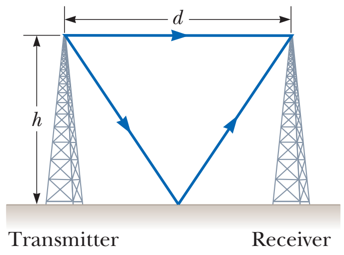


Figure P37.61 Problems 61 and 62.

62. Figure P37.61 shows a radio-wave transmitter and a receiver separated by a distance d and both a distance h above the ground. The receiver can receive signals both directly from the transmitter and indirectly from signals that reflect from the ground. Assume the ground is level between the transmitter and receiver and a 180° phase shift occurs upon reflection. Determine the longest wavelengths that interfere (a) constructively and (b) destructively.

63. In a Newton's-rings experiment, a plano-convex glass ($n = 1.52$) lens having radius $r = 5.00$ cm is placed on a flat plate as shown in Figure P37.63 (page 1158). When light of wavelength $\lambda = 650$ nm is incident normally, 55 bright rings are observed, with the last one precisely on the edge of the lens. (a) What is the radius R of curvature of the convex surface of the lens? (b) What is the focal length of the lens?

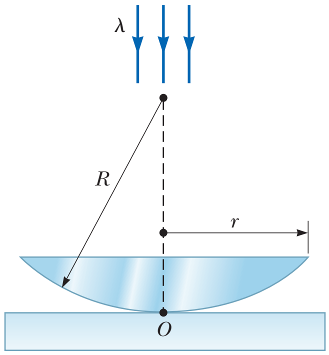


Figure P37.63

64. *Why is the following situation impossible?* A piece of transparent material having an index of refraction $n = 1.50$ is cut into the shape of a wedge as shown in Figure P37.64. Both the top and bottom surfaces of the wedge are in contact with air. Monochromatic light of wavelength $\lambda = 632.8$ nm is normally incident from above, and the wedge is viewed from above. Let $h = 1.00$ mm represent the height of the wedge and $\ell = 0.500$ m its length. A thin-film interference pattern appears in the wedge due to reflection from the top and bottom surfaces. You have been given the task of counting the number of bright fringes that appear in the entire length ℓ of the wedge. You find this task tedious, and your concentration is broken by a noisy distraction after accurately counting 5 000 bright fringes.

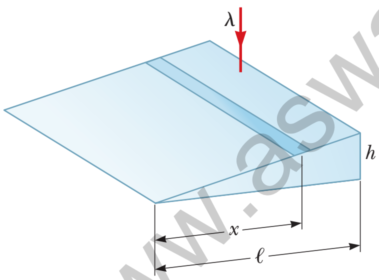


Figure P37.64

65. A plano-concave lens having index of refraction 1.50 is placed on a flat glass plate as shown in Figure P37.65. Its curved surface, with radius of curvature 8.00 m, is on the bottom. The lens is illuminated from above with yellow sodium light of wavelength 589 nm, and a series of concentric bright and dark rings is observed by reflection. The interference pattern has a dark spot at the center that is surrounded by 50 dark rings, the largest of which is at the outer edge of the lens. (a) What is the thickness of the air layer at the center of the interference pattern? (b) Calculate the radius of the outermost dark ring. (c) Find the focal length of the lens.

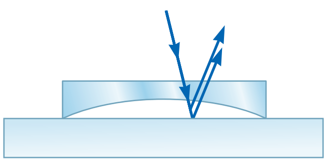


Figure P37.65

66. A plano-convex lens has index of refraction n. The curved side of the lens has radius of curvature R and rests on a flat glass surface of the same index of refraction, with a film of index n_{film} between them, as shown in Figure 37.66. The lens is illuminated from above by light of wavelength λ. Show that the dark Newton's rings have radii given approximately by

$$r \approx \sqrt{\frac{m\lambda R}{n_{\text{film}}}}$$

where $r << R$ and m is an integer.

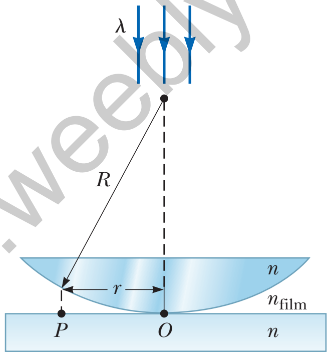


Figure P37.66

67. Interference fringes are produced using Lloyd's mirror and a source S of wavelength $\lambda = 606$ nm as shown in Figure P37.67. Fringes separated by $\Delta y = 1.20$ mm are formed on a screen a distance $L = 2.00$ m from the source. Find the vertical distance h of the source above the reflecting surface.

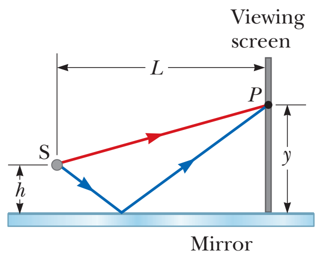


Figure P37.67

68. The quantity nt in Equations 37.17 and 37.18 is called the *optical path length* corresponding to the geometrical distance t and is analogous to the quantity δ in Equation 37.1, the path difference. The optical path length is proportional to n because a larger index of refraction shortens the wavelength, so more cycles of a wave fit into a particular geometrical distance. (a) Assume a mixture of corn syrup and water is prepared in a tank, with its index of refraction n increasing uniformly from 1.33 at $y = 20.0$ cm at the top to 1.90 at $y = 0$. Write the index of refraction $n(y)$ as a function of y.

(b) Compute the optical path length corresponding to the 20.0-cm height of the tank by calculating

$$\int_0^{20 \text{ cm}} n(y)\, dy$$

(c) Suppose a narrow beam of light is directed into the mixture at a nonzero angle with respect to the normal to the surface of the mixture. Qualitatively describe its path.

69. M Astronomers observe a 60.0-MHz radio source both directly and by reflection from the sea as shown in Figure P37.17. If the receiving dish is 20.0 m above sea level, what is the angle of the radio source above the horizon at first maximum?

70. Figure CQ37.2 shows an unbroken soap film in a circular frame. The film thickness increases from top to bottom, slowly at first and then rapidly. As a simpler model, consider a soap film (n = 1.33) contained within a rectangular wire frame. The frame is held vertically so that the film drains downward and forms a wedge with flat faces. The thickness of the film at the top is essentially zero. The film is viewed in reflected white light with near-normal incidence, and the first violet (λ = 420 nm) interference band is observed 3.00 cm from the top edge of the film. (a) Locate the first red (λ = 680 nm) interference band. (b) Determine the film thickness at the positions of the violet and red bands. (c) What is the wedge angle of the film?

Challenge Problems

71. Our discussion of the techniques for determining constructive and destructive interference by reflection from a thin film in air has been confined to rays striking the film at nearly normal incidence. **What If?** Assume a ray is incident at an angle of 30.0° (relative to the normal) on a film with index of refraction 1.38 surrounded by vacuum. Calculate the minimum thickness for constructive interference of sodium light with a wavelength of 590 nm.

72. The condition for constructive interference by reflection from a thin film in air as developed in Section 37.5 assumes nearly normal incidence. **What If?** Suppose the light is incident on the film at a nonzero angle θ_1 (relative to the normal). The index of refraction of the film is n, and the film is surrounded by vacuum. Find the condition for constructive interference that relates the thickness t of the film, the index of refraction n of the film, the wavelength λ of the light, and the angle of incidence θ_1.

73. Both sides of a uniform film that has index of refraction n and thickness d are in contact with air. For normal incidence of light, an intensity minimum is observed in the reflected light at λ_2 and an intensity maximum is observed at λ_1, where $\lambda_1 > \lambda_2$. (a) Assuming no intensity minima are observed between λ_1 and λ_2, find an expression for the integer m in Equations 37.17 and 37.18 in terms of the wavelengths λ_1 and λ_2. (b) Assuming $n = 1.40$, $\lambda_1 = 500$ nm, and $\lambda_2 = 370$ nm, determine the best estimate for the thickness of the film.

74. Slit 1 of a double slit is wider than slit 2 so that the light from slit 1 has an amplitude 3.00 times that of the light from slit 2. Show that Equation 37.13 is replaced by the equation $I = I_{max}(1 + 3\cos^2 \phi/2)$ for this situation.

75. Monochromatic light of wavelength 620 nm passes through a very narrow slit S and then strikes a screen in which are two parallel slits, S_1 and S_2, as shown in Figure P37.75. Slit S_1 is directly in line with S and at a distance of L = 1.20 m away from S, whereas S_2 is displaced a distance d to one side. The light is detected at point P on a second screen, equidistant from S_1 and S_2. When either slit S_1 or S_2 is open, equal light intensities are measured at point P. When both slits are open, the intensity is three times larger. Find the minimum possible value for the slit separation d.

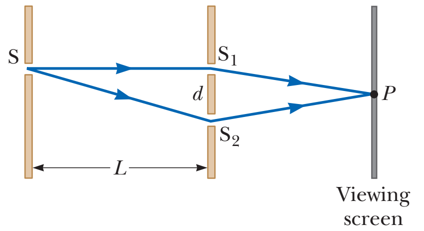


Figure P37.75

76. A plano-convex lens having a radius of curvature of r = 4.00 m is placed on a concave glass surface whose radius of curvature is R = 12.0 m as shown in Figure P37.76. Assuming 500-nm light is incident normal to the flat surface of the lens, determine the radius of the 100th bright ring.

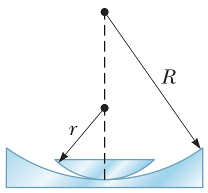


Figure P37.76

CHAPTER

38

Diffraction Patterns and Polarization

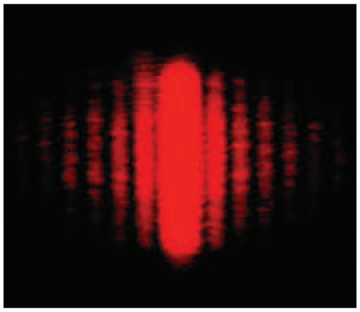
The Hubble Space Telescope does its viewing above the atmosphere and does not suffer from the atmospheric blurring, caused by air turbulence, that plagues ground-based telescopes. Despite this advantage, it does have limitations due to diffraction effects. In this chapter, we show how the wave nature of light limits the ability of any optical system to distinguish between closely spaced objects. *(NASA Hubble Space Telescope Collection)*

When plane light waves pass through a small aperture in an opaque barrier, the aperture acts as if it were a point source of light, with waves entering the shadow region behind the barrier. This phenomenon, known as diffraction, was first mentioned in Section 35.3, and can be described only with a wave model for light. In this chapter, we investigate the features of the *diffraction pattern* that occurs when the light from the aperture is allowed to fall upon a screen.

In Chapter 34, we learned that electromagnetic waves are transverse. That is, the electric and magnetic field vectors associated with electromagnetic waves are perpendicular to the direction of wave propagation. In this chapter, we show that under certain conditions these transverse waves with electric field vectors in all possible transverse directions can be *polarized* in various ways. In other words, only certain directions of the electric field vectors are present in the polarized wave.

38.1 Introduction to Diffraction Patterns

In Sections 35.3 and 37.1, we discussed that light of wavelength comparable to or larger than the width of a slit spreads out in all forward directions upon passing through the slit. This phenomenon is called *diffraction*. When light passes through a narrow slit, it spreads beyond the narrow path defined by the slit into regions that would be in shadow if light traveled in straight lines. Other waves, such as sound waves and water waves, also have this property of spreading when passing through apertures or by sharp edges.



Douglas C. Johnson/California State Polytechnic University, Pomona

Figure 38.1 The diffraction pattern that appears on a screen when light passes through a narrow vertical slit. The pattern consists of a broad central fringe and a series of less intense and narrower side fringes.

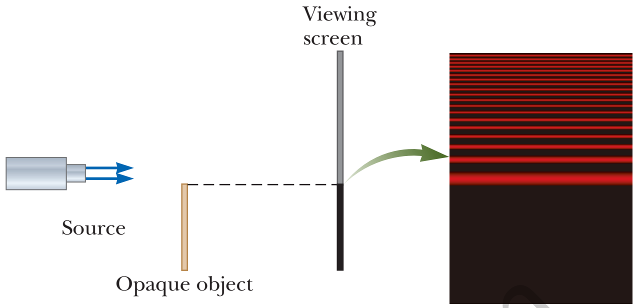
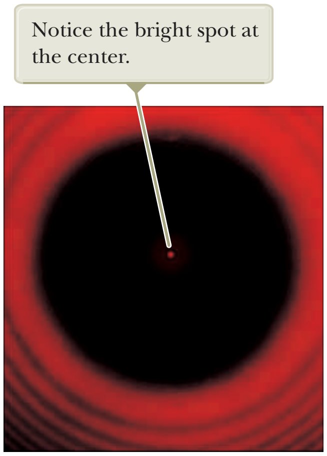


Figure 38.2 Light from a small source passes by the edge of an opaque object and continues on to a screen. A diffraction pattern consisting of bright and dark fringes appears on the screen in the region above the edge of the object.

You might expect that the light passing through a small opening would simply result in a broad region of light on a screen due to the spreading of the light as it passes through the opening. We find something more interesting, however. A **diffraction pattern** consisting of light and dark areas is observed, somewhat similar to the interference patterns discussed earlier. For example, when a narrow slit is placed between a distant light source (or a laser beam) and a screen, the light produces a diffraction pattern like that shown in Figure 38.1. The pattern consists of a broad, intense central band (called the **central maximum**) flanked by a series of narrower, less intense additional bands (called **side maxima** or **secondary maxima**) and a series of intervening dark bands (or **minima**). Figure 38.2 shows a diffraction pattern associated with light passing by the edge of an object. Again we see bright and dark fringes, which is reminiscent of an interference pattern.

Figure 38.3 shows a diffraction pattern associated with the shadow of a penny. A bright spot occurs at the center, and circular fringes extend outward from the shadow's edge. We can explain the central bright spot by using the wave theory of light, which predicts constructive interference at this point. From the viewpoint of ray optics (in which light is viewed as rays traveling in straight lines), we expect the center of the shadow to be dark because that part of the viewing screen is completely shielded by the penny.

Shortly before the central bright spot was first observed, one of the supporters of ray optics, Simeon Poisson, argued that if Augustin Fresnel's wave theory of light were valid, a central bright spot should be observed in the shadow of a circular object illuminated by a point source of light. To Poisson's astonishment, the spot was observed by Dominique Arago shortly thereafter. Therefore, Poisson's prediction reinforced the wave theory rather than disproving it.



P. M. Rinard, *Am. J. Phys.* 44: 70 1976

Figure 38.3 Diffraction pattern created by the illumination of a penny, with the penny positioned midway between the screen and light source.

38.2 Diffraction Patterns from Narrow Slits

Let's consider a common situation, that of light passing through a narrow opening modeled as a slit and projected onto a screen. To simplify our analysis, we assume the observing screen is far from the slit and the rays reaching the screen are approximately parallel. (This situation can also be achieved experimentally by using a converging lens to focus the parallel rays on a nearby screen.) In this model, the pattern on the screen is called a **Fraunhofer diffraction pattern.**[1]

Figure 38.4a (page 1162) shows light entering a single slit from the left and diffracting as it propagates toward a screen. Figure 38.4b shows the fringe structure of

[1]If the screen is brought close to the slit (and no lens is used), the pattern is a *Fresnel* diffraction pattern. The Fresnel pattern is more difficult to analyze, so we shall restrict our discussion to Fraunhofer diffraction.

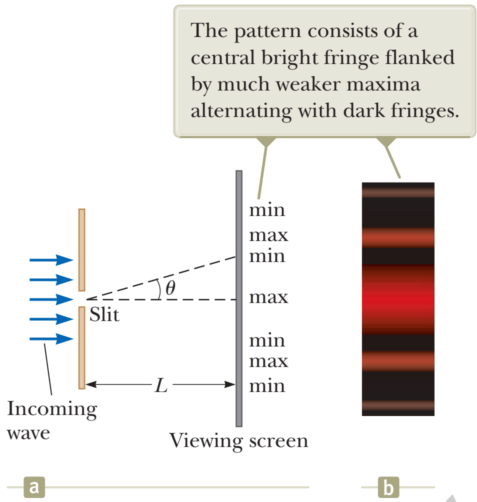


Figure 38.4 (a) Geometry for analyzing the Fraunhofer diffraction pattern of a single slit. (Drawing not to scale.) (b) Simulation of a single-slit Fraunhofer diffraction pattern.

Pitfall Prevention 38.1

Diffraction Versus Diffraction Pattern *Diffraction* refers to the general behavior of waves spreading out as they pass through a slit. We used diffraction in explaining the existence of an interference pattern in Chapter 37. A *diffraction pattern* is actually a misnomer, but is deeply entrenched in the language of physics. The diffraction pattern seen on a screen when a single slit is illuminated is actually another interference pattern. The interference is between parts of the incident light illuminating different regions of the slit.

a Fraunhofer diffraction pattern. A bright fringe is observed along the axis at $\theta = 0$, with alternating dark and bright fringes on each side of the central bright fringe.

Until now, we have assumed slits are point sources of light. In this section, we abandon that assumption and see how the finite width of slits is the basis for understanding Fraunhofer diffraction. We can explain some important features of this phenomenon by examining waves coming from various portions of the slit as shown in Figure 38.5. According to Huygens's principle, each portion of the slit acts as a source of light waves. Hence, light from one portion of the slit can interfere with light from another portion, and the resultant light intensity on a viewing screen depends on the direction θ. Based on this analysis, we recognize that a diffraction pattern is actually an interference pattern in which the different sources of light are different portions of the single slit! Therefore, the diffraction patterns we discuss in this chapter are applications of the waves in interference analysis model.

To analyze the diffraction pattern, let's divide the slit into two halves as shown in Figure 38.5. Keeping in mind that all the waves are in phase as they leave the slit, consider rays 1 and 3. As these two rays travel toward a viewing screen far to the right of the figure, ray 1 travels farther than ray 3 by an amount equal to the path difference $(a/2)\sin\theta$, where a is the width of the slit. Similarly, the path difference between rays 2 and 4 is also $(a/2)\sin\theta$, as is that between rays 3 and 5. If this path difference is exactly half a wavelength (corresponding to a phase difference of 180°), the pairs of waves cancel each other and destructive interference results. This cancellation occurs for any two rays that originate at points separated by half the slit width because the phase difference between two such points is 180°. Therefore, waves from the upper half of the slit interfere destructively with waves from the lower half when

$$\frac{a}{2}\sin\theta = \frac{\lambda}{2}$$

or, if we consider waves at angle θ both above the dashed line in Figure 38.5 and below,

$$\sin\theta = \pm\frac{\lambda}{a}$$

Dividing the slit into four equal parts and using similar reasoning, we find that the viewing screen is also dark when

$$\sin\theta = \pm 2\frac{\lambda}{a}$$

Likewise, dividing the slit into six equal parts shows that darkness occurs on the screen when

$$\sin\theta = \pm 3\frac{\lambda}{a}$$

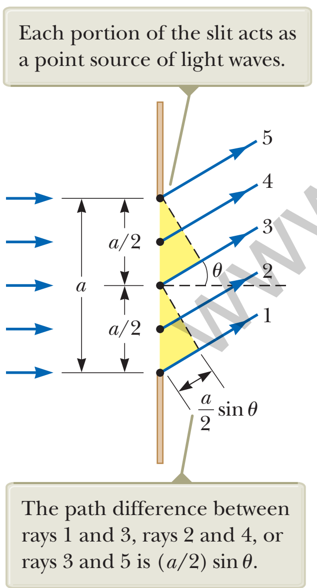


Figure 38.5 Paths of light rays that encounter a narrow slit of width a and diffract toward a screen in the direction described by angle θ (not to scale).

Therefore, the general condition for destructive interference is

$$\sin\theta_{\text{dark}} = m\frac{\lambda}{a} \quad m = \pm1, \pm2, \pm3, \ldots \tag{38.1}$$

◀ **Condition for destructive interference for a single slit**

This equation gives the values of θ_{dark} for which the diffraction pattern has zero light intensity, that is, when a dark fringe is formed. It tells us nothing, however, about the variation in light intensity along the screen. The general features of the intensity distribution are shown in Figure 38.4. A broad, central bright fringe is observed; this fringe is flanked by much weaker bright fringes alternating with dark fringes. The various dark fringes occur at the values of θ_{dark} that satisfy Equation 38.1. Each bright-fringe peak lies approximately halfway between its bordering dark-fringe minima. Notice that the central bright maximum is twice as wide as the secondary maxima. There is no central dark fringe, represented by the absence of $m = 0$ in Equation 38.1.

Pitfall Prevention 38.2

Similar Equation Warning! Equation 38.1 has exactly the same form as Equation 37.2, with d, the slit separation, used in Equation 37.2 and a, the slit width, used in Equation 38.1. Equation 37.2, however, describes the *bright* regions in a two-slit interference pattern, whereas Equation 38.1 describes the *dark* regions in a single-slit diffraction pattern.

Quick Quiz 38.1 Suppose the slit width in Figure 38.4 is made half as wide. Does the central bright fringe **(a)** become wider, **(b)** remain the same, or **(c)** become narrower?

Example 38.1 Where Are the Dark Fringes? AM

Light of wavelength 580 nm is incident on a slit having a width of 0.300 mm. The viewing screen is 2.00 m from the slit. Find the width of the central bright fringe.

SOLUTION

Conceptualize Based on the problem statement, we imagine a single-slit diffraction pattern similar to that in Figure 38.4.

Categorize We categorize this example as a straightforward application of our discussion of single-slit diffraction patterns, which comes from the *waves in interference* analysis model.

Analyze Evaluate Equation 38.1 for the two dark fringes that flank the central bright fringe, which correspond to $m = \pm1$:

$$\sin\theta_{\text{dark}} = \pm\frac{\lambda}{a}$$

Let y represent the vertical position along the viewing screen in Figure 38.4a, measured from the point on the screen directly behind the slit. Then, $\tan\theta_{\text{dark}} = y_1/L$, where the subscript 1 refers to the first dark fringe. Because θ_{dark} is very small, we can use the approximation $\sin\theta_{\text{dark}} \approx \tan\theta_{\text{dark}}$; therefore, $y_1 = L\sin\theta_{\text{dark}}$.

The width of the central bright fringe is twice the absolute value of y_1:

$$2|y_1| = 2|L\sin\theta_{\text{dark}}| = 2\left|\pm L\frac{\lambda}{a}\right| = 2L\frac{\lambda}{a} = 2(2.00\text{ m})\frac{580\times10^{-9}\text{ m}}{0.300\times10^{-3}\text{ m}}$$

$$= 7.73\times10^{-3}\text{ m} = \boxed{7.73\text{ mm}}$$

Finalize Notice that this value is much greater than the width of the slit. Let's explore below what happens if we change the slit width.

WHAT IF? What if the slit width is increased by an order of magnitude to 3.00 mm? What happens to the diffraction pattern?

Answer Based on Equation 38.1, we expect that the angles at which the dark bands appear will decrease as a increases. Therefore, the diffraction pattern narrows.

Repeat the calculation with the larger slit width:

$$2|y_1| = 2L\frac{\lambda}{a} = 2(2.00\text{ m})\frac{580\times10^{-9}\text{ m}}{3.00\times10^{-3}\text{ m}} = 7.73\times10^{-4}\text{ m} = \boxed{0.773\text{ mm}}$$

Notice that this result is *smaller* than the width of the slit. In general, for large values of a, the various maxima and minima are so closely spaced that only a large, central bright area resembling the geometric image of the slit is observed. This concept is very important in the performance of optical instruments such as telescopes.

Intensity of Single-Slit Diffraction Patterns

Analysis of the intensity variation in a diffraction pattern from a single slit of width a shows that the intensity is given by

Intensity of a single-slit Fraunhofer diffraction pattern

$$I = I_{\text{max}}\left[\frac{\sin(\pi a \sin\theta/\lambda)}{\pi a \sin\theta/\lambda}\right]^2 \quad \textbf{(38.2)}$$

where I_{max} is the intensity at $\theta = 0$ (the central maximum) and λ is the wavelength of light used to illuminate the slit. This expression shows that *minima* occur when

$$\frac{\pi a \sin\theta_{\text{dark}}}{\lambda} = m\pi$$

or

Condition for intensity minima for a single slit

$$\sin\theta_{\text{dark}} = m\frac{\lambda}{a} \qquad m = \pm1, \pm2, \pm3, \ldots$$

in agreement with Equation 38.1.

Figure 38.6a represents a plot of the intensity in the single-slit pattern as given by Equation 38.2, and Figure 38.6b is a simulation of a single-slit Fraunhofer diffraction pattern. Notice that most of the light intensity is concentrated in the central bright fringe.

Intensity of Two-Slit Diffraction Patterns

When more than one slit is present, we must consider not only diffraction patterns due to the individual slits but also the interference patterns due to the waves coming from different slits. Notice the curved dashed lines in Figure 37.7 in Chapter 37, which indicate a decrease in intensity of the interference maxima as θ increases. This decrease is due to a diffraction pattern. The interference patterns in that figure are located entirely within the central bright fringe of the diffraction pattern, so the only hint of the diffraction pattern we see is the falloff in intensity toward the outside of the pattern. To determine the effects of both two-slit interference and a single-slit diffraction pattern from each slit from a wider viewpoint than that in Figure 37.7, we combine Equations 37.14 and 38.2:

$$I = I_{\text{max}}\cos^2\left(\frac{\pi d \sin\theta}{\lambda}\right)\left[\frac{\sin(\pi a \sin\theta/\lambda)}{\pi a \sin\theta/\lambda}\right]^2 \quad \textbf{(38.3)}$$

Although this expression looks complicated, it merely represents the single-slit diffraction pattern (the factor in square brackets) acting as an "envelope" for a two-slit interference pattern (the cosine-squared factor) as shown in Figure 38.7. The broken

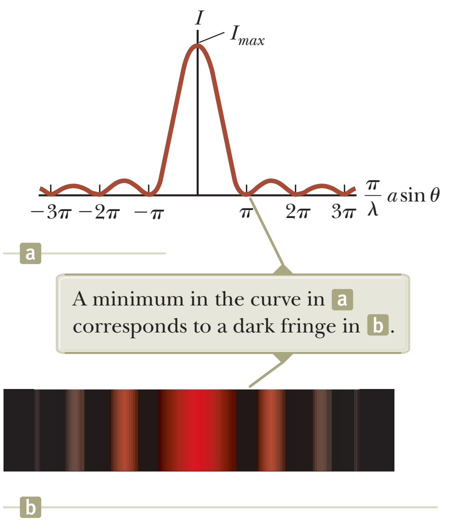


Figure 38.6 (a) A plot of light intensity I versus $(\pi/\lambda)a\sin\theta$ for the single-slit Fraunhofer diffraction pattern. (b) Simulation of a single-slit Fraunhofer diffraction pattern.

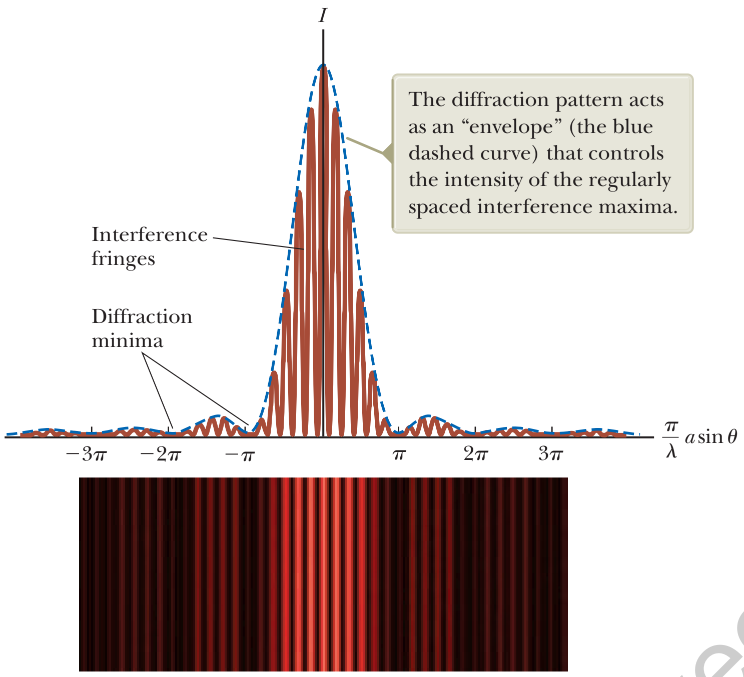


Figure 38.7 The combined effects of two-slit and single-slit interference. This pattern is produced when 650-nm light waves pass through two 3.0-μm slits that are 18 μm apart.

blue curve in Figure 38.7 represents the factor in square brackets in Equation 38.3. The cosine-squared factor by itself would give a series of peaks all with the same height as the highest peak of the red-brown curve in Figure 38.7. Because of the effect of the square-bracket factor, however, these peaks vary in height as shown.

Equation 37.2 indicates the conditions for interference maxima as $d \sin\theta = m\lambda$, where d is the distance between the two slits. Equation 38.1 specifies that the first diffraction minimum occurs when $a \sin\theta = \lambda$, where a is the slit width. Dividing Equation 37.2 by Equation 38.1 (with $m = 1$) allows us to determine which interference maximum coincides with the first diffraction minimum:

$$\frac{d \sin\theta}{a \sin\theta} = \frac{m\lambda}{\lambda}$$

$$\frac{d}{a} = m \qquad \textbf{(38.4)}$$

In Figure 38.7, $d/a = 18\ \mu\text{m}/3.0\ \mu\text{m} = 6$. Therefore, the sixth interference maximum (if we count the central maximum as $m = 0$) is aligned with the first diffraction minimum and is dark.

Quick Quiz 38.2 Consider the central peak in the diffraction envelope in Figure 38.7 and look closely at the horizontal scale. Suppose the wavelength of the light is changed to 450 nm. What happens to this central peak? **(a)** The width of the peak decreases, and the number of interference fringes it encloses decreases. **(b)** The width of the peak decreases, and the number of interference fringes it encloses increases. **(c)** The width of the peak decreases, and the number of interference fringes it encloses remains the same. **(d)** The width of the peak increases, and the number of interference fringes it encloses decreases. **(e)** The width of the peak increases, and the number of interference fringes it encloses increases. **(f)** The width of the peak increases, and the number of interference fringes it encloses remains the same. **(g)** The width of the peak remains the same, and the number of interference fringes it encloses decreases. **(h)** The width of the peak remains the same, and the number of interference fringes it encloses increases. **(i)** The width of the peak remains the same, and the number of interference fringes it encloses remains the same.

38.3 Resolution of Single-Slit and Circular Apertures

The ability of optical systems to distinguish between closely spaced objects is limited because of the wave nature of light. To understand this limitation, consider Figure 38.8, which shows two light sources far from a narrow slit of width a. The sources can be two noncoherent point sources S_1 and S_2; for example, they could be two distant stars. If no interference occurred between light passing through different parts of the slit, two distinct bright spots (or images) would be observed on the viewing screen. Because of such interference, however, each source is imaged as a bright central region flanked by weaker bright and dark fringes, a diffraction pattern. What is observed on the screen is the sum of two diffraction patterns: one from S_1 and the other from S_2.

If the two sources are far enough apart to keep their central maxima from overlapping as in Figure 38.8a, their images can be distinguished and are said to be *resolved*. If the sources are close together as in Figure 38.8b, however, the two central maxima overlap and the images are not resolved. To determine whether two images are resolved, the following condition is often used:

> When the central maximum of one image falls on the first minimum of another image, the images are said to be just resolved. This limiting condition of resolution is known as **Rayleigh's criterion.**

From Rayleigh's criterion, we can determine the minimum angular separation θ_{min} subtended by the sources at the slit in Figure 38.8 for which the images are just resolved. Equation 38.1 indicates that the first minimum in a single-slit diffraction pattern occurs at the angle for which

$$\sin\theta = \frac{\lambda}{a}$$

where a is the width of the slit. According to Rayleigh's criterion, this expression gives the smallest angular separation for which the two images are resolved. Because $\lambda << a$ in most situations, $\sin\theta$ is small and we can use the approximation $\sin\theta \approx \theta$. Therefore, the limiting angle of resolution for a slit of width a is

$$\theta_{\text{min}} = \frac{\lambda}{a} \qquad \textbf{(38.5)}$$

where θ_{min} is expressed in radians. Hence, the angle subtended by the two sources at the slit must be greater than λ/a if the images are to be resolved.

Many optical systems use circular apertures rather than slits. The diffraction pattern of a circular aperture as shown in the photographs of Figure 38.9 consists of

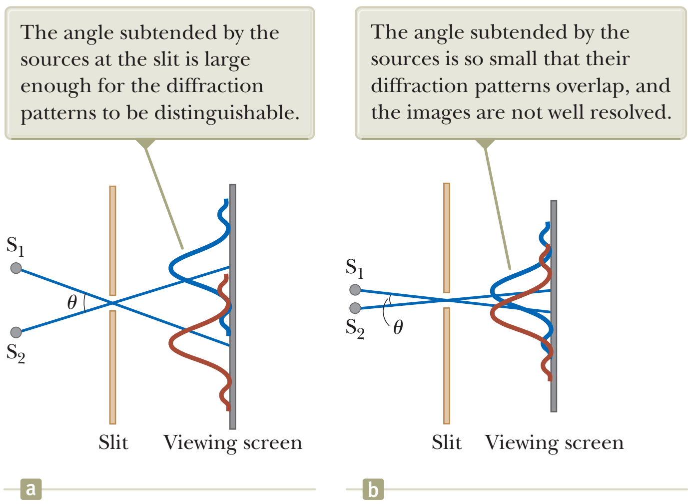


Figure 38.8 Two point sources far from a narrow slit each produce a diffraction pattern. (a) The sources are separated by a large angle. (b) The sources are separated by a small angle. (Notice that the angles are greatly exaggerated. The drawing is not to scale.)

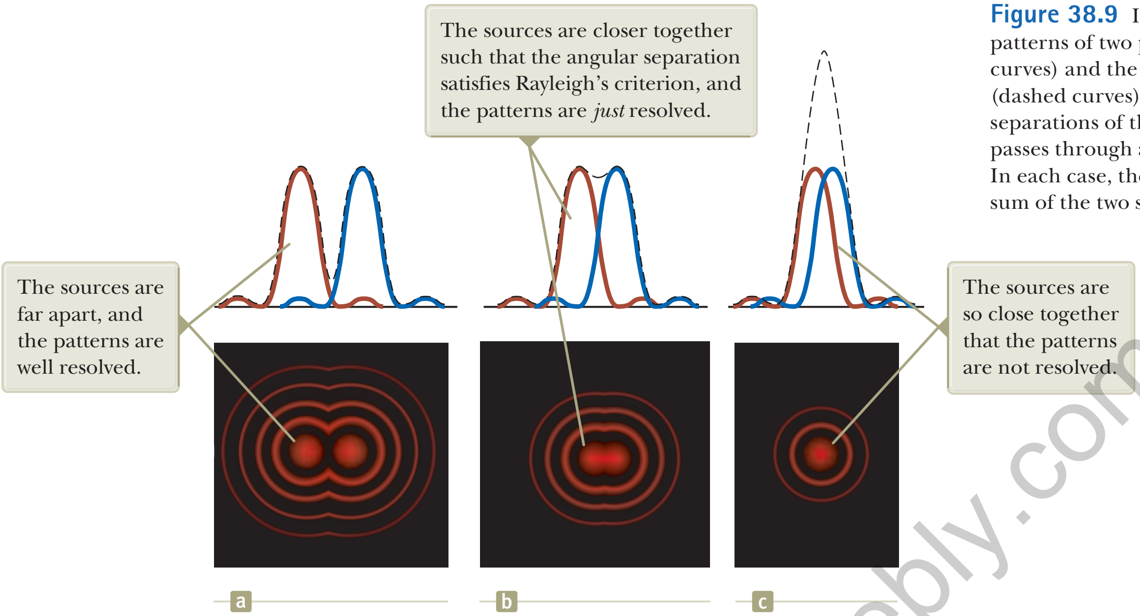


Figure 38.9 Individual diffraction patterns of two point sources (solid curves) and the resultant patterns (dashed curves) for various angular separations of the sources as the light passes through a circular aperture. In each case, the dashed curve is the sum of the two solid curves.

a central circular bright disk surrounded by progressively fainter bright and dark rings. Figure 38.9 shows diffraction patterns for three situations in which light from two point sources passes through a circular aperture. When the sources are far apart, their images are well resolved (Fig. 38.9a). When the angular separation of the sources satisfies Rayleigh's criterion, the images are just resolved (Fig. 38.9b). Finally, when the sources are close together, the images are said to be unresolved (Fig. 38.9c) and the pattern looks like that of a single source.

Analysis shows that the limiting angle of resolution of the circular aperture is

$$\theta_{\min} = 1.22\,\frac{\lambda}{D} \qquad \textbf{(38.6)}$$

◀ **Limiting angle of resolution for a circular aperture**

where D is the diameter of the aperture. This expression is similar to Equation 38.5 except for the factor 1.22, which arises from a mathematical analysis of diffraction from the circular aperture.

Quick Quiz 38.3 Cat's eyes have pupils that can be modeled as vertical slits. At night, would cats be more successful in resolving **(a)** headlights on a distant car or **(b)** vertically separated lights on the mast of a distant boat?

Quick Quiz 38.4 Suppose you are observing a binary star with a telescope and are having difficulty resolving the two stars. You decide to use a colored filter to maximize the resolution. (A filter of a given color transmits only that color of light.) What color filter should you choose? **(a)** blue **(b)** green **(c)** yellow **(d)** red

Example 38.2 Resolution of the Eye

Light of wavelength 500 nm, near the center of the visible spectrum, enters a human eye. Although pupil diameter varies from person to person, let's estimate a daytime diameter of 2 mm.

(A) Estimate the limiting angle of resolution for this eye, assuming its resolution is limited only by diffraction.

SOLUTION

Conceptualize Identify the pupil of the eye as the aperture through which the light travels. Light passing through this small aperture causes diffraction patterns to occur on the retina.

Categorize We determine the result using equations developed in this section, so we categorize this example as a substitution problem.

continued

38.2 continued

Use Equation 38.6, taking $\lambda = 500$ nm and $D = 2$ mm:

$$\theta_{\min} = 1.22\frac{\lambda}{D} = 1.22\left(\frac{5.00\times10^{-7}\text{ m}}{2\times10^{-3}\text{ m}}\right)$$

$$= \boxed{3\times10^{-4}\text{ rad}} \approx \boxed{1\text{ min of arc}}$$

(B) Determine the minimum separation distance d between two point sources that the eye can distinguish if the point sources are a distance $L = 25$ cm from the observer (Fig. 38.10).

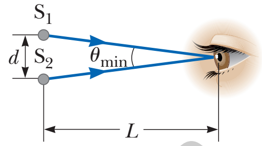


Figure 38.10 (Example 38.2) Two point sources separated by a distance d as observed by the eye.

SOLUTION

Noting that $\theta_{\min}$ is small, find d:

$$\sin\theta_{\min} \approx \theta_{\min} \approx \frac{d}{L} \rightarrow d = L\theta_{\min}$$

Substitute numerical values:

$$d = (25\text{ cm})(3\times10^{-4}\text{ rad}) = \boxed{8\times10^{-3}\text{ cm}}$$

This result is approximately equal to the thickness of a human hair.

Example 38.3 Resolution of a Telescope

Each of the two telescopes at the Keck Observatory on the dormant Mauna Kea volcano in Hawaii has an effective diameter of 10 m. What is its limiting angle of resolution for 600-nm light?

SOLUTION

Conceptualize Identify the aperture through which the light travels as the opening of the telescope. Light passing through this aperture causes diffraction patterns to occur in the final image.

Categorize We determine the result using equations developed in this section, so we categorize this example as a substitution problem.

Use Equation 38.6, taking $\lambda = 6.00\times10^{-7}$ m and $D = 10$ m:

$$\theta_{\min} = 1.22\frac{\lambda}{D} = 1.22\left(\frac{6.00\times10^{-7}\text{ m}}{10\text{ m}}\right)$$

$$= \boxed{7.3\times10^{-8}\text{ rad}} \approx \boxed{0.015\text{ s of arc}}$$

Any two stars that subtend an angle greater than or equal to this value are resolved (if atmospheric conditions are ideal).

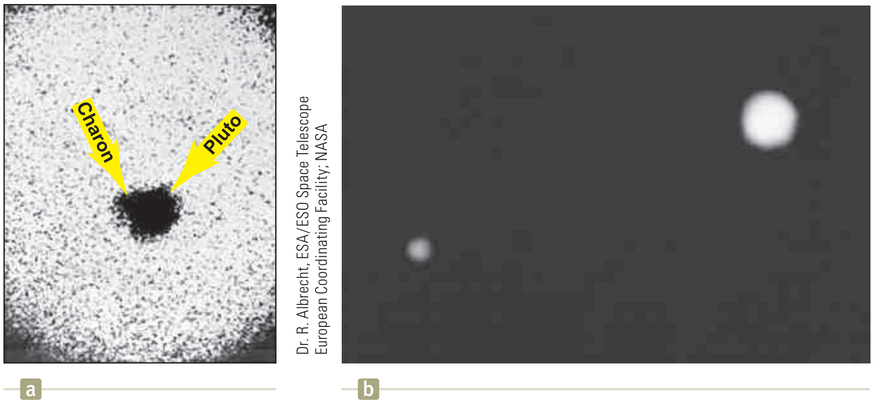
WHAT IF? What if we consider radio telescopes? They are much larger in diameter than optical telescopes, but do they have better angular resolutions than optical telescopes? For example, the radio telescope at Arecibo, Puerto Rico, has a diameter of 305 m and is designed to detect radio waves of 0.75-m wavelength. How does its resolution compare with that of one of the Keck telescopes?

Answer The increase in diameter might suggest that radio telescopes would have better resolution than a Keck telescope, but Equation 38.6 shows that $\theta_{\min}$ depends on *both* diameter and wavelength. Calculating the minimum angle of resolution for the radio telescope, we find

$$\theta_{\min} = 1.22\frac{\lambda}{D} = 1.22\left(\frac{0.75\text{ m}}{305\text{ m}}\right)$$

$$= 3.0\times10^{-3}\text{ rad} \approx 10\text{ min of arc}$$

This limiting angle of resolution is measured in *minutes* of arc rather than the *seconds* of arc for the optical telescope. Therefore, the change in wavelength more than compensates for the increase in diameter. The limiting angle of resolution for the Arecibo radio telescope is more than 40 000 times larger (that is, *worse*) than the Keck minimum.



Courtesy U.S. Naval Observatory/James W. Christy

Dr. R. Albrecht, ESA/ESO Space Telescope European Coordinating Facility; NASA

Figure 38.11 (a) The photograph on which Charon, the moon of Pluto, was discovered in 1978. From an Earth-based telescope, atmospheric blurring results in Charon appearing only as a subtle bump on the edge of Pluto. (b) A Hubble Space Telescope photo of Pluto and Charon, clearly resolving the two objects.

A telescope such as the one discussed in Example 38.3 can never reach its diffraction limit because the limiting angle of resolution is always set by atmospheric blurring at optical wavelengths. This seeing limit is usually about 1 s of arc and is never smaller than about 0.1 s of arc. The atmospheric blurring is caused by variations in index of refraction with temperature variations in the air. This blurring is one reason for the superiority of photographs from orbiting telescopes, which view celestial objects from a position above the atmosphere.

As an example of the effects of atmospheric blurring, consider telescopic images of Pluto and its moon, Charon. Figure 38.11a, an image taken in 1978, represents the discovery of Charon. In this photograph, taken from an Earth-based telescope, atmospheric turbulence causes the image of Charon to appear only as a bump on the edge of Pluto. In comparison, Figure 38.11b shows a photograph taken from the Hubble Space Telescope. Without the problems of atmospheric turbulence, Pluto and its moon are clearly resolved.

38.4 The Diffraction Grating

The **diffraction grating,** a useful device for analyzing light sources, consists of a large number of equally spaced parallel slits. A *transmission grating* can be made by cutting parallel grooves on a glass plate with a precision ruling machine. The spaces between the grooves are transparent to the light and hence act as separate slits. A *reflection grating* can be made by cutting parallel grooves on the surface of a reflective material. The reflection of light from the spaces between the grooves is specular, and the reflection from the grooves cut into the material is diffuse. Therefore, the spaces between the grooves act as parallel sources of reflected light like the slits in a transmission grating. Current technology can produce gratings that have very small slit spacings. For example, a typical grating ruled with 5 000 grooves/cm has a slit spacing $d = (1/5\,000)$ cm $= 2.00 \times 10^{-4}$ cm.

A section of a diffraction grating is illustrated in Figure 38.12 (page 1170). A plane wave is incident from the left, normal to the plane of the grating. The pattern observed on the screen far to the right of the grating is the result of the combined effects of interference and diffraction. Each slit produces diffraction, and the diffracted beams interfere with one another to produce the final pattern.

The waves from all slits are in phase as they leave the slits. For an arbitrary direction θ measured from the horizontal, however, the waves must travel different path lengths before reaching the screen. Notice in Figure 38.12 that the path difference δ between rays from any two adjacent slits is equal to $d \sin \theta$. If this path difference equals one wavelength or some integral multiple of a wavelength, waves from all slits are in phase at the screen and a bright fringe is observed. Therefore, the condition for *maxima* in the interference pattern at the angle θ_{bright} is

$$d \sin \theta_{\text{bright}} = m\lambda \quad m = 0, \pm 1, \pm 2, \pm 3, \ldots \tag{38.7}$$

◀ **Condition for interference maxima for a grating**

Pitfall Prevention 38.3

A Diffraction Grating Is an Interference Grating As with *diffraction pattern, diffraction grating* is a misnomer, but is deeply entrenched in the language of physics. The diffraction grating depends on diffraction in the same way as the double slit, spreading the light so that light from different slits can interfere. It would be more correct to call it an *interference grating,* but *diffraction grating* is the name in use.

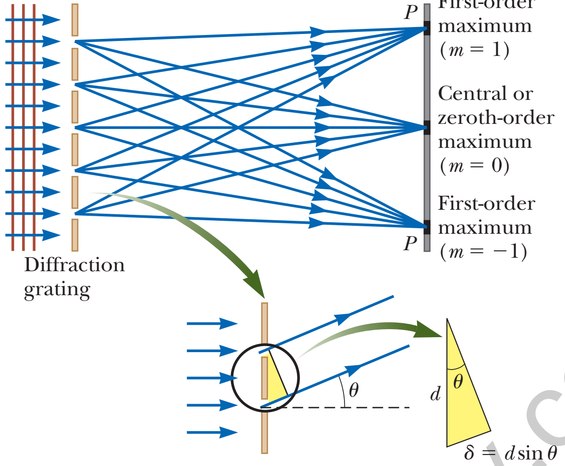


Figure 38.12 Side view of a diffraction grating. The slit separation is d, and the path difference between adjacent slits is $d \sin \theta$.

We can use this expression to calculate the wavelength if we know the grating spacing d and the angle θ_{bright}. If the incident radiation contains several wavelengths, the mth-order maximum for each wavelength occurs at a specific angle. All wavelengths are seen at $\theta = 0$, corresponding to $m = 0$, the zeroth-order maximum. The first-order maximum ($m = 1$) is observed at an angle that satisfies the relationship $\sin \theta_{\text{bright}} = \lambda/d$, the second-order maximum ($m = 2$) is observed at a larger angle θ_{bright}, and so on. For the small values of d typical in a diffraction grating, the angles θ_{bright} are large, as we see in Example 38.5.

The intensity distribution for a diffraction grating obtained with the use of a monochromatic source is shown in Figure 38.13. Notice the sharpness of the principal maxima and the broadness of the dark areas compared with the broad bright fringes characteristic of the two-slit interference pattern (see Fig. 37.6). You should also review Figure 37.7, which shows that the width of the intensity maxima decreases as the number of slits increases. Because the principal maxima are so sharp, they are much brighter than two-slit interference maxima.

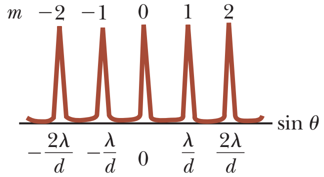


Figure 38.13 Intensity versus $\sin \theta$ for a diffraction grating. The zeroth-, first-, and second-order maxima are shown.

Quick Quiz 38.5 Ultraviolet light of wavelength 350 nm is incident on a diffraction grating with slit spacing d and forms an interference pattern on a screen a distance L away. The angular positions θ_{bright} of the interference maxima are large. The locations of the bright fringes are marked on the screen. Now red light of wavelength 700 nm is used with a diffraction grating to form another diffraction pattern on the screen. Will the bright fringes of this pattern be located at the marks on the screen if **(a)** the screen is moved to a distance $2L$ from the grating, **(b)** the screen is moved to a distance $L/2$ from the grating, **(c)** the grating is replaced with one of slit spacing $2d$, **(d)** the grating is replaced with one of slit spacing $d/2$, or **(e)** nothing is changed?

Conceptual Example 38.4 A Compact Disc Is a Diffraction Grating

Light reflected from the surface of a compact disc is multicolored as shown in Figure 38.14. The colors and their intensities depend on the orientation of the CD relative to the eye and relative to the light source. Explain how that works.

SOLUTION

The surface of a CD has a spiral grooved track (with adjacent grooves having a separation on

Carlos E. Santa Maria/Shutterstock.com

Figure 38.14 (Conceptual Example 38.4) A compact disc observed under white light. The colors observed in the reflected light and their intensities depend on the orientation of the CD relative to the eye and relative to the light source.

38.4 continued

the order of 1 μm). Therefore, the surface acts as a reflection grating. The light reflecting from the regions between these closely spaced grooves interferes constructively only in certain directions that depend on the wavelength and the direction of the incident light. Any section of the CD serves as a diffraction grating for white light, sending different colors in different directions. The different colors you see upon viewing one section change when the light source, the CD, or you change position. This change in position causes the angle of incidence or the angle of the diffracted light to be altered.

Example 38.5 The Orders of a Diffraction Grating

Monochromatic light from a helium–neon laser (λ = 632.8 nm) is incident normally on a diffraction grating containing 6 000 grooves per centimeter. Find the angles at which the first- and second-order maxima are observed.

SOLUTION

Conceptualize Study Figure 38.12 and imagine that the light coming from the left originates from the helium–neon laser. Let's evaluate the possible values of the angle θ for constructive interference.

Categorize We determine results using equations developed in this section, so we categorize this example as a substitution problem.

Calculate the slit separation as the inverse of the number of grooves per centimeter:

$$d = \frac{1}{6\,000}\text{ cm} = 1.667 \times 10^{-4}\text{ cm} = 1\,667\text{ nm}$$

Solve Equation 38.7 for sin θ and substitute numerical values for the first-order maximum (m = 1) to find θ_1:

$$\sin\theta_1 = \frac{(1)\lambda}{d} = \frac{632.8\text{ nm}}{1\,667\text{ nm}} = 0.379\,7$$

$$\theta_1 = \boxed{22.31°}$$

Repeat for the second-order maximum (m = 2):

$$\sin\theta_2 = \frac{(2)\lambda}{d} = \frac{2(632.8\text{ nm})}{1\,667\text{ nm}} = 0.759\,4$$

$$\theta_2 = \boxed{49.41°}$$

WHAT IF? What if you looked for the third-order maximum? Would you find it?

Answer For m = 3, we find sin θ_3 = 1.139. Because sin θ cannot exceed unity, this result does not represent a realistic solution. Hence, only zeroth-, first-, and second-order maxima can be observed for this situation.

Applications of Diffraction Gratings

A schematic drawing of a simple apparatus used to measure angles in a diffraction pattern is shown in Figure 38.15 (page 1172). This apparatus is a *diffraction grating spectrometer.* The light to be analyzed passes through a slit, and a collimated beam of light is incident on the grating. The diffracted light leaves the grating at angles that satisfy Equation 38.7, and a telescope is used to view the image of the slit. The wavelength can be determined by measuring the precise angles at which the images of the slit appear for the various orders.

The spectrometer is a useful tool in *atomic spectroscopy,* in which the light from an atom is analyzed to find the wavelength components. These wavelength components can be used to identify the atom. We shall investigate atomic spectra in Chapter 42 of the extended version of this text.

Another application of diffraction gratings is the *grating light valve* (GLV), which competes in some video display applications with the digital micromirror devices (DMDs) discussed in Section 35.4. A GLV is a silicon microchip fitted with an array

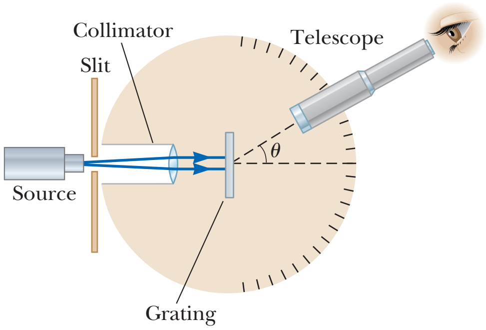


Figure 38.15 Diagram of a diffraction grating spectrometer. The collimated beam incident on the grating is spread into its various wavelength components with constructive interference for a particular wavelength occurring at the angles θ_{bright} that satisfy the equation $d \sin \theta_{\text{bright}} = m\lambda$, where $m = 0, \pm1, \pm2, \ldots$.

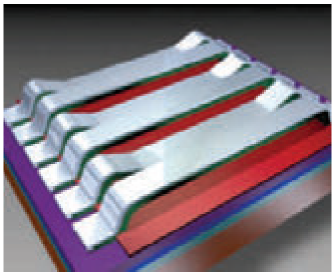


Figure 38.16 A small portion of a grating light valve. The alternating reflective ribbons at different levels act as a diffraction grating, offering very high-speed control of the direction of light toward a digital display device.

of parallel silicon nitride ribbons coated with a thin layer of aluminum (Fig. 38.16). Each ribbon is approximately 20 μm long and 5 μm wide and is separated from the silicon substrate by an air gap on the order of 100 nm. With no voltage applied, all ribbons are at the same level. In this situation, the array of ribbons acts as a flat surface, specularly reflecting incident light.

When a voltage is applied between a ribbon and the electrode on the silicon substrate, an electric force pulls the ribbon downward, closer to the substrate. Alternate ribbons can be pulled down, while those in between remain in an elevated configuration. As a result, the array of ribbons acts as a diffraction grating such that the constructive interference for a particular wavelength of light can be directed toward a screen or other optical display system. If one uses three such devices—one each for red, blue, and green light—full-color display is possible.

In addition to its use in video display, the GLV has found applications in laser optical navigation sensor technology, computer-to-plate commercial printing, and other types of imaging.

Another interesting application of diffraction gratings is **holography,** the production of three-dimensional images of objects. The physics of holography was developed by Dennis Gabor (1900–1979) in 1948 and resulted in the Nobel Prize in Physics for Gabor in 1971. The requirement of coherent light for holography delayed the realization of holographic images from Gabor's work until the development of lasers in the 1960s. Figure 38.17 shows a single hologram viewed from two different positions and the three-dimensional character of its image. Notice in particular the difference in the view through the magnifying glass in Figures 38.17a and 38.17b.

Figure 38.18 shows how a hologram is made. Light from the laser is split into two parts by a half-silvered mirror at *B*. One part of the beam reflects off the object to be photographed and strikes an ordinary photographic film. The other half of the beam is diverged by lens L_2, reflects from mirrors M_1 and M_2, and finally

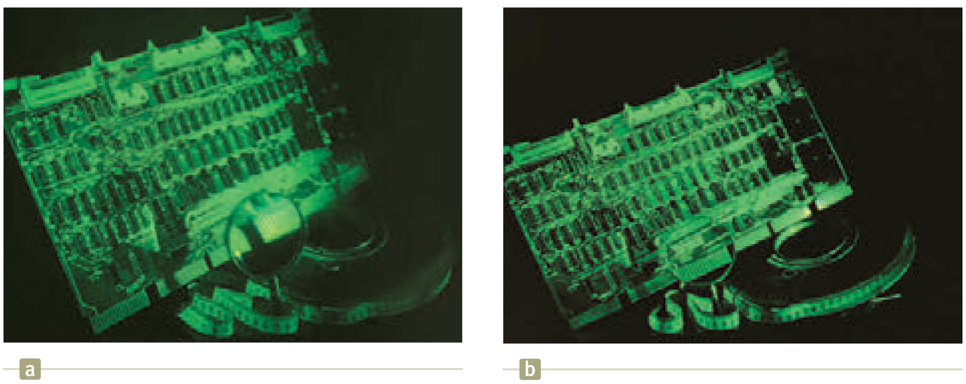


Figure 38.17 In this hologram, a circuit board is shown from two different views. Notice the difference in the appearance of the measuring tape and the view through the magnifying lens in (a) and (b).

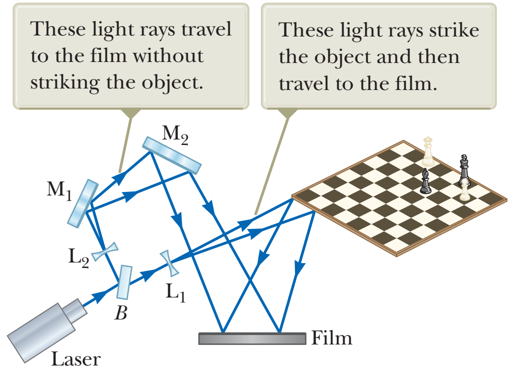


Figure 38.18 Experimental arrangement for producing a hologram.

strikes the film. The two beams overlap to form an extremely complicated interference pattern on the film. Such an interference pattern can be produced only if the phase relationship of the two waves is constant throughout the exposure of the film. This condition is met by illuminating the scene with light coming through a pinhole or with coherent laser radiation. The hologram records not only the intensity of the light scattered from the object (as in a conventional photograph), but also the phase difference between the reference beam and the beam scattered from the object. Because of this phase difference, an interference pattern is formed that produces an image in which all three-dimensional information available from the perspective of any point on the hologram is preserved.

In a normal photographic image, a lens is used to focus the image so that each point on the object corresponds to a single point on the photograph. Notice that there is no lens used in Figure 38.18 to focus the light onto the film. Therefore, light from each point on the object reaches *all* points on the film. As a result, each region of the photographic film on which the hologram is recorded contains information about all illuminated points on the object, which leads to a remarkable result: if a small section of the hologram is cut from the film, the complete image can be formed from the small piece! (The quality of the image is reduced, but the entire image is present.)

A hologram is best viewed by allowing coherent light to pass through the developed film as one looks back along the direction from which the beam comes. The interference pattern on the film acts as a diffraction grating. Figure 38.19 shows two rays of light striking and passing through the film. For each ray, the $m = 0$ and $m = \pm 1$ rays in the diffraction pattern are shown emerging from the right side of the film. The $m = +1$ rays converge to form a real image of the scene, which is not the image that is normally viewed. By extending the light rays corresponding to $m = -1$ behind the film, we see that there is a virtual image located there, with light coming from it in exactly the same way that light came from the actual object

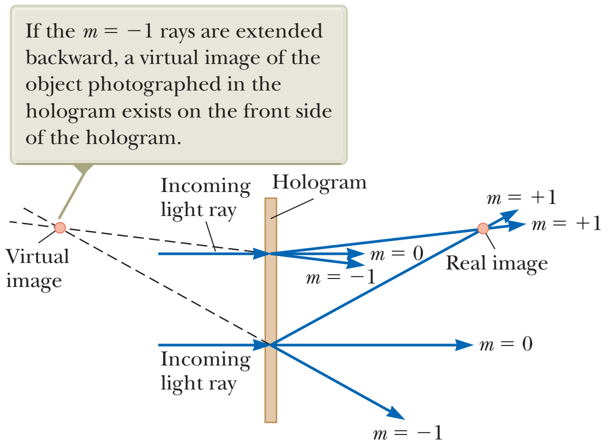


Figure 38.19 Two light rays strike a hologram at normal incidence. For each ray, outgoing rays corresponding to $m = 0$ and $m = \pm1$ are shown.

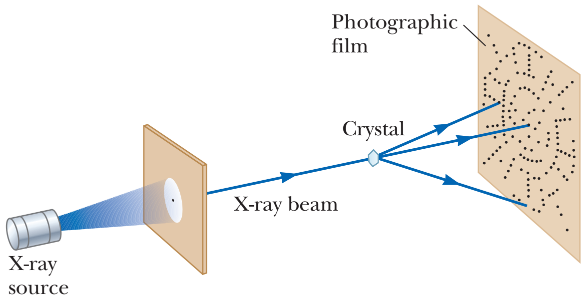


Figure 38.20 Schematic diagram of the technique used to observe the diffraction of x-rays by a crystal. The array of spots formed on the film is called a Laue pattern.

when the film was exposed. This image is what one sees when looking through the holographic film.

Holograms are finding a number of applications. You may have a hologram on your credit card. This special type of hologram is called a *rainbow hologram* and is designed to be viewed in reflected white light.

38.5 Diffraction of X-Rays by Crystals

In principle, the wavelength of any electromagnetic wave can be determined if a grating of the proper spacing (on the order of λ) is available. X-rays, discovered by Wilhelm Roentgen (1845–1923) in 1895, are electromagnetic waves of very short wavelength (on the order of 0.1 nm). It would be impossible to construct a grating having such a small spacing by the cutting process described at the beginning of Section 38.4. The atomic spacing in a solid is known to be about 0.1 nm, however. In 1913, Max von Laue (1879–1960) suggested that the regular array of atoms in a crystal could act as a three-dimensional diffraction grating for x-rays. Subsequent experiments confirmed this prediction. The diffraction patterns from crystals are complex because of the three-dimensional nature of the crystal structure. Nevertheless, x-ray diffraction has proved to be an invaluable technique for elucidating these structures and for understanding the structure of matter.

Figure 38.20 shows one experimental arrangement for observing x-ray diffraction from a crystal. A collimated beam of monochromatic x-rays is incident on a crystal. The diffracted beams are very intense in certain directions, corresponding to constructive interference from waves reflected from layers of atoms in the crystal. The diffracted beams, which can be detected by a photographic film, form an array of spots known as a *Laue pattern* as in Figure 38.21a. One can deduce the crystalline structure by analyzing the positions and intensities of the various spots in the pattern. Figure 38.21b shows a Laue pattern from a crystalline enzyme, using a wide range of wavelengths so that a swirling pattern results.

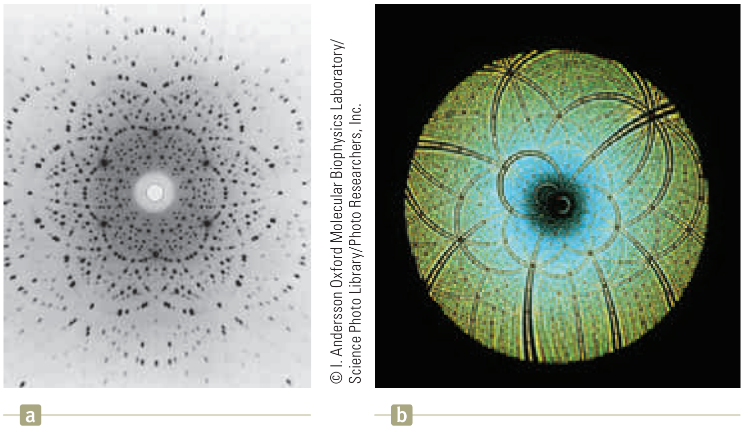


Figure 38.21 (a) A Laue pattern of a single crystal of the mineral beryl (beryllium aluminum silicate). Each dot represents a point of constructive interference. (b) A Laue pattern of the enzyme Rubisco, produced with a wide-band x-ray spectrum. This enzyme is present in plants and takes part in the process of photosynthesis. The Laue pattern is used to determine the crystal structure of Rubisco.

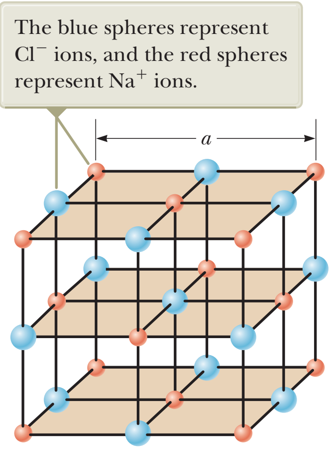


Figure 38.22 Crystalline structure of sodium chloride (NaCl). The length of the cube edge is $a = 0.562\ 737$ nm.

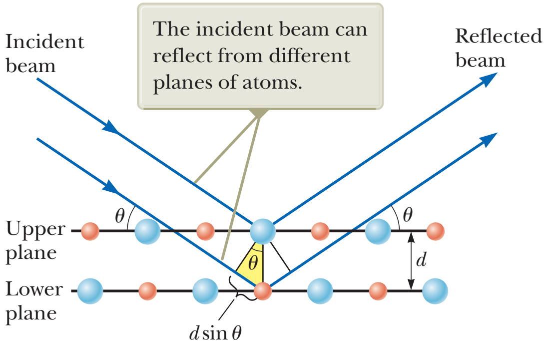


Figure 38.23 A two-dimensional description of the reflection of an x-ray beam from two parallel crystalline planes separated by a distance d. The beam reflected from the lower plane travels farther than the beam reflected from the upper plane by a distance $2d \sin\theta$.

The arrangement of atoms in a crystal of sodium chloride (NaCl) is shown in Figure 38.22. Each unit cell (the geometric solid that repeats throughout the crystal) is a cube having an edge length a. A careful examination of the NaCl structure shows that the ions lie in discrete planes (the shaded areas in Fig. 38.22). Now suppose an incident x-ray beam makes an angle θ with one of the planes as in Figure 38.23. The beam can be reflected from both the upper plane and the lower one, but the beam reflected from the lower plane travels farther than the beam reflected from the upper plane. The effective path difference is $2d \sin\theta$. The two beams reinforce each other (constructive interference) when this path difference equals some integer multiple of λ. The same is true for reflection from the entire family of parallel planes. Hence, the condition for *constructive* interference (maxima in the reflected beam) is

$$2d \sin\theta = m\lambda \quad m = 1, 2, 3, \ldots \tag{38.8}$$

◀ **Bragg's law**

This condition is known as **Bragg's law,** after W. L. Bragg (1890–1971), who first derived the relationship. If the wavelength and diffraction angle are measured, Equation 38.8 can be used to calculate the spacing between atomic planes.

Pitfall Prevention 38.4

Different Angles Notice in Figure 38.23 that the angle θ is measured from the reflecting surface rather than from the normal as in the case of the law of reflection in Chapter 35. With slits and diffraction gratings, we also measured the angle θ from the normal to the array of slits. Because of historical tradition, the angle is measured differently in Bragg diffraction, so interpret Equation 38.8 with care.

38.6 Polarization of Light Waves

In Chapter 34, we described the transverse nature of light and all other electromagnetic waves. Polarization, discussed in this section, is firm evidence of this transverse nature.

An ordinary beam of light consists of a large number of waves emitted by the atoms of the light source. Each atom produces a wave having some particular orientation of the electric field vector $\vec{\mathbf{E}}$, corresponding to the direction of atomic vibration. The *direction of polarization* of each individual wave is defined to be the direction in which the electric field is vibrating. In Figure 38.24, this direction happens to lie along the y axis. All individual electromagnetic waves traveling in the x direction have an $\vec{\mathbf{E}}$ vector parallel to the yz plane, but this vector could be at any possible angle with respect to the y axis. Because all directions of vibration from a wave source are possible, the resultant electromagnetic wave is a superposition of waves vibrating in many different directions. The result is an **unpolarized** light beam, represented in Figure 38.25a (page 1176). The direction of wave propagation in this figure is perpendicular to the page. The arrows show a few possible

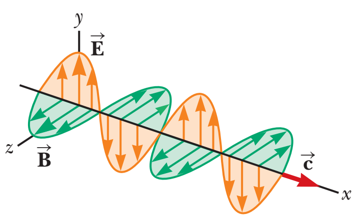


Figure 38.24 Schematic diagram of an electromagnetic wave propagating at velocity $\vec{\mathbf{c}}$ in the x direction. The electric field vibrates in the xy plane, and the magnetic field vibrates in the xz plane.

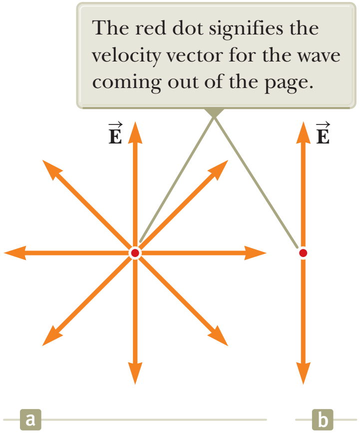


Figure 38.25 (a) A representation of an unpolarized light beam viewed along the direction of propagation. The transverse electric field can vibrate in any direction in the plane of the page with equal probability. (b) A linearly polarized light beam with the electric field vibrating in the vertical direction.

directions of the electric field vectors for the individual waves making up the resultant beam. At any given point and at some instant of time, all these individual electric field vectors add to give one resultant electric field vector.

As noted in Section 34.3, a wave is said to be **linearly polarized** if the resultant electric field $\vec{\mathbf{E}}$ vibrates in the same direction *at all times* at a particular point as shown in Figure 38.25b. (Sometimes, such a wave is described as *plane-polarized,* or simply *polarized.*) The plane formed by $\vec{\mathbf{E}}$ and the direction of propagation is called the *plane of polarization* of the wave. If the wave in Figure 38.24 represents the resultant of all individual waves, the plane of polarization is the *xy* plane.

A linearly polarized beam can be obtained from an unpolarized beam by removing all waves from the beam except those whose electric field vectors oscillate in a single plane. We now discuss four processes for producing polarized light from unpolarized light.

Polarization by Selective Absorption

The most common technique for producing polarized light is to use a material that transmits waves whose electric fields vibrate in a plane parallel to a certain direction and that absorbs waves whose electric fields vibrate in all other directions.

In 1938, E. H. Land (1909–1991) discovered a material, which he called *Polaroid,* that polarizes light through selective absorption. This material is fabricated in thin sheets of long-chain hydrocarbons. The sheets are stretched during manufacture so that the long-chain molecules align. After a sheet is dipped into a solution containing iodine, the molecules become good electrical conductors. Conduction takes place primarily along the hydrocarbon chains because electrons can move easily only along the chains. If light whose electric field vector is parallel to the chains is incident on the material, the electric field accelerates electrons along the chains and energy is absorbed from the radiation. Therefore, the light does not pass through the material. Light whose electric field vector is perpendicular to the chains passes through the material because electrons cannot move from one molecule to the next. As a result, when unpolarized light is incident on the material, the exiting light is polarized perpendicular to the molecular chains.

It is common to refer to the direction perpendicular to the molecular chains as the *transmission axis.* In an ideal polarizer, all light with $\vec{\mathbf{E}}$ parallel to the transmission axis is transmitted and all light with $\vec{\mathbf{E}}$ perpendicular to the transmission axis is absorbed.

Figure 38.26 represents an unpolarized light beam incident on a first polarizing sheet, called the *polarizer.* Because the transmission axis is oriented vertically in the figure, the light transmitted through this sheet is polarized vertically. A second polarizing sheet, called the *analyzer,* intercepts the beam. In Figure 38.26, the analyzer transmission axis is set at an angle θ to the polarizer axis. We call the electric field vector of the first transmitted beam $\vec{\mathbf{E}}_0$. The component of $\vec{\mathbf{E}}_0$ perpendicular to the analyzer axis is completely absorbed. The component of $\vec{\mathbf{E}}_0$ parallel to the

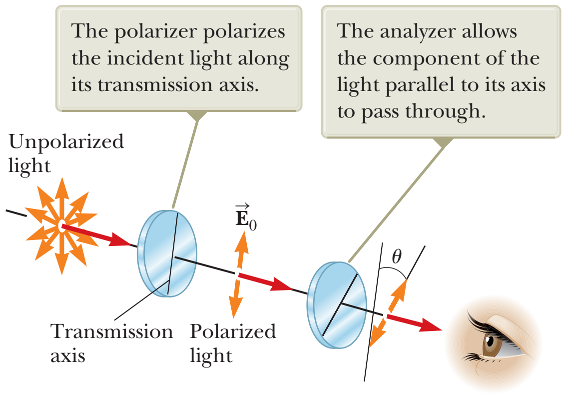


Figure 38.26 Two polarizing sheets whose transmission axes make an angle θ with each other. Only a fraction of the polarized light incident on the analyzer is transmitted through it.

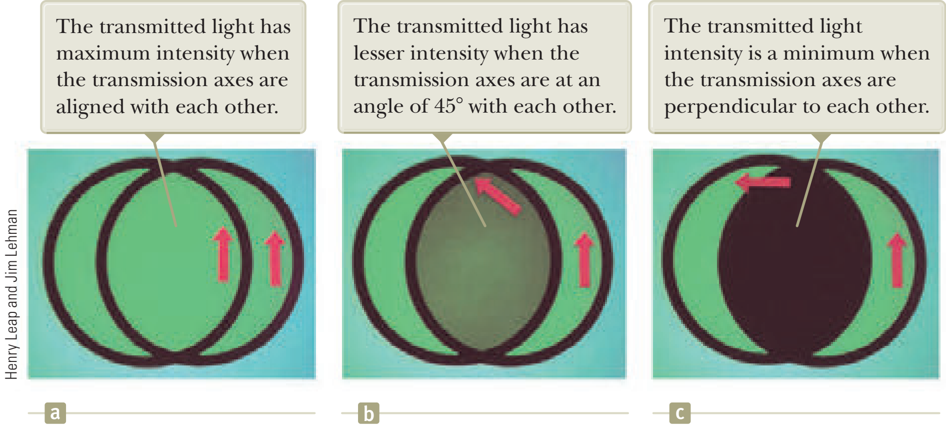


Figure 38.27 The intensity of light transmitted through two polarizers depends on the relative orientation of their transmission axes. The red arrows indicate the transmission axes of the polarizers.

analyzer axis, which is transmitted through the analyzer, is $E_0 \cos\theta$. Because the intensity of the transmitted beam varies as the square of its magnitude, we conclude that the intensity I of the (polarized) beam transmitted through the analyzer varies as

$$I = I_{\max}\cos^2\theta \tag{38.9}$$

◀ Malus's law

where $I_{\max}$ is the intensity of the polarized beam incident on the analyzer. This expression, known as **Malus's law,**[2] applies to any two polarizing materials whose transmission axes are at an angle θ to each other. This expression shows that the intensity of the transmitted beam is maximum when the transmission axes are parallel ($\theta = 0$ or $180°$) and is zero (complete absorption by the analyzer) when the transmission axes are perpendicular to each other. This variation in transmitted intensity through a pair of polarizing sheets is illustrated in Figure 38.27. Because the average value of $\cos^2\theta$ is $\frac{1}{2}$, the intensity of initially unpolarized light is reduced by a factor of one-half as the light passes through a single ideal polarizer.

Polarization by Reflection

When an unpolarized light beam is reflected from a surface, the polarization of the reflected light depends on the angle of incidence. If the angle of incidence is 0°, the reflected beam is unpolarized. For other angles of incidence, the reflected light is polarized to some extent, and for one particular angle of incidence, the reflected light is completely polarized. Let's now investigate reflection at that special angle.

Suppose an unpolarized light beam is incident on a surface as in Figure 38.28a (page 1178). Each individual electric field vector can be resolved into two components: one parallel to the surface (and perpendicular to the page in Fig. 38.28, represented by the dots) and the other (represented by the orange arrows) perpendicular both to the first component and to the direction of propagation. Therefore, the polarization of the entire beam can be described by two electric field components in these directions. It is found that the parallel component represented by the dots reflects more strongly than the other component represented by the arrows, resulting in a partially polarized reflected beam. Furthermore, the refracted beam is also partially polarized.

Now suppose the angle of incidence θ_1 is varied until the angle between the reflected and refracted beams is 90° as in Figure 38.28b. At this particular angle of incidence, the reflected beam is completely polarized (with its electric field vector parallel to the surface) and the refracted beam is still only partially polarized. The angle of incidence at which this polarization occurs is called the **polarizing angle** θ_p.

[2]Named after its discoverer, E. L. Malus (1775–1812). Malus discovered that reflected light was polarized by viewing it through a calcite ($CaCO_3$) crystal.

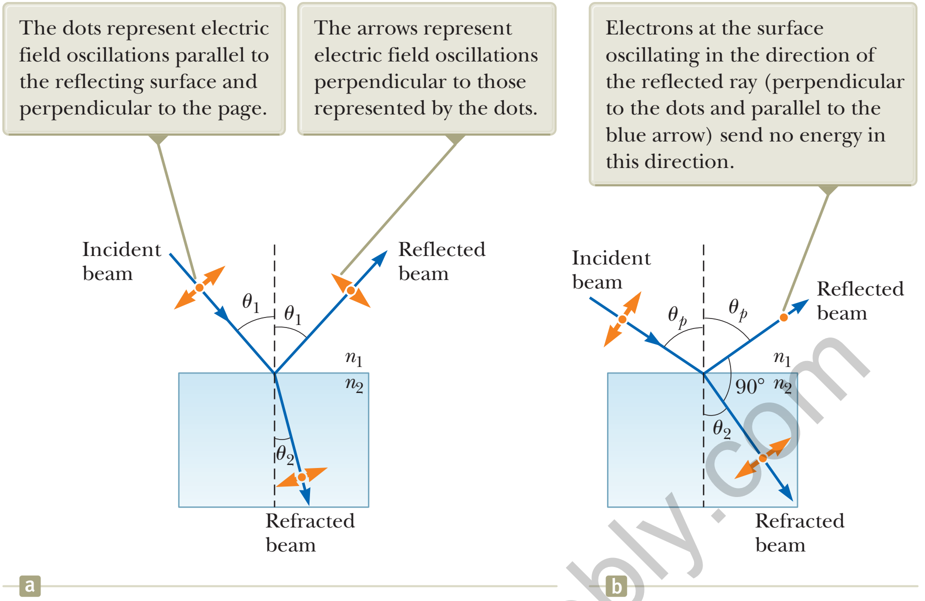


Figure 38.28 (a) When unpolarized light is incident on a reflecting surface, the reflected and refracted beams are partially polarized. (b) The reflected beam is completely polarized when the angle of incidence equals the polarizing angle θ_p, which satisfies the equation $n_2/n_1 = \tan\theta_p$. At this incident angle, the reflected and refracted rays are perpendicular to each other.

We can obtain an expression relating the polarizing angle to the index of refraction of the reflecting substance by using Figure 38.28b. From this figure, we see that $\theta_p + 90° + \theta_2 = 180°$; therefore, $\theta_2 = 90° - \theta_p$. Using Snell's law of refraction (Eq. 35.8) gives

$$\frac{n_2}{n_1} = \frac{\sin\theta_1}{\sin\theta_2} = \frac{\sin\theta_p}{\sin\theta_2}$$

Because $\sin\theta_2 = \sin(90° - \theta_p) = \cos\theta_p$, we can write this expression as $n_2/n_1 = \sin\theta_p/\cos\theta_p$, which means that

Brewster's law ▶

$$\tan\theta_p = \frac{n_2}{n_1} \tag{38.10}$$

This expression is called **Brewster's law,** and the polarizing angle θ_p is sometimes called **Brewster's angle,** after its discoverer, David Brewster (1781–1868). Because n varies with wavelength for a given substance, Brewster's angle is also a function of wavelength.

We can understand polarization by reflection by imagining that the electric field in the incident light sets electrons at the surface of the material in Figure 38.28b into oscillation. The component directions of oscillation are (1) parallel to the arrows shown on the refracted beam of light and therefore parallel to the reflected beam and (2) perpendicular to the page. The oscillating electrons act as dipole antennas radiating light with a polarization parallel to the direction of oscillation. Consult Figure 34.12, which shows the pattern of radiation from a dipole antenna. Notice that there is no radiation at an angle of $\theta = 0$, that is, along the oscillation direction of the antenna. Therefore, for the oscillations in direction 1, there is no radiation in the direction along the reflected ray. For oscillations in direction 2, the electrons radiate light with a polarization perpendicular to the page. Therefore, the light reflected from the surface at this angle is completely polarized parallel to the surface.

Polarization by reflection is a common phenomenon. Sunlight reflected from water, glass, and snow is partially polarized. If the surface is horizontal, the electric field vector of the reflected light has a strong horizontal component. Sunglasses made of polarizing material reduce the glare of reflected light. The transmission axes of such lenses are oriented vertically so that they absorb the strong horizontal component of the reflected light. If you rotate sunglasses through 90°, they are not as effective at blocking the glare from shiny horizontal surfaces.

Polarization by Double Refraction

Solids can be classified on the basis of internal structure. Those in which the atoms are arranged in a specific order are called *crystalline;* the NaCl structure of Figure 38.22 is one example of a crystalline solid. Those solids in which the atoms are distributed randomly are called *amorphous.* When light travels through an amorphous material such as glass, it travels with a speed that is the same in all directions. That is, glass has a single index of refraction. In certain crystalline materials such as calcite and quartz, however, the speed of light is not the same in all directions. In these materials, the speed of light depends on the direction of propagation *and* on the plane of polarization of the light. Such materials are characterized by two indices of refraction. Hence, they are often referred to as **double-refracting** or **birefringent** materials.

When unpolarized light enters a birefringent material, it may split into an **ordinary (O) ray** and an **extraordinary (E) ray.** These two rays have mutually perpendicular polarizations and travel at different speeds through the material. The two speeds correspond to two indices of refraction, n_O for the ordinary ray and n_E for the extraordinary ray.

There is one direction, called the **optic axis,** along which the ordinary and extraordinary rays have the same speed. If light enters a birefringent material at an angle to the optic axis, however, the different indices of refraction will cause the two polarized rays to split and travel in different directions as shown in Figure 38.29.

The index of refraction n_O for the ordinary ray is the same in all directions. If one could place a point source of light inside the crystal as in Figure 38.30, the ordinary waves would spread out from the source as spheres. The index of refraction n_E varies with the direction of propagation. A point source sends out an extraordinary wave having wave fronts that are elliptical in cross section. The difference in speed for the two rays is a maximum in the direction perpendicular to the optic axis. For example, in calcite, $n_O = 1.658$ at a wavelength of 589.3 nm and n_E varies from 1.658 along the optic axis to 1.486 perpendicular to the optic axis. Values for n_O and the extreme value of n_E for various double-refracting crystals are given in Table 38.1.

If you place a calcite crystal on a sheet of paper and then look through the crystal at any writing on the paper, you would see two images as shown in Figure 38.31. As can be seen from Figure 38.29, these two images correspond to one formed by the ordinary ray and one formed by the extraordinary ray. If the two images are viewed through a sheet of rotating polarizing glass, they alternately appear and disappear because the ordinary and extraordinary rays are plane-polarized along mutually perpendicular directions.

Some materials such as glass and plastic become birefringent when stressed. Suppose an unstressed piece of plastic is placed between a polarizer and an analyzer so that light passes from polarizer to plastic to analyzer. When the plastic is unstressed and the analyzer axis is perpendicular to the polarizer axis, none of the polarized light passes through the analyzer. In other words, the unstressed plastic has no effect on the light passing through it. If the plastic is stressed, however, regions of greatest stress become birefringent and the polarization of the light passing through the plastic changes. Hence, a series of bright and dark bands is observed in the transmitted light, with the bright bands corresponding to regions of greatest stress.

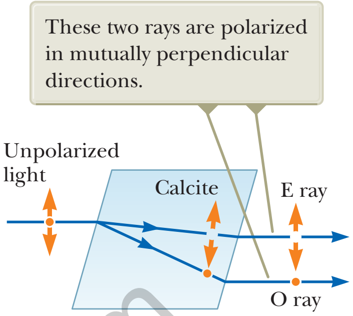


Figure 38.29 Unpolarized light incident at an angle to the optic axis in a calcite crystal splits into an ordinary (O) ray and an extraordinary (E) ray (not to scale).

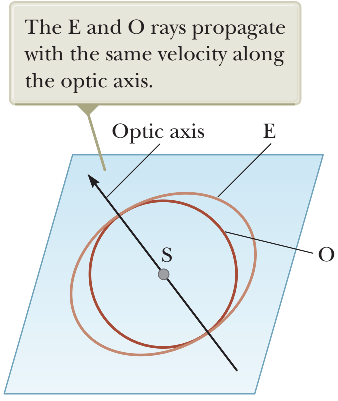


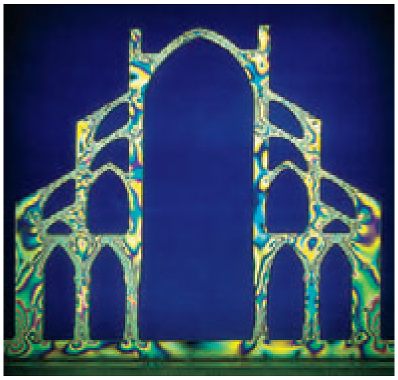
Figure 38.30 A point source S inside a double-refracting crystal produces a spherical wave front corresponding to the ordinary (O) ray and an elliptical wave front corresponding to the extraordinary (E) ray.

Henry Leap and Jim Lehman

Figure 38.31 A calcite crystal produces a double image because it is a birefringent (double-refracting) material.

Table 38.1 **Indices of Refraction for Some Double-Refracting Crystals at a Wavelength of 589.3 nm**

Crystal	n_O	n_E	n_O/n_E
Calcite ($CaCO_3$)	1.658	1.486	1.116
Quartz (SiO_2)	1.544	1.553	0.994
Sodium nitrate ($NaNO_3$)	1.587	1.336	1.188
Sodium sulfite ($NaSO_3$)	1.565	1.515	1.033
Zinc chloride ($ZnCl_2$)	1.687	1.713	0.985
Zinc sulfide (ZnS)	2.356	2.378	0.991



Peter Aprahamian/Science Photo Library, Photo Researchers, Inc.

Figure 38.32 A plastic model of an arch structure under load conditions. The pattern is produced when the plastic model is viewed between a polarizer and analyzer oriented perpendicular to each other. Such patterns are useful in the optimal design of architectural components.

Engineers often use this technique, called *optical stress analysis,* in designing structures ranging from bridges to small tools. They build a plastic model and analyze it under different load conditions to determine regions of potential weakness and failure under stress. An example of a plastic model under stress is shown in Figure 38.32.

Polarization by Scattering

When light is incident on any material, the electrons in the material can absorb and reradiate part of the light. Such absorption and reradiation of light by electrons in the gas molecules that make up air is what causes sunlight reaching an observer on the Earth to be partially polarized. You can observe this effect—called **scattering**—by looking directly up at the sky through a pair of sunglasses whose lenses are made of polarizing material. Less light passes through at certain orientations of the lenses than at others.

Figure 38.33 illustrates how sunlight becomes polarized when it is scattered. The phenomenon is similar to that creating completely polarized light upon reflection from a surface at Brewster's angle. An unpolarized beam of sunlight traveling in the horizontal direction (parallel to the ground) strikes a molecule of one of the gases that make up air, setting the electrons of the molecule into vibration. These vibrating charges act like the vibrating charges in an antenna. The horizontal component of the electric field vector in the incident wave results in a horizontal component of the vibration of the charges, and the vertical component of the vector results in a vertical component of vibration. If the observer in Figure 38.33 is looking straight up (perpendicular to the original direction of propagation of the light), the vertical oscillations of the charges send no radiation toward the observer. Therefore, the observer sees light that is completely polarized in the horizontal direction as indicated by the orange arrows. If the observer looks in other directions, the light is partially polarized in the horizontal direction.

Variations in the color of scattered light in the atmosphere can be understood as follows. When light of various wavelengths λ is incident on gas molecules of diameter d, where $d << \lambda$, the relative intensity of the scattered light varies as $1/\lambda^4$. The condition $d << \lambda$ is satisfied for scattering from oxygen (O_2) and nitrogen (N_2) molecules in the atmosphere, whose diameters are about 0.2 nm. Hence, short wavelengths (violet light) are scattered more efficiently than long wavelengths (red light). Therefore, when sunlight is scattered by gas molecules in the air, the short-wavelength radiation (violet) is scattered more intensely than the long-wavelength radiation (red).

When you look up into the sky in a direction that is not toward the Sun, you see the scattered light, which is predominantly violet. Your eyes, however, are not very sensitive to violet light. Light of the next color in the spectrum, blue, is scattered with less intensity than violet, but your eyes are far more sensitive to blue light than to violet light. Hence, you see a blue sky. If you look toward the west at sunset (or toward the east at sunrise), you are looking in a direction toward the Sun and are seeing light that has passed through a large distance of air. Most of the blue light has been scattered by the air between you and the Sun. The light that survives this

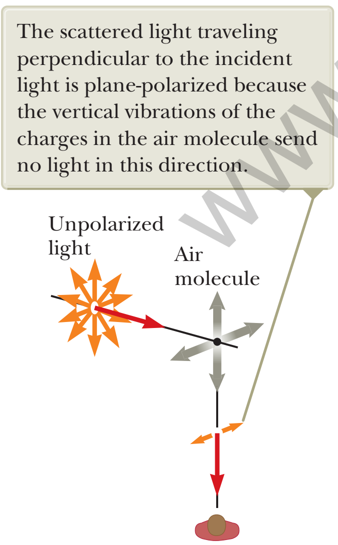


Figure 38.33 The scattering of unpolarized sunlight by air molecules.

trip through the air to you has had much of its blue component scattered and is therefore heavily weighted toward the red end of the spectrum; as a result, you see the red and orange colors of sunset (or sunrise).

Optical Activity

Many important applications of polarized light involve materials that display **optical activity.** A material is said to be optically active if it rotates the plane of polarization of any light transmitted through the material. The angle through which the light is rotated by a specific material depends on the length of the path through the material and on concentration if the material is in solution. One optically active material is a solution of the common sugar dextrose. A standard method for determining the concentration of sugar solutions is to measure the rotation produced by a fixed length of the solution.

Molecular asymmetry determines whether a material is optically active. For example, some proteins are optically active because of their spiral shape.

The liquid crystal displays found in most calculators have their optical activity changed by the application of electric potential across different parts of the display. Try using a pair of polarizing sunglasses to investigate the polarization used in the display of your calculator.

Quick Quiz 38.6 A polarizer for microwaves can be made as a grid of parallel metal wires approximately 1 cm apart. Is the electric field vector for microwaves transmitted through this polarizer **(a)** parallel or **(b)** perpendicular to the metal wires?

Quick Quiz 38.7 You are walking down a long hallway that has many light fixtures in the ceiling and a very shiny, newly waxed floor. When looking at the floor, you see reflections of every light fixture. Now you put on sunglasses that are polarized. Some of the reflections of the light fixtures can no longer be seen. (Try it!) Are the reflections that disappear those **(a)** nearest to you, **(b)** farthest from you, or **(c)** at an intermediate distance from you?

Summary

Concepts and Principles

Diffraction is the deviation of light from a straight-line path when the light passes through an aperture or around an obstacle. Diffraction is due to the wave nature of light.

The **Fraunhofer diffraction pattern** produced by a single slit of width a on a distant screen consists of a central bright fringe and alternating bright and dark fringes of much lower intensities. The angles θ_{dark} at which the diffraction pattern has zero intensity, corresponding to destructive interference, are given by

$$\sin\theta_{\text{dark}} = m\frac{\lambda}{a} \qquad m = \pm1, \pm2, \pm3, \ldots \qquad \textbf{(38.1)}$$

Rayleigh's criterion, which is a limiting condition of resolution, states that two images formed by an aperture are just distinguishable if the central maximum of the diffraction pattern for one image falls on the first minimum of the diffraction pattern for the other image. The limiting angle of resolution for a slit of width a is $\theta_{\text{min}} = \lambda/a$, and the limiting angle of resolution for a circular aperture of diameter D is given by $\theta_{\text{min}} = 1.22\lambda/D$.

A **diffraction grating** consists of a large number of equally spaced, identical slits. The condition for intensity maxima in the interference pattern of a diffraction grating for normal incidence is

$$d\sin\theta_{\text{bright}} = m\lambda \qquad m = 0, \pm1, \pm2, \pm3, \ldots \qquad \textbf{(38.7)}$$

where d is the spacing between adjacent slits and m is the order number of the intensity maximum.

continued

When polarized light of intensity I_{max} is emitted by a polarizer and then is incident on an analyzer, the light transmitted through the analyzer has an intensity equal to $I_{max} \cos^2 \theta$, where θ is the angle between the polarizer and analyzer transmission axes.

In general, reflected light is partially polarized. Reflected light, however, is completely polarized when the angle of incidence is such that the angle between the reflected and refracted beams is 90°. This angle of incidence, called the **polarizing angle** θ_p, satisfies **Brewster's law:**

$$\tan \theta_p = \frac{n_2}{n_1} \quad \textbf{(38.10)}$$

where n_1 is the index of refraction of the medium in which the light initially travels and n_2 is the index of refraction of the reflecting medium.

Objective Questions

$\boxed{1.}$ denotes answer available in *Student Solutions Manual/Study Guide*

1. Certain sunglasses use a polarizing material to reduce the intensity of light reflected as glare from water or automobile windshields. What orientation should the polarizing filters have to be most effective? (a) The polarizers should absorb light with its electric field horizontal. (b) The polarizers should absorb light with its electric field vertical. (c) The polarizers should absorb both horizontal and vertical electric fields. (d) The polarizers should not absorb either horizontal or vertical electric fields.

2. What is most likely to happen to a beam of light when it reflects from a shiny metallic surface at an arbitrary angle? Choose the best answer. (a) It is totally absorbed by the surface. (b) It is totally polarized. (c) It is unpolarized. (d) It is partially polarized. (e) More information is required.

3. In Figure 38.4, assume the slit is in a barrier that is opaque to x-rays as well as to visible light. The photograph in Figure 38.4b shows the diffraction pattern produced with visible light. What will happen if the experiment is repeated with x-rays as the incoming wave and with no other changes? (a) The diffraction pattern is similar. (b) There is no noticeable diffraction pattern but rather a projected shadow of high intensity on the screen, having the same width as the slit. (c) The central maximum is much wider, and the minima occur at larger angles than with visible light. (d) No x-rays reach the screen.

4. A Fraunhofer diffraction pattern is produced on a screen located 1.00 m from a single slit. If a light source of wavelength 5.00×10^{-7} m is used and the distance from the center of the central bright fringe to the first dark fringe is 5.00×10^{-3} m, what is the slit width? (a) 0.010 0 mm (b) 0.100 mm (c) 0.200 mm (d) 1.00 mm (e) 0.005 00 mm

$\boxed{5.}$ Consider a wave passing through a single slit. What happens to the width of the central maximum of its diffraction pattern as the slit is made half as wide? (a) It becomes one-fourth as wide. (b) It becomes one-half as wide. (c) Its width does not change. (d) It becomes twice as wide. (e) It becomes four times as wide.

6. Assume Figure 38.1 was photographed with red light of a single wavelength λ_0. The light passed through a single slit of width a and traveled distance L to the screen where the photograph was made. Consider the width of the central bright fringe, measured between the centers of the dark fringes on both sides of it. Rank from largest to smallest the widths of the central fringe in the following situations and note any cases of equality. (a) The experiment is performed as photographed. (b) The experiment is performed with light whose frequency is increased by 50%. (c) The experiment is performed with light whose wavelength is increased by 50%. (d) The experiment is performed with the original light and with a slit of width $2a$. (e) The experiment is performed with the original light and slit and with distance $2L$ to the screen.

7. If plane polarized light is sent through two polarizers, the first at 45° to the original plane of polarization and the second at 90° to the original plane of polarization, what fraction of the original polarized intensity passes through the last polarizer? (a) 0 (b) $\frac{1}{4}$ (c) $\frac{1}{2}$ (d) $\frac{1}{8}$ (e) $\frac{1}{10}$

8. Why is it advantageous to use a large-diameter objective lens in a telescope? (a) It diffracts the light more effectively than smaller-diameter objective lenses. (b) It increases its magnification. (c) It enables you to see more objects in the field of view. (d) It reflects unwanted wavelengths. (e) It increases its resolution.

9. What combination of optical phenomena causes the bright colored patterns sometimes seen on wet streets covered with a layer of oil? Choose the best answer. (a) diffraction and polarization (b) interference and diffraction (c) polarization and reflection (d) refraction and diffraction (e) reflection and interference

10. When you receive a chest x-ray at a hospital, the x-rays pass through a set of parallel ribs in your chest. Do your ribs act as a diffraction grating for x-rays? (a) Yes. They produce diffracted beams that can be observed separately. (b) Not to a measurable extent. The ribs are too far apart. (c) Essentially not. The ribs are too close together. (d) Essentially not. The ribs are too few in number. (e) Absolutely not. X-rays cannot diffract.

11. When unpolarized light passes through a diffraction grating, does it become polarized? (a) No, it does not. (b) Yes, it does, with the transmission axis parallel to the slits or grooves in the grating. (c) Yes, it does, with the transmission axis perpendicular to the slits or grooves in the grating. (d) It possibly does because an electric field above some threshold is blocked out by the grating if the field is perpendicular to the slits.

12. Off in the distance, you see the headlights of a car, but they are indistinguishable from the single headlight of a motorcycle. Assume the car's headlights are now switched from low beam to high beam so that the light intensity you receive becomes three times greater. What then happens to your ability to resolve the two light sources? (a) It increases by a factor of 9. (b) It increases by a factor of 3. (c) It remains the same. (d) It becomes one-third as good. (e) It becomes one-ninth as good.

Conceptual Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. The atoms in a crystal lie in planes separated by a few tenths of a nanometer. Can they produce a diffraction pattern for visible light as they do for x-rays? Explain your answer with reference to Bragg's law.

2. Holding your hand at arm's length, you can readily block sunlight from reaching your eyes. Why can you not block sound from reaching your ears this way?

3. How could the index of refraction of a flat piece of opaque obsidian glass be determined?

4. (a) Is light from the sky polarized? (b) Why is it that clouds seen through Polaroid glasses stand out in bold contrast to the sky?

5. A laser beam is incident at a shallow angle on a horizontal machinist's ruler that has a finely calibrated scale. The engraved rulings on the scale give rise to a diffraction pattern on a vertical screen. Discuss how you can use this technique to obtain a measure of the wavelength of the laser light.

6. If a coin is glued to a glass sheet and this arrangement is held in front of a laser beam, the projected shadow has diffraction rings around its edge and a bright spot in the center. How are these effects possible?

7. Fingerprints left on a piece of glass such as a windowpane often show colored spectra like that from a diffraction grating. Why?

8. A laser produces a beam a few millimeters wide, with uniform intensity across its width. A hair is stretched vertically across the front of the laser to cross the beam. (a) How is the diffraction pattern it produces on a distant screen related to that of a vertical slit equal in width to the hair? (b) How could you determine the width of the hair from measurements of its diffraction pattern?

9. A radio station serves listeners in a city to the northeast of its broadcast site. It broadcasts from three adjacent towers on a mountain ridge, along a line running east to west, in what's called a *phased array*. Show that by introducing time delays among the signals the individual towers radiate, the station can maximize net intensity in the direction toward the city (and in the opposite direction) and minimize the signal transmitted in other directions.

10. John William Strutt, Lord Rayleigh (1842–1919), invented an improved foghorn. To warn ships of a coastline, a foghorn should radiate sound in a wide horizontal sheet over the ocean's surface. It should not waste energy by broadcasting sound upward or downward. Rayleigh's foghorn trumpet is shown in two possible configurations, horizontal and vertical, in Figure CQ38.10. Which is the correct orientation? Decide whether the long dimension of the rectangular opening should be horizontal or vertical and argue for your decision.

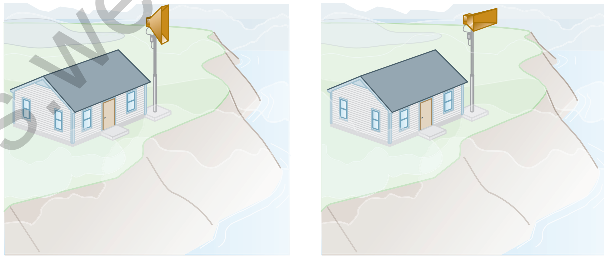


Figure CQ38.10

11. Why can you hear around corners, but not see around corners?

12. Figure CQ38.12 shows a megaphone in use. Construct a theoretical description of how a megaphone works. You may assume the sound of your voice radiates just through the opening of your mouth. Most of the information in speech is carried not in a signal at the fundamental frequency, but in noises and in harmonics, with frequencies of a few thousand hertz. Does your theory allow any prediction that is simple to test?

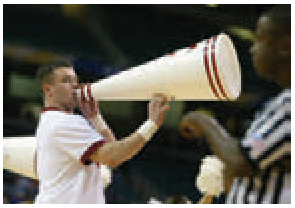


Figure CQ38.12

Problems

WebAssign The problems found in this chapter may be assigned online in Enhanced WebAssign

1. straightforward; **2.** intermediate; **3.** challenging

1. full solution available in the *Student Solutions Manual/Study Guide*

AMT Analysis Model tutorial available in Enhanced WebAssign

GP Guided Problem

M Master It tutorial available in Enhanced WebAssign

W Watch It video solution available in Enhanced WebAssign

Section 38.2 Diffraction Patterns from Narrow Slits

1. Light of wavelength 587.5 nm illuminates a slit of width 0.75 mm. (a) At what distance from the slit should a screen be placed if the first minimum in the diffraction pattern is to be 0.85 mm from the central maximum? (b) Calculate the width of the central maximum.

2. Helium–neon laser light (λ = 632.8 nm) is sent through a 0.300-mm-wide single slit. What is the width of the central maximum on a screen 1.00 m from the slit?

3. Sound with a frequency 650 Hz from a distant source passes through a doorway 1.10 m wide in a sound-absorbing wall. Find (a) the number and (b) the angular directions of the diffraction minima at listening positions along a line parallel to the wall.

4. A horizontal laser beam of wavelength 632.8 nm has a circular cross section 2.00 mm in diameter. A rectangular aperture is to be placed in the center of the beam so that when the light falls perpendicularly on a wall 4.50 m away, the central maximum fills a rectangle 110 mm wide and 6.00 mm high. The dimensions are measured between the minima bracketing the central maximum. Find the required (a) width and (b) height of the aperture. (c) Is the longer dimension of the central bright patch in the diffraction pattern horizontal or vertical? (d) Is the longer dimension of the aperture horizontal or vertical? (e) Explain the relationship between these two rectangles, using a diagram.

5. Coherent microwaves of wavelength 5.00 cm enter a tall, narrow window in a building otherwise essentially opaque to the microwaves. If the window is 36.0 cm wide, what is the distance from the central maximum to the first-order minimum along a wall 6.50 m from the window?

6. Light of wavelength 540 nm passes through a slit of width 0.200 mm. (a) The width of the central maximum on a screen is 8.10 mm. How far is the screen from the slit? (b) Determine the width of the first bright fringe to the side of the central maximum.

7. A screen is placed 50.0 cm from a single slit, which is illuminated with light of wavelength 690 nm. If the distance between the first and third minima in the diffraction pattern is 3.00 mm, what is the width of the slit?

8. A screen is placed a distance L from a single slit of width a, which is illuminated with light of wavelength λ. Assume $L >> a$. If the distance between the minima for $m = m_1$ and $m = m_2$ in the diffraction pattern is Δy, what is the width of the slit?

9. Assume light of wavelength 650 nm passes through two slits 3.00 μm wide, with their centers 9.00 μm apart. Make a sketch of the combined diffraction and interference pattern in the form of a graph of intensity versus $\phi = (\pi a \sin\theta)/\lambda$. You may use Figure 38.7 as a starting point.

10. **What If?** Suppose light strikes a single slit of width a at an angle β from the perpendicular direction as shown in Figure P38.10. Show that Equation 38.1, the condition for destructive interference, must be modified to read

$$\sin\theta_{\text{dark}} = m\frac{\lambda}{a} - \sin\beta$$
$$m = \pm1, \pm2, \pm3, \ldots$$

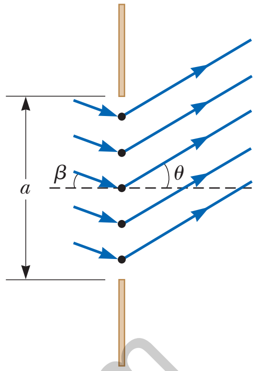


Figure P38.10

11. A diffraction pattern is formed on a screen 120 cm away from a 0.400-mm-wide slit. Monochromatic 546.1-nm light is used. Calculate the fractional intensity $I/I_{\max}$ at a point on the screen 4.10 mm from the center of the principal maximum.

12. Coherent light of wavelength 501.5 nm is sent through two parallel slits in an opaque material. Each slit is 0.700 μm wide. Their centers are 2.80 μm apart. The light then falls on a semicylindrical screen, with its axis at the midline between the slits. We would like to describe the appearance of the pattern of light visible on the screen. (a) Find the direction for each two-slit interference maximum on the screen as an angle away from the bisector of the line joining the slits. (b) How many angles are there that represent two-slit interference maxima? (c) Find the direction for each single-slit interference minimum on the screen as an angle away from the bisector of the line joining the slits. (d) How many angles are there that represent single-slit interference minima? (e) How many of the angles in part (d) are identical to those in part (a)? (f) How many bright fringes are visible on the screen? (g) If the intensity of the central fringe is $I_{\max}$, what is the intensity of the last fringe visible on the screen?

13. A beam of monochromatic light is incident on a single slit of width 0.600 mm. A diffraction pattern forms on a wall 1.30 m beyond the slit. The distance between the positions of zero intensity on both sides of the central maximum is 2.00 mm. Calculate the wavelength of the light.

Section 38.3 Resolution of Single-Slit and Circular Apertures

Note: In Problems 14, 19, 22, 23, and 67, you may use the Rayleigh criterion for the limiting angle of resolution of an eye. The standard may be overly optimistic for human vision.

14. The pupil of a cat's eye narrows to a vertical slit of width 0.500 mm in daylight. Assume the average wavelength of the light is 500 nm. What is the angular resolution for horizontally separated mice?

15. The angular resolution of a radio telescope is to be 0.100° when the incident waves have a wavelength of 3.00 mm. What minimum diameter is required for the telescope's receiving dish?

16. A *pinhole camera* has a small circular aperture of diameter D. Light from distant objects passes through the aperture into an otherwise dark box, falling on a screen at the other end of the box. The aperture in a pinhole camera has diameter D = 0.600 mm. Two

point sources of light of wavelength 550 nm are at a distance L from the hole. The separation between the sources is 2.80 cm, and they are just resolved by the camera. What is L?

17. The objective lens of a certain refracting telescope has a diameter of 58.0 cm. The telescope is mounted in a satellite that orbits the Earth at an altitude of 270 km to view objects on the Earth's surface. Assuming an average wavelength of 500 nm, find the minimum distance between two objects on the ground if their images are to be resolved by this lens.

18. Yellow light of wavelength 589 nm is used to view an object under a microscope. The objective lens diameter is 9.00 mm. (a) What is the limiting angle of resolution? (b) Suppose it is possible to use visible light of any wavelength. What color should you choose to give the smallest possible angle of resolution, and what is this angle? (c) Suppose water fills the space between the object and the objective. What effect does this change have on the resolving power when 589-nm light is used?

19. What is the approximate size of the smallest object on the Earth that astronauts can resolve by eye when they are orbiting 250 km above the Earth? Assume λ = 500 nm and a pupil diameter of 5.00 mm.

20. M A helium–neon laser emits light that has a wavelength of 632.8 nm. The circular aperture through which the beam emerges has a diameter of 0.500 cm. Estimate the diameter of the beam 10.0 km from the laser.

21. M To increase the resolving power of a microscope, the object and the objective are immersed in oil (n = 1.5). If the limiting angle of resolution without the oil is 0.60 μrad, what is the limiting angle of resolution with the oil? *Hint:* The oil changes the wavelength of the light.

22. Narrow, parallel, glowing gas-filled tubes in a variety of colors form block letters to spell out the name of a nightclub. Adjacent tubes are all 2.80 cm apart. The tubes forming one letter are filled with neon and radiate predominantly red light with a wavelength of 640 nm. For another letter, the tubes emit predominantly blue light at 440 nm. The pupil of a dark-adapted viewer's eye is 5.20 mm in diameter. (a) Which color is easier to resolve? State how you decide. (b) If she is in a certain range of distances away, the viewer can resolve the separate tubes of one color but not the other. The viewer's distance must be in what range for her to resolve the tubes of only one of these two colors?

23. M Impressionist painter Georges Seurat created paintings with an enormous number of dots of pure pigment, each of which was approximately 2.00 mm in diameter. The idea was to have colors such as red and green next to each other to form a scintillating canvas, such as in his masterpiece, *A Sunday Afternoon on the Island of La Grande Jatte* (Fig. P38.23). Assume λ = 500 nm and a pupil diameter of 5.00 mm. Beyond what distance would a viewer be unable to discern individual dots on the canvas?

© SuperStock/SuperStock

Figure P38.23

24. W A circular radar antenna on a Coast Guard ship has a diameter of 2.10 m and radiates at a frequency of 15.0 GHz. Two small boats are located 9.00 km away from the ship. How close together could the boats be and still be detected as two objects?

Section 38.4 The Diffraction Grating

Note: In the following problems, assume the light is incident normally on the gratings.

25. A helium–neon laser (λ = 632.8 nm) is used to calibrate a diffraction grating. If the first-order maximum occurs at 20.5°, what is the spacing between adjacent grooves in the grating?

26. M White light is spread out into its spectral components by a diffraction grating. If the grating has 2 000 grooves per centimeter, at what angle does red light of wavelength 640 nm appear in first order?

27. Consider an array of parallel wires with uniform spacing of 1.30 cm between centers. In air at 20.0°C, ultrasound with a frequency of 37.2 kHz from a distant source is incident perpendicular to the array. (a) Find the number of directions on the other side of the array in which there is a maximum of intensity. (b) Find the angle for each of these directions relative to the direction of the incident beam.

28. W Three discrete spectral lines occur at angles of 10.1°, 13.7°, and 14.8° in the first-order spectrum of a grating spectrometer. (a) If the grating has 3 660 slits/cm, what are the wavelengths of the light? (b) At what angles are these lines found in the second-order spectrum?

29. The laser in a compact disc player must precisely follow the spiral track on the CD, along which the distance between one loop of the spiral and the next is only about 1.25 μm. Figure P38.29 (page 1186) shows how a diffraction grating is used to provide information to keep the beam on track. The laser light passes through a diffraction grating before it reaches the CD. The strong central maximum of the diffraction pattern is used to read the information in the track of pits. The two first-order side maxima are designed to fall on the flat surfaces on both sides of the information track and are used for steering. As long as both beams are reflecting from smooth, nonpitted surfaces, they are detected with constant high intensity. If the main beam wanders off the track, however, one of the side beams begins to strike pits on the information track and the reflected light diminishes. This change is used with an electronic circuit to guide the beam back to the desired location. Assume the laser light has a wavelength of 780 nm and the diffraction grating is positioned 6.90 μm from the disk. Assume the first-order beams are to fall on the CD 0.400 μm on either side of the information track. What should be the number of grooves per millimeter in the grating?

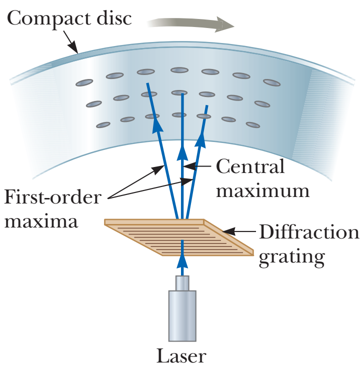


Figure P38.29

30. A grating with 250 grooves/mm is used with an incandescent light source. Assume the visible spectrum to range in wavelength from 400 nm to 700 nm. In how many orders can one see (a) the entire visible spectrum and (b) the short-wavelength region of the visible spectrum?

31. A diffraction grating has 4 200 rulings/cm. On a screen 2.00 m from the grating, it is found that for a particular order *m*, the maxima corresponding to two closely spaced wavelengths of sodium (589.0 nm and 589.6 nm) are separated by 1.54 mm. Determine the value of *m*.

32. M The hydrogen spectrum includes a red line at 656 nm and a blue-violet line at 434 nm. What are the angular separations between these two spectral lines for all visible orders obtained with a diffraction grating that has 4 500 grooves/cm?

33. W Light from an argon laser strikes a diffraction grating that has 5 310 grooves per centimeter. The central and first-order principal maxima are separated by 0.488 m on a wall 1.72 m from the grating. Determine the wavelength of the laser light.

34. Show that whenever white light is passed through a diffraction grating of any spacing size, the violet end of the spectrum in the third order on a screen always overlaps the red end of the spectrum in the second order.

35. Light of wavelength 500 nm is incident normally on a diffraction grating. If the third-order maximum of the diffraction pattern is observed at 32.0°, (a) what is the number of rulings per centimeter for the grating? (b) Determine the total number of primary maxima that can be observed in this situation.

36. A wide beam of laser light with a wavelength of 632.8 nm is directed through several narrow parallel slits, separated by 1.20 mm, and falls on a sheet of photographic film 1.40 m away. The exposure time is chosen so that the film stays unexposed everywhere except at the central region of each bright fringe. (a) Find the distance between these interference maxima. The film is printed as a transparency; it is opaque everywhere except at the exposed lines. Next, the same beam of laser light is directed through the transparency and allowed to fall on a screen 1.40 m beyond. (b) Argue that several narrow, parallel, bright regions, separated by 1.20 mm, appear on the screen as real images of the original slits. (A similar train of thought, at a soccer game, led Dennis Gabor to invent holography.)

37. A beam of bright red light of wavelength 654 nm passes through a diffraction grating. Enclosing the space beyond the grating is a large semicylindrical screen centered on the grating, with its axis parallel to the slits in the grating. Fifteen bright spots appear on the screen. Find (a) the maximum and (b) the minimum possible values for the slit separation in the diffraction grating.

Section 38.5 Diffraction of X-Rays by Crystals

38. M If the spacing between planes of atoms in a NaCl crystal is 0.281 nm, what is the predicted angle at which 0.140-nm x-rays are diffracted in a first-order maximum?

39. AMT M Potassium iodide (KI) has the same crystalline structure as NaCl, with atomic planes separated by 0.353 nm. A monochromatic x-ray beam shows a first-order diffraction maximum when the grazing angle is 7.60°. Calculate the x-ray wavelength.

40. Monochromatic x-rays (λ = 0.166 nm) from a nickel target are incident on a potassium chloride (KCl) crystal surface. The spacing between planes of atoms in KCl is 0.314 nm. At what angle (relative to the surface) should the beam be directed for a second-order maximum to be observed?

41. The first-order diffraction maximum is observed at 12.6° for a crystal having a spacing between planes of atoms of 0.250 nm. (a) What wavelength x-ray is used to observe this first-order pattern? (b) How many orders can be observed for this crystal at this wavelength?

Section 38.6 Polarization of Light Waves

Problem 62 in Chapter 34 can be assigned with this section.

42. *Why is the following situation impossible?* A technician is measuring the index of refraction of a solid material by observing the polarization of light reflected from its surface. She notices that when a light beam is projected from air onto the material surface, the reflected light is totally polarized parallel to the surface when the incident angle is 41.0°.

43. M Plane-polarized light is incident on a single polarizing disk with the direction of $\vec{\mathbf{E}}_0$ parallel to the direction of the transmission axis. Through what angle should the disk be rotated so that the intensity in the transmitted beam is reduced by a factor of (a) 3.00, (b) 5.00, and (c) 10.0?

44. AMT The angle of incidence of a light beam onto a reflecting surface is continuously variable. The reflected ray in air is completely polarized when the angle of incidence is 48.0°. What is the index of refraction of the reflecting material?

45. Unpolarized light passes through two ideal Polaroid sheets. The axis of the first is vertical, and the axis of the second is at 30.0° to the vertical. What fraction of the incident light is transmitted?

46. Two handheld radio transceivers with dipole antennas are separated by a large fixed distance. If the transmitting antenna is vertical, what fraction of the maximum received power will appear in the receiving antenna when it is inclined from the vertical by (a) 15.0°, (b) 45.0°, and (c) 90.0°?

47. You use a sequence of ideal polarizing filters, each with its axis making the same angle with the axis of the previous filter, to rotate the plane of polarization of a polarized light beam by a total of 45.0°. You wish to have an intensity reduction no larger than 10.0%. (a) How many polarizers do you need to achieve your goal? (b) What is the angle between adjacent polarizers?

48. An unpolarized beam of light is incident on a stack of ideal polarizing filters. The axis of the first filter is perpendicular to the axis of the last filter in the stack. Find the fraction by which the transmitted beam's intensity is reduced in the three following cases. (a) Three filters are in the stack, each with its transmission axis at 45.0° relative to the preceding filter. (b) Four filters are in the stack, each with its transmission axis at 30.0° relative to the preceding filter. (c) Seven filters are in the stack, each with its transmission axis at 15.0° relative to the preceding filter. (d) Comment on comparing the answers to parts (a), (b), and (c).

49. The critical angle for total internal reflection for sapphire surrounded by air is 34.4°. Calculate the polarizing angle for sapphire.

50. For a particular transparent medium surrounded by air, find the polarizing angle θ_p in terms of the critical angle for total internal reflection θ_c.

51. M Three polarizing plates whose planes are parallel are centered on a common axis. The directions of the transmission axes relative to the common vertical direction are shown in Figure P38.51. A linearly polarized beam of light with plane of polarization parallel to the vertical reference direction is incident from the left onto the first disk with intensity I_i = 10.0 units (arbitrary). Calculate the transmitted intensity I_f when θ_1 = 20.0°, θ_2 = 40.0°, and θ_3 = 60.0°. *Hint:* Make repeated use of Malus's law.

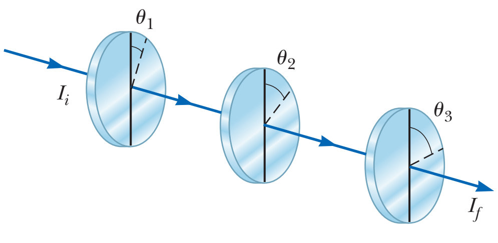


Figure P38.51

52. Two polarizing sheets are placed together with their transmission axes crossed so that no light is transmitted. A third sheet is inserted between them with its transmission axis at an angle of 45.0° with respect to each of the other axes. Find the fraction of incident unpolarized light intensity transmitted by the three-sheet combination. (Assume each polarizing sheet is ideal.)

Additional Problems

53. In a single-slit diffraction pattern, assuming each side maximum is halfway between the adjacent minima, find the ratio of the intensity of (a) the first-order side maximum and (b) the second-order side maximum to the intensity of the central maximum.

54. Laser light with a wavelength of 632.8 nm is directed through one slit or two slits and allowed to fall on a screen 2.60 m beyond. Figure P38.54 shows the pattern on the screen, with a centimeter ruler below it. (a) Did the light pass through one slit or two slits? Explain how you can determine the answer. (b) If one slit, find its width. If two slits, find the distance between their centers.

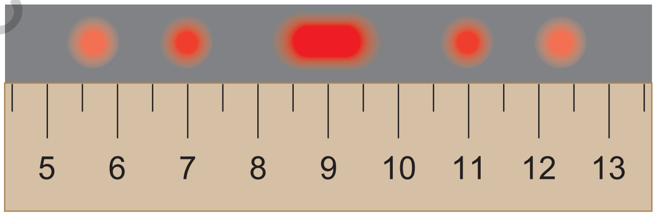


Figure P38.54

55. In water of uniform depth, a wide pier is supported on pilings in several parallel rows 2.80 m apart. Ocean waves of uniform wavelength roll in, moving in a direction that makes an angle of 80.0° with the rows of pilings. Find the three longest wavelengths of waves that are strongly reflected by the pilings.

56. The second-order dark fringe in a single-slit diffraction pattern is 1.40 mm from the center of the central maximum. Assuming the screen is 85.0 cm from a slit of width 0.800 mm and assuming monochromatic incident light, calculate the wavelength of the incident light.

57. Light from a helium–neon laser (λ = 632.8 nm) is incident on a single slit. What is the maximum width of the slit for which no diffraction minima are observed?

58. W Two motorcycles separated laterally by 2.30 m are approaching an observer wearing night-vision goggles sensitive to infrared light of wavelength 885 nm. (a) Assume the light propagates through perfectly steady and uniform air. What aperture diameter is required if the motorcycles' headlights are to be resolved at a distance of 12.0 km? (b) Comment on how realistic the assumption in part (a) is.

59. The *Very Large Array* (VLA) is a set of 27 radio telescope dishes in Catron and Socorro counties, New Mexico (Fig. P38.59). The antennas can be moved apart on railroad tracks, and their combined signals give the resolving power of a synthetic aperture 36.0 km in diameter. (a) If the detectors are tuned to a frequency of 1.40 GHz, what is the angular resolution of the VLA? (b) Clouds of interstellar hydrogen radiate at the frequency used in part (a). What must be the separation distance of two clouds at the center of the galaxy, 26 000 light-years away, if they are to be resolved? (c) **What If?** As the telescope looks up, a circling hawk looks down. Assume the hawk is most sensitive to green light having a wavelength of 500 nm and has a pupil of diameter 12.0 mm. Find the angular resolution of the hawk's eye. (d) A mouse is on the ground 30.0 m below. By what distance must the mouse's whiskers be separated if the hawk can resolve them?

Figure P38.59

60. Two wavelengths λ and $\lambda + \Delta\lambda$ (with $\Delta\lambda << \lambda$) are incident on a diffraction grating. Show that the angular separation between the spectral lines in the mth-order spectrum is

$$\Delta\theta = \frac{\Delta\lambda}{\sqrt{(d/m)^2 - \lambda^2}}$$

where d is the slit spacing and m is the order number.

61. **Review.** A beam of 541-nm light is incident on a diffraction grating that has 400 grooves/mm. (a) Determine the angle of the second-order ray. (b) **What If?** If the entire apparatus is immersed in water, what is the new second-order angle of diffraction? (c) Show that the two diffracted rays of parts (a) and (b) are related through the law of refraction.

62. *Why is the following situation impossible?* A technician is sending laser light of wavelength 632.8 nm through a pair of slits separated by 30.0 μm. Each slit is of width 2.00 μm. The screen on which he projects the pattern is not wide enough, so light from the m = 15 interference maximum misses the edge of the screen and passes into the next lab station, startling a coworker.

63. A 750-nm light beam in air hits the flat surface of a certain liquid, and the beam is split into a reflected ray and a refracted ray. If the reflected ray is completely polarized when it is at 36.0° with respect to the surface, what is the wavelength of the refracted ray?

64. Iridescent peacock feathers are shown in Figure P38.64a. The surface of one microscopic barbule is composed of transparent keratin that supports rods of dark brown melanin in a regular lattice, represented in Figure P38.64b. (Your fingernails are made of keratin, and melanin is the dark pigment giving color to human skin.) In a portion of the feather that can appear turquoise (blue-green), assume the melanin rods are uniformly separated by 0.25 μm, with air between them. (a) Explain how this structure can appear turquoise when it contains no blue or green pigment. (b) Explain how it can also appear violet if light falls on it in a different direction. (c) Explain how it can present different colors to your two eyes simultaneously, which is a characteristic of iridescence. (d) A compact disc can appear to be any color of the rainbow. Explain why the portion of the feather in Figure P38.64b cannot appear yellow or red. (e) What could be different about the array of melanin rods in a portion of the feather that does appear to be red?

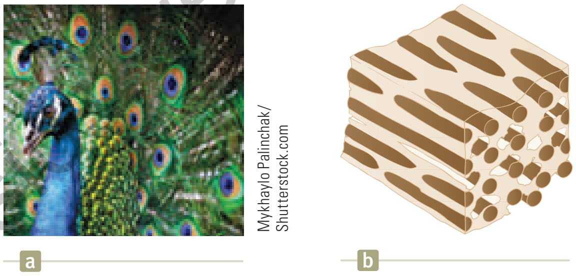


Figure P38.64

65. AMT Light in air strikes a water surface at the polarizing angle. The part of the beam refracted into the water strikes a submerged slab of material with refractive index n = 1.62 as shown in Figure P38.65. The light reflected from the upper surface of the slab is completely polarized. Find the angle θ between the water surface and the surface of the slab.

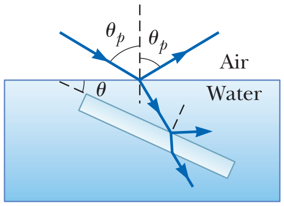


Figure P38.65
Problems 65 and 66.

66. Light in air (assume n = 1) strikes the surface of a liquid of index of refraction n_ℓ at the polarizing angle. The part of the beam refracted into the liquid strikes a submerged slab of material with refractive index n as shown in Figure P38.65. The light reflected from the upper surface of the slab is completely polarized. Find the angle θ between the water surface and the surface of the slab as a function of n and n_ℓ.

67. An American standard analog television picture (non-HDTV), also known as NTSC, is composed of approximately 485 visible horizontal lines of varying light intensity. Assume your ability to resolve the lines is limited only by the Rayleigh criterion, the pupils of your eyes are 5.00 mm in diameter, and the average wavelength of the light coming from the screen is 550 nm. Calculate the ratio of the minimum viewing distance to the vertical dimension of the picture such that you will not be able to resolve the lines.

68. A pinhole camera has a small circular aperture of diameter D. Light from distant objects passes through the aperture into an otherwise dark box, falling on a screen located a distance L away. If D is too large, the display on the screen will be fuzzy because a bright point in the field of view will send light onto a circle of diameter slightly larger than D. On the other hand, if D is too small, diffraction will blur the display on the screen. The screen shows a reasonably sharp image if the diameter of the central disk of the diffraction pattern, specified by Equation 38.6, is equal to D at the screen. (a) Show that for monochromatic light with plane wave fronts and $L >> D$, the condition for a sharp view is fulfilled if $D^2 = 2.44\lambda L$. (b) Find the optimum pinhole diameter for 500-nm light projected onto a screen 15.0 cm away.

69. The *scale* of a map is a number of kilometers per centimeter specifying the distance on the ground that any distance on the map represents. The scale of a spectrum is its *dispersion,* a number of nanometers per centimeter, specifying the change in wavelength that a distance across the spectrum represents. You must know the dispersion if you want to compare one spectrum with another or make a measurement of, for example, a Doppler shift. Let y represent the position relative to the center of a diffraction pattern projected onto a flat screen at distance L by a diffraction grating with slit spacing d. The dispersion is $d\lambda/dy$. (a) Prove that the dispersion is given by

$$\frac{d\lambda}{dy} = \frac{L^2 d}{m(L^2 + y^2)^{3/2}}$$

(b) A light with a mean wavelength of 550 nm is analyzed with a grating having 8 000 rulings/cm and projected onto a screen 2.40 m away. Calculate the dispersion in first order.

70. (a) Light traveling in a medium of index of refraction n_1 is incident at an angle θ on the surface of a medium of index n_2. The angle between reflected and refracted rays is β. Show that

$$\tan\theta = \frac{n_2 \sin\beta}{n_1 - n_2 \cos\beta}$$

(b) **What If?** Show that this expression for $\tan\theta$ reduces to Brewster's law when $\beta = 90°$.

71. The intensity of light in a diffraction pattern of a single slit is described by the equation

$$I = I_{\max}\frac{\sin^2\phi}{\phi^2}$$

where $\phi = (\pi a \sin\theta)/\lambda$. The central maximum is at $\phi = 0$, and the side maxima are *approximately* at $\phi = (m + \frac{1}{2})\pi$ for $m = 1, 2, 3, \ldots$. Determine more precisely (a) the location of the first side maximum, where $m = 1$, and (b) the location of the second side maximum. *Suggestion:* Observe in Figure 38.6a that the graph of intensity versus ϕ has a horizontal tangent at maxima and also at minima.

72. How much diffraction spreading does a light beam undergo? One quantitative answer is the *full width at half maximum* of the central maximum of the single-slit Fraunhofer diffraction pattern. You can evaluate this angle of spreading in this problem. (a) In Equation 38.2, define $\phi = \pi a \sin\theta/\lambda$ and show that at the point where $I = 0.5I_{\max}$ we must have $\phi = \sqrt{2}\sin\phi$. (b) Let $y_1 = \sin\phi$ and $y_2 = \phi/\sqrt{2}$. Plot y_1 and y_2 on the same set of axes over a range from $\phi = 1$ rad to $\phi = \pi/2$ rad. Determine ϕ from the point of intersection of the two curves. (c) Then show that if the fraction λ/a is not large, the angular full width at half maximum of the central diffraction maximum is $\theta = 0.885\lambda/a$. (d) **What If?** Another method to solve the transcendental equation $\phi = \sqrt{2}\sin\phi$ in part (a) is to guess a first value of ϕ, use a computer or calculator to see how nearly it fits, and continue to update your estimate until the equation balances. How many steps (iterations) does this process take?

73. Two closely spaced wavelengths of light are incident on a diffraction grating. (a) Starting with Equation 38.7, show that the angular dispersion of the grating is given by

$$\frac{d\theta}{d\lambda} = \frac{m}{d\cos\theta}$$

(b) A square grating 2.00 cm on each side containing 8 000 equally spaced slits is used to analyze the spectrum of mercury. Two closely spaced lines emitted by this element have wavelengths of 579.065 nm and 576.959 nm. What is the angular separation of these two wavelengths in the second-order spectrum?

74. Light of wavelength 632.8 nm illuminates a single slit, and a diffraction pattern is formed on a screen 1.00 m from the slit. (a) Using the data in the following table, plot relative intensity versus position. Choose an appropriate value for the slit width a and, on the same graph used for the experimental data, plot the theoretical expression for the relative intensity

$$\frac{I}{I_{\max}} = \frac{\sin^2\phi}{\phi^2}$$

where $\phi = (\pi a \sin\theta)/\lambda$. (b) What value of a gives the best fit of theory and experiment?

Position Relative to Central Maximum (mm)	Relative Intensity
0	1.00
0.8	0.95
1.6	0.80
3.2	0.39
4.8	0.079
6.5	0.003
8.1	0.036
9.7	0.043
11.3	0.013
12.9	0.000 3
14.5	0.012
16.1	0.015
17.7	0.004 4
19.3	0.000 3

Challenge Problems

75. Figure P38.75a is a three-dimensional sketch of a birefringent crystal. The dotted lines illustrate how a thin, parallel-faced slab of material could be cut from the larger specimen with the crystal's optic axis parallel to the faces of the plate. A section cut from the crystal in this manner is known as a *retardation plate.* When a beam of light is incident on the plate perpendicular to the direction of the optic axis as shown in Figure P38.75b, the O ray and the E ray travel along a single straight line, but with different speeds. The figure shows the wave fronts for the two rays. (a) Let the thickness of the plate be d. Show that the phase difference between the O ray and the E ray after traveling the thickness of the plate is

$$\theta = \frac{2\pi d}{\lambda}|n_O - n_E|$$

where λ is the wavelength in air. (b) In a particular case, the incident light has a wavelength of 550 nm. Find the minimum value of d for a quartz plate for which $\theta = \pi/2$. Such a plate is called a *quarter-wave plate.* Use values of n_O and n_E from Table 38.1.

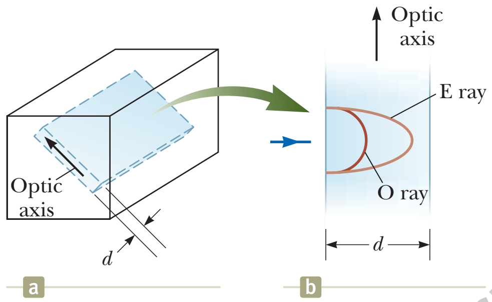


Figure P38.75

76. A spy satellite can consist of a large-diameter concave mirror forming an image on a digital-camera detector and sending the picture to a ground receiver by radio waves. In effect, it is an astronomical telescope in orbit, looking down instead of up. (a) Can a spy satellite read a license plate? (b) Can it read the date on a dime? Argue for your answers by making an order-of-magnitude calculation, specifying the data you estimate.

77. Suppose the single slit in Figure 38.4 is 6.00 cm wide and in front of a microwave source operating at 7.50 GHz. (a) Calculate the angle for the first minimum in the diffraction pattern. (b) What is the relative intensity I/I_{max} at $\theta = 15.0°$? (c) Assume two such sources, separated laterally by 20.0 cm, are behind the slit. What must be the maximum distance between the plane of the sources and the slit if the diffraction patterns are to be resolved? In this case, the approximation $\sin\theta \approx \tan\theta$ is not valid because of the relatively small value of a/λ.

78. In Figure P38.78, suppose the transmission axes of the left and right polarizing disks are perpendicular to each other. Also, let the center disk be rotated on the common axis with an angular speed ω. Show that if unpolarized light is incident on the left disk with an intensity I_{max}, the intensity of the beam emerging from the right disk is

$$I = \tfrac{1}{16}I_{max}(1 - \cos 4\omega t)$$

This result means that the intensity of the emerging beam is modulated at a rate four times the rate of rotation of the center disk. *Suggestion:* Use the trigonometric identities $\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta)$ and $\sin^2\theta = \frac{1}{2}(1 - \cos 2\theta)$.

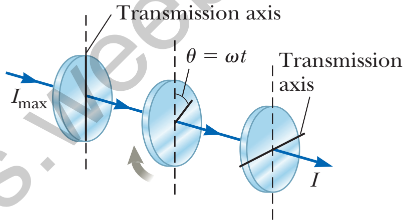


Figure P38.78

79. Consider a light wave passing through a slit and propagating toward a distant screen. Figure P38.79 shows the intensity variation for the pattern on the screen. Give a mathematical argument that more than 90% of the transmitted energy is in the central maximum of the diffraction pattern. *Suggestion:* You are not expected to calculate the precise percentage, but explain the steps of your reasoning. You may use the identification

$$\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8}$$

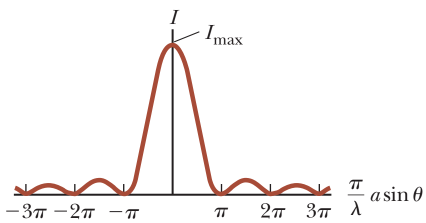
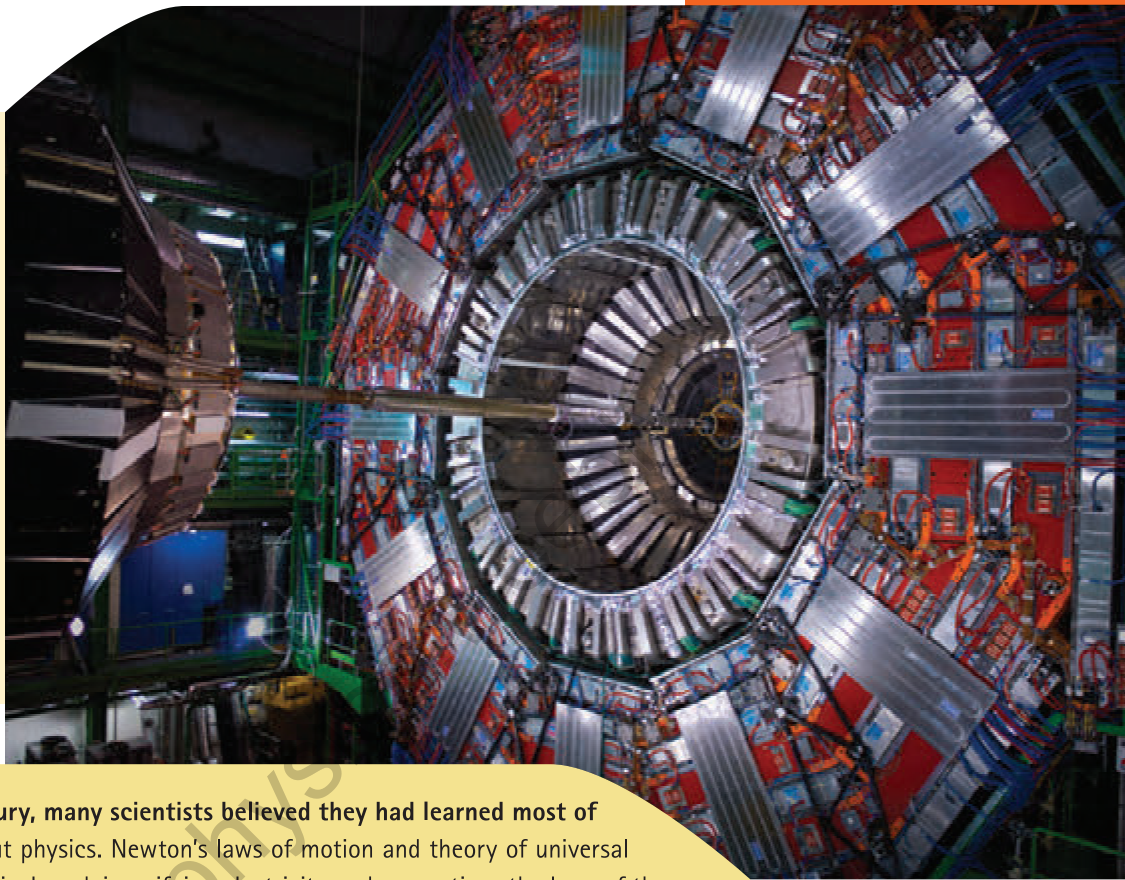


Figure P38.79

Modern Physics

PART 6



The Compact Muon Solenoid (CMS) Detector is part of the Large Hadron Collider at the European Laboratory for Particle Physics operated by CERN. It is one of several detectors that search for elementary particles. For a sense of scale, the green structure to the left of the detector and extending to the top is five stories high. *(CERN)*

At the end of the 19th century, many scientists believed they had learned most of what there was to know about physics. Newton's laws of motion and theory of universal gravitation, Maxwell's theoretical work in unifying electricity and magnetism, the laws of thermodynamics and kinetic theory, and the principles of optics were highly successful in explaining a variety of phenomena.

At the turn of the 20th century, however, a major revolution shook the world of physics. In 1900, Max Planck provided the basic ideas that led to the formulation of the quantum theory, and in 1905, Albert Einstein formulated his special theory of relativity. The excitement of the times is captured in Einstein's own words: "It was a marvelous time to be alive." Both theories were to have a profound effect on our understanding of nature. Within a few decades, they inspired new developments in the fields of atomic physics, nuclear physics, and condensed-matter physics.

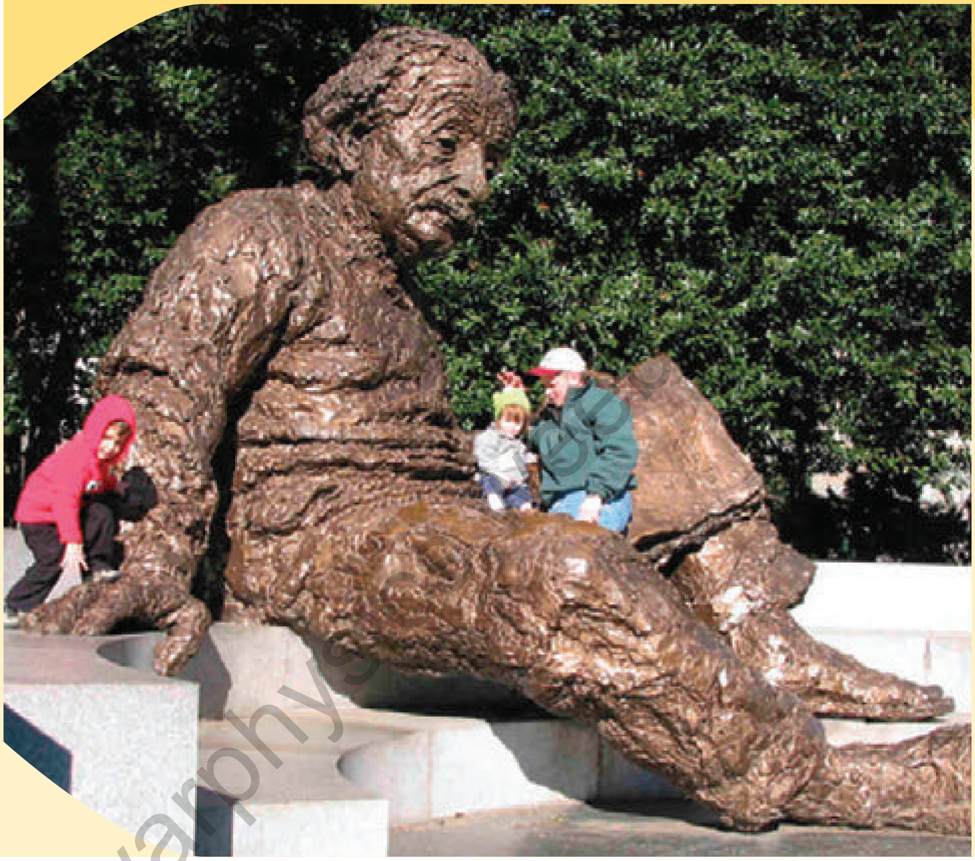
In Chapter 39, we shall introduce the special theory of relativity. The theory provides us with a new and deeper view of physical laws. Although the predictions of this theory often violate our common sense, the theory correctly describes the results of experiments involving speeds near the speed of light. The extended version of this textbook, *Physics for Scientists and Engineers with Modern Physics,* covers the basic concepts of quantum mechanics and their application to atomic and molecular physics. In addition, we introduce condensed matter physics, nuclear physics, particle physics, and cosmology in the extended version.

Even though the physics that was developed during the 20th century has led to a multitude of important technological achievements, the story is still incomplete. Discoveries will continue to evolve during our lifetimes, and many of these discoveries will deepen or refine our understanding of nature and the Universe around us. It is still a "marvelous time to be alive." ■

CHAPTER

39

Relativity



Standing on the shoulders of a giant. David Serway, son of one of the authors, watches over two of his children, Nathan and Kaitlyn, as they frolic in the arms of Albert Einstein's statue at the Einstein memorial in Washington, D.C. It is well known that Einstein, the principal architect of relativity, was very fond of children. *(Emily Serway)*

Our everyday experiences and observations involve objects that move at speeds much less than the speed of light. Newtonian mechanics was formulated by observing and describing the motion of such objects, and this formalism is very successful in describing a wide range of phenomena that occur at low speeds. Nonetheless, it fails to describe properly the motion of objects whose speeds approach that of light.

Experimentally, the predictions of Newtonian theory can be tested at high speeds by accelerating electrons or other charged particles through a large electric potential difference. For example, it is possible to accelerate an electron to a speed of $0.99c$ (where c is the speed of light) by using a potential difference of several million volts. According to Newtonian mechanics, if the potential difference is increased by a factor of 4, the electron's kinetic energy is four times greater and its speed should double to $1.98c$. Experiments show, however, that the speed of the electron—as well as the speed of any other object in the Universe—always remains less than the speed of light, regardless of the size of the accelerating voltage. Because it places no upper limit on speed, Newtonian mechanics is contrary to modern experimental results and is clearly a limited theory.

In 1905, at the age of only 26, Einstein published his special theory of relativity. Regarding the theory, Einstein wrote:

> The relativity theory arose from necessity, from serious and deep contradictions in the old theory from which there seemed no escape. The strength of the new theory lies in the consistency and simplicity with which it solves all these difficulties.[1]

Although Einstein made many other important contributions to science, the special theory of relativity alone represents one of the greatest intellectual achievements of all time. With this theory, experimental observations can be correctly predicted over the range of speeds from $v = 0$ to speeds approaching the speed of light. At low speeds, Einstein's theory reduces to Newtonian mechanics as a limiting situation. It is important to recognize that Einstein was working on electromagnetism when he developed the special theory of relativity. He was convinced that Maxwell's equations were correct, and to reconcile them with one of his postulates, he was forced into the revolutionary notion of assuming that space and time are not absolute.

This chapter gives an introduction to the special theory of relativity, with emphasis on some of its predictions. In addition to its well-known and essential role in theoretical physics, the special theory of relativity has practical applications, including the design of nuclear power plants and modern global positioning system (GPS) units. These devices depend on relativistic principles for proper design and operation.

39.1 The Principle of Galilean Relativity

To describe a physical event, we must establish a frame of reference. You should recall from Chapter 5 that an inertial frame of reference is one in which an object is observed to have no acceleration when no forces act on it. Furthermore, any frame moving with constant velocity with respect to an inertial frame must also be an inertial frame.

There is no absolute inertial reference frame. Therefore, the results of an experiment performed in a vehicle moving with uniform velocity must be identical to the results of the same experiment performed in a stationary vehicle. The formal statement of this result is called the **principle of Galilean relativity:**

The laws of mechanics must be the same in all inertial frames of reference.

◀ **Principle of Galilean relativity**

Let's consider an observation that illustrates the equivalence of the laws of mechanics in different inertial frames. The pickup truck in Figure 39.1a moves with a

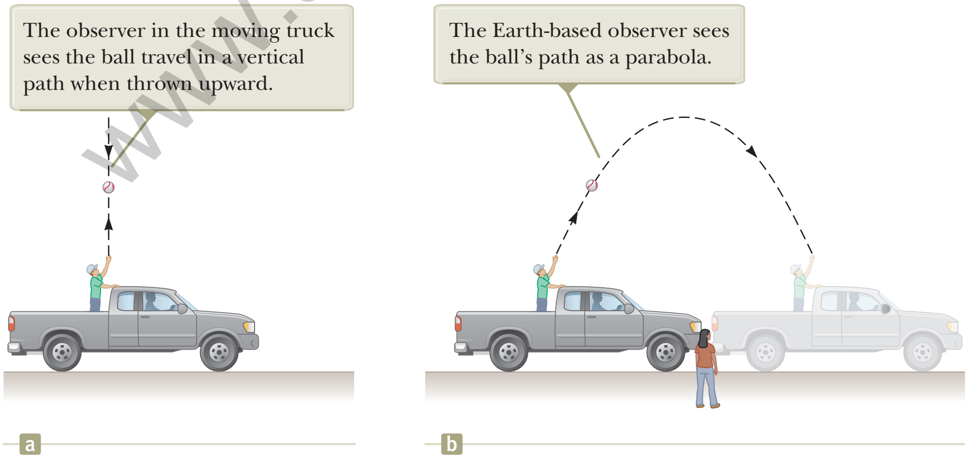


Figure 39.1 Two observers watch the path of a thrown ball and obtain different results.

[1]A. Einstein and L. Infield, *The Evolution of Physics* (New York: Simon and Schuster, 1961).

constant velocity with respect to the ground. If a passenger in the truck throws a ball straight up and if air resistance is neglected, the passenger observes that the ball moves in a vertical path. The motion of the ball appears to be precisely the same as if the ball were thrown by a person at rest on the Earth. The law of universal gravitation and the equations of motion under constant acceleration are obeyed whether the truck is at rest or in uniform motion.

Consider also an observer on the ground as in Figure 39.1b. Both observers agree on the laws of physics: the observer in the truck throws a ball straight up, and it rises and falls back into his hand according to the particle under constant acceleration model. Do the observers agree on the path of the ball thrown by the observer in the truck? The observer on the ground sees the path of the ball as a parabola as illustrated in Figure 39.1b, whereas, as mentioned earlier, the observer in the truck sees the ball move in a vertical path. Furthermore, according to the observer on the ground, the ball has a horizontal component of velocity equal to the velocity of the truck, and the horizontal motion of the ball is described by the particle under constant velocity model. Although the two observers disagree on certain aspects of the situation, they agree on the validity of Newton's laws and on the results of applying appropriate analysis models that we have learned. This agreement implies that no mechanical experiment can detect any difference between the two inertial frames. The only thing that can be detected is the relative motion of one frame with respect to the other.

Quick Quiz 39.1 Which observer in Figure 39.1 sees the ball's *correct* path? **(a)** the observer in the truck **(b)** the observer on the ground **(c)** both observers

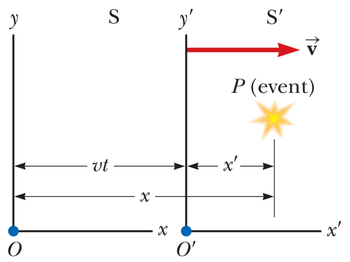


Figure 39.2 An event occurs at a point P. The event is seen by two observers in inertial frames S and S′, where S′ moves with a velocity $\vec{v}$ relative to S.

Suppose some physical phenomenon, which we call an *event,* occurs and is observed by an observer at rest in an inertial reference frame. The wording "in a frame" means that the observer is at rest with respect to the origin of that frame. The event's location and time of occurrence can be specified by the four coordinates (x, y, z, t). We would like to be able to transform these coordinates from those of an observer in one inertial frame to those of another observer in a frame moving with uniform relative velocity compared with the first frame.

Consider two inertial frames S and S′ (Fig. 39.2). The S′ frame moves with a constant velocity $\vec{v}$ along the common x and x' axes, where $\vec{v}$ is measured relative to S. We assume the origins of S and S′ coincide at $t = 0$ and an event occurs at point P in space at some instant of time. For simplicity, we show the observer O in the S frame and the observer O' in the S′ frame as blue dots at the origins of their coordinate frames in Figure 39.2, but that is not necessary: either observer could be at any fixed location in his or her frame. Observer O describes the event with space–time coordinates (x, y, z, t), whereas observer O' in S′ uses the coordinates (x', y', z', t') to describe the same event. Model the origin of S′ as a particle under constant velocity relative to the origin of S. As we see from the geometry in Figure 39.2, the relationships among these various coordinates can be written

Galilean transformation equations

$$x' = x - vt \qquad y' = y \qquad z' = z \qquad t' = t \tag{39.1}$$

These equations are the **Galilean space–time transformation equations.** Note that time is assumed to be the same in both inertial frames. That is, within the framework of classical mechanics, all clocks run at the same rate, regardless of their velocity, so the time at which an event occurs for an observer in S is the same as the time for the same event in S′. Consequently, the time interval between two successive events should be the same for both observers. Although this assumption may seem obvious, it turns out to be incorrect in situations where v is comparable to the speed of light.

Now suppose a particle moves through a displacement of magnitude dx along the x axis in a time interval dt as measured by an observer in S. It follows from Equations 39.1 that the corresponding displacement dx' measured by an observer in S′ is

$dx' = dx - v\,dt$, where frame S′ is moving with speed v in the x direction relative to frame S. Because $dt = dt'$, we find that

$$\frac{dx'}{dt'} = \frac{dx}{dt} - v$$

or

$$u'_x = u_x - v \tag{39.2}$$

where u_x and u'_x are the x components of the velocity of the particle measured by observers in S and S′, respectively. (We use the symbol $\vec{\mathbf{u}}$ rather than $\vec{\mathbf{v}}$ for particle velocity because $\vec{\mathbf{v}}$ is already used for the relative velocity of two reference frames.) Equation 39.2 is the **Galilean velocity transformation equation.** It is consistent with our intuitive notion of time and space as well as with our discussions in Section 4.6. As we shall soon see, however, it leads to serious contradictions when applied to electromagnetic waves.

Quick Quiz 39.2 A baseball pitcher with a 90-mi/h fastball throws a ball while standing on a railroad flatcar moving at 110 mi/h. The ball is thrown in the same direction as that of the velocity of the train. If you apply the Galilean velocity transformation equation to this situation, is the speed of the ball relative to the Earth **(a)** 90 mi/h, **(b)** 110 mi/h, **(c)** 20 mi/h, **(d)** 200 mi/h, or **(e)** impossible to determine?

Pitfall Prevention 39.1

The Relationship Between the S and S′ Frames Many of the mathematical representations in this chapter are true *only* for the specified relationship between the S and S′ frames. The x and x' axes coincide, except their origins are different. The y and y' axes (and the z and z' axes) are parallel, but they only coincide at one instant due to the time-varying position of the origin of S′ with respect to that of S. We choose the time $t = 0$ to be the instant at which the origins of the two coordinate systems coincide. If the S′ frame is moving in the positive x direction relative to S, then v is positive; otherwise, it is negative.

The Speed of Light

It is quite natural to ask whether the principle of Galilean relativity also applies to electricity, magnetism, and optics. Experiments indicate that the answer is no. Recall from Chapter 34 that Maxwell showed that the speed of light in free space is $c = 3.00 \times 10^8$ m/s. Physicists of the late 1800s thought light waves move through a medium called the *ether* and the speed of light is c only in a special, absolute frame at rest with respect to the ether. The Galilean velocity transformation equation was expected to hold for observations of light made by an observer in any frame moving at speed v relative to the absolute ether frame. That is, if light travels along the x axis and an observer moves with velocity $\vec{\mathbf{v}}$ along the x axis, the observer measures the light to have speed $c \pm v$, depending on the directions of travel of the observer and the light.

Because the existence of a preferred, absolute ether frame would show that light is similar to other classical waves and that Newtonian ideas of an absolute frame are true, considerable importance was attached to establishing the existence of the ether frame. Prior to the late 1800s, experiments involving light traveling in media moving at the highest laboratory speeds attainable at that time were not capable of detecting differences as small as that between c and $c \pm v$. Starting in about 1880, scientists decided to use the Earth as the moving frame in an attempt to improve their chances of detecting these small changes in the speed of light.

Observers fixed on the Earth can take the view that they are stationary and that the absolute ether frame containing the medium for light propagation moves past them with speed v. Determining the speed of light under these circumstances is similar to determining the speed of an aircraft traveling in a moving air current, or wind; consequently, we speak of an "ether wind" blowing through our apparatus fixed to the Earth.

A direct method for detecting an ether wind would use an apparatus fixed to the Earth to measure the ether wind's influence on the speed of light. If v is the speed of the ether relative to the Earth, light should have its maximum speed $c + v$ when propagating downwind as in Figure 39.3a. Likewise, the speed of light should have its minimum value $c - v$ when the light is propagating upwind as in Figure 39.3b and an intermediate value $(c^2 - v^2)^{1/2}$ when the light is directed such that it travels perpendicular to the ether wind as in Figure 39.3c. In this latter case, the vector $\vec{\mathbf{c}}$

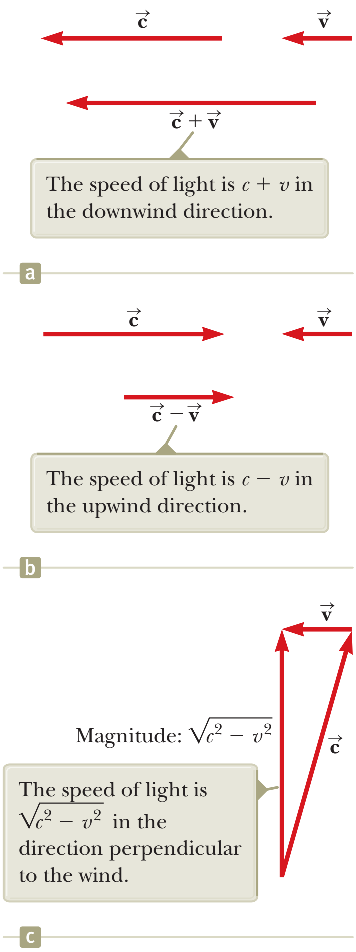


Figure 39.3 If the velocity of the ether wind relative to the Earth is $\vec{\mathbf{v}}$ and the velocity of light relative to the ether is $\vec{\mathbf{c}}$, the speed of light relative to the Earth depends on the direction of the Earth's velocity.

must be aimed upstream so that the resultant velocity is perpendicular to the wind, like the boat in Figure 4.21b. If the Sun is assumed to be at rest in the ether, the velocity of the ether wind would be equal to the orbital velocity of the Earth around the Sun, which has a magnitude of approximately 30 km/s or 3×10^4 m/s. Because $c = 3 \times 10^8$ m/s, it is necessary to detect a change in speed of approximately 1 part in 10^4 for measurements in the upwind or downwind directions. Although such a change is experimentally measurable, all attempts to detect such changes and establish the existence of the ether wind (and hence the absolute frame) proved futile! We shall discuss the classic experimental search for the ether in Section 39.2.

The principle of Galilean relativity refers only to the laws of mechanics. If it is assumed the laws of electricity and magnetism are the same in all inertial frames, a paradox concerning the speed of light immediately arises. That can be understood by recognizing that Maxwell's equations imply that the speed of light always has the fixed value 3.00×10^8 m/s in all inertial frames, a result in direct contradiction to what is expected based on the Galilean velocity transformation equation. According to Galilean relativity, the speed of light should *not* be the same in all inertial frames.

To resolve this contradiction in theories, we must conclude that either (1) the laws of electricity and magnetism are not the same in all inertial frames or (2) the Galilean velocity transformation equation is incorrect. If we assume the first alternative, a preferred reference frame in which the speed of light has the value c must exist and the measured speed must be greater or less than this value in any other reference frame, in accordance with the Galilean velocity transformation equation. If we assume the second alternative, we must abandon the notions of absolute time and absolute length that form the basis of the Galilean space–time transformation equations.

39.2 The Michelson–Morley Experiment

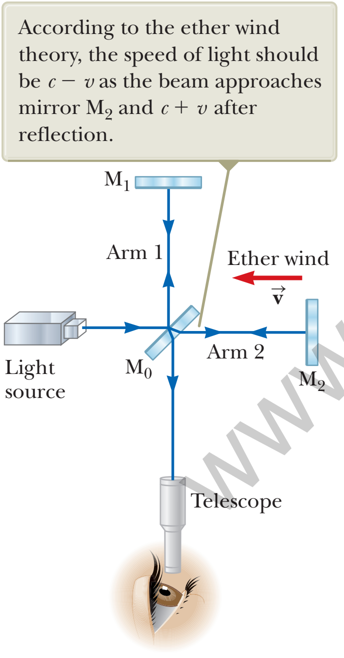


Figure 39.4 A Michelson interferometer is used in an attempt to detect the ether wind.

The most famous experiment designed to detect small changes in the speed of light was first performed in 1881 by A. A. Michelson (see Section 37.6) and later repeated under various conditions by Michelson and Edward W. Morley (1838–1923). As we shall see, the outcome of the experiment contradicted the ether hypothesis.

The experiment was designed to determine the velocity of the Earth relative to that of the hypothetical ether. The experimental tool used was the Michelson interferometer, which was discussed in Section 37.6 and is shown again in Figure 39.4. Arm 2 is aligned along the direction of the Earth's motion through space. The Earth moving through the ether at speed v is equivalent to the ether flowing past the Earth in the opposite direction with speed v. This ether wind blowing in the direction opposite the direction of the Earth's motion should cause the speed of light measured in the Earth frame to be $c - v$ as the light approaches mirror M_2 and $c + v$ after reflection, where c is the speed of light in the ether frame.

The two light beams reflect from M_1 and M_2 and recombine, and an interference pattern is formed as discussed in Section 37.6. The interference pattern is then observed while the interferometer is rotated through an angle of 90°. This rotation interchanges the speed of the ether wind between the arms of the interferometer. The rotation should cause the fringe pattern to shift slightly but measurably. Measurements failed, however, to show any change in the interference pattern! The Michelson–Morley experiment was repeated at different times of the year when the ether wind was expected to change direction and magnitude, but the results were always the same: no fringe shift of the magnitude required was *ever* observed.[2]

The negative results of the Michelson–Morley experiment not only contradicted the ether hypothesis, but also showed that it is impossible to measure the absolute

[2]From an Earth-based observer's point of view, changes in the Earth's speed and direction of motion in the course of a year are viewed as ether wind shifts. Even if the speed of the Earth with respect to the ether were zero at some time, six months later the speed of the Earth would be 60 km/s with respect to the ether and as a result a fringe shift should be noticed. No shift has ever been observed, however.

velocity of the Earth with respect to the ether frame. Einstein, however, offered a postulate for his special theory of relativity that places quite a different interpretation on these null results. In later years, when more was known about the nature of light, the idea of an ether that permeates all of space was abandoned. Light is now understood to be an electromagnetic wave, which requires no medium for its propagation. As a result, the idea of an ether in which these waves travel became unnecessary.

Details of the Michelson–Morley Experiment

To understand the outcome of the Michelson–Morley experiment, let's assume the two arms of the interferometer in Figure 39.4 are of equal length L. We shall analyze the situation as if there were an ether wind because that is what Michelson and Morley expected to find. As noted above, the speed of the light beam along arm 2 should be $c - v$ as the beam approaches M_2 and $c + v$ after the beam is reflected. We model a pulse of light as a particle under constant speed. Therefore, the time interval for travel to the right for the pulse is $\Delta t = L/(c - v)$ and the time interval for travel to the left is $\Delta t = L/(c + v)$. The total time interval for the round trip along arm 2 is

$$\Delta t_{\text{arm 2}} = \frac{L}{c+v} + \frac{L}{c-v} = \frac{2Lc}{c^2 - v^2} = \frac{2L}{c}\left(1 - \frac{v^2}{c^2}\right)^{-1}$$

Now consider the light beam traveling along arm 1, perpendicular to the ether wind. Because the speed of the beam relative to the Earth is $(c^2 - v^2)^{1/2}$ in this case (see Fig. 39.3c), the time interval for travel for each half of the trip is $\Delta t = L/(c^2 - v^2)^{1/2}$ and the total time interval for the round trip is

$$\Delta t_{\text{arm 1}} = \frac{2L}{(c^2 - v^2)^{1/2}} = \frac{2L}{c}\left(1 - \frac{v^2}{c^2}\right)^{-1/2}$$

The time difference Δt between the horizontal round trip (arm 2) and the vertical round trip (arm 1) is

$$\Delta t = \Delta t_{\text{arm 2}} - \Delta t_{\text{arm 1}} = \frac{2L}{c}\left[\left(1 - \frac{v^2}{c^2}\right)^{-1} - \left(1 - \frac{v^2}{c^2}\right)^{-1/2}\right]$$

Because $v^2/c^2 << 1$, we can simplify this expression by using the following binomial expansion after dropping all terms higher than second order:

$$(1 - x)^n \approx 1 - nx \quad (\text{for } x << 1)$$

In our case, $x = v^2/c^2$, and we find that

$$\Delta t = \Delta t_{\text{arm 2}} - \Delta t_{\text{arm 1}} \approx \frac{Lv^2}{c^3} \tag{39.3}$$

This time difference between the two instants at which the reflected beams arrive at the viewing telescope gives rise to a phase difference between the beams, producing an interference pattern when they combine at the position of the telescope. A shift in the interference pattern should be detected when the interferometer is rotated through 90° in a horizontal plane so that the two beams exchange roles. This rotation results in a time difference twice that given by Equation 39.3. Therefore, the path difference that corresponds to this time difference is

$$\Delta d = c(2\,\Delta t) = \frac{2Lv^2}{c^2}$$

Because a change in path length of one wavelength corresponds to a shift of one fringe, the corresponding fringe shift is equal to this path difference divided by the wavelength of the light:

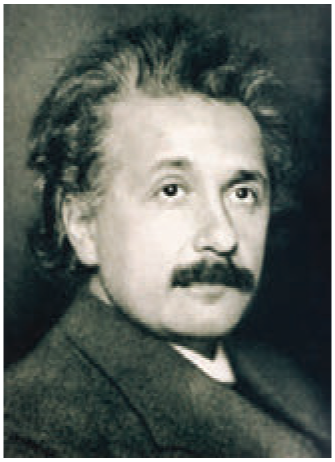
$$\text{Shift} = \frac{2Lv^2}{\lambda c^2} \tag{39.4}$$

In the experiments by Michelson and Morley, each light beam was reflected by mirrors many times to give an effective path length L of approximately 11 m. Using this value, taking v to be equal to 3.0×10^4 m/s (the speed of the Earth around the Sun), and using 500 nm for the wavelength of the light, we expect a fringe shift of

$$\text{Shift} = \frac{2(11\text{ m})(3.0 \times 10^4\text{ m/s})^2}{(5.0 \times 10^{-7}\text{ m})(3.0 \times 10^8\text{ m/s})^2} = 0.44$$

The instrument used by Michelson and Morley could detect shifts as small as 0.01 fringe, but it detected no shift whatsoever in the fringe pattern! The experiment has been repeated many times since by different scientists under a wide variety of conditions, and no fringe shift has ever been detected. Therefore, it was concluded that the motion of the Earth with respect to the postulated ether cannot be detected.

Many efforts were made to explain the null results of the Michelson–Morley experiment and to save the ether frame concept and the Galilean velocity transformation equation for light. All proposals resulting from these efforts have been shown to be wrong. No experiment in the history of physics received such valiant efforts to explain the absence of an expected result as did the Michelson–Morley experiment. The stage was set for Einstein, who solved the problem in 1905 with his special theory of relativity.



Albert Einstein
German-American Physicist (1879–1955)
Einstein, one of the greatest physicists of all time, was born in Ulm, Germany. In 1905, at age 26, he published four scientific papers that revolutionized physics. Two of these papers were concerned with what is now considered his most important contribution: the special theory of relativity.

In 1916, Einstein published his work on the general theory of relativity. The most dramatic prediction of this theory is the degree to which light is deflected by a gravitational field. Measurements made by astronomers on bright stars in the vicinity of the eclipsed Sun in 1919 confirmed Einstein's prediction, and Einstein became a world celebrity as a result. Einstein was deeply disturbed by the development of quantum mechanics in the 1920s despite his own role as a scientific revolutionary. In particular, he could never accept the probabilistic view of events in nature that is a central feature of quantum theory. The last few decades of his life were devoted to an unsuccessful search for a unified theory that would combine gravitation and electromagnetism.

39.3 Einstein's Principle of Relativity

In the previous section, we noted the impossibility of measuring the speed of the ether with respect to the Earth and the failure of the Galilean velocity transformation equation in the case of light. Einstein proposed a theory that boldly removed these difficulties and at the same time completely altered our notion of space and time.[3] He based his special theory of relativity on two postulates:

1. **The principle of relativity:** The laws of physics must be the same in all inertial reference frames.
2. **The constancy of the speed of light:** The speed of light in vacuum has the same value, $c = 3.00 \times 10^8$ m/s, in all inertial frames, regardless of the velocity of the observer or the velocity of the source emitting the light.

The first postulate asserts that *all* the laws of physics—those dealing with mechanics, electricity and magnetism, optics, thermodynamics, and so on—are the same in all reference frames moving with constant velocity relative to one another. This postulate is a generalization of the principle of Galilean relativity, which refers only to the laws of mechanics. From an experimental point of view, Einstein's principle of relativity means that any kind of experiment (measuring the speed of light, for example) performed in a laboratory at rest must give the same result when performed in a laboratory moving at a constant velocity with respect to the first one. Hence, no preferred inertial reference frame exists, and it is impossible to detect absolute motion.

Note that postulate 2 is required by postulate 1: if the speed of light were not the same in all inertial frames, measurements of different speeds would make it possible to distinguish between inertial frames. As a result, a preferred, absolute frame could be identified, in contradiction to postulate 1.

Although the Michelson–Morley experiment was performed before Einstein published his work on relativity, it is not clear whether or not Einstein was aware of the details of the experiment. Nonetheless, the null result of the experiment can be readily understood within the framework of Einstein's theory. According to

[3] A. Einstein, "On the Electrodynamics of Moving Bodies," *Ann. Physik* **17**:891, 1905. For an English translation of this article and other publications by Einstein, see the book by H. Lorentz, A. Einstein, H. Minkowski, and H. Weyl, *The Principle of Relativity* (New York: Dover, 1958).

his principle of relativity, the premises of the Michelson–Morley experiment were incorrect. In the process of trying to explain the expected results, we stated that when light traveled against the ether wind, its speed was $c - v$, in accordance with the Galilean velocity transformation equation. If the state of motion of the observer or of the source has no influence on the value found for the speed of light, however, one always measures the value to be c. Likewise, the light makes the return trip after reflection from the mirror at speed c, not at speed $c + v$. Therefore, the motion of the Earth does not influence the interference pattern observed in the Michelson–Morley experiment, and a null result should be expected.

If we accept Einstein's theory of relativity, we must conclude that relative motion is unimportant when measuring the speed of light. At the same time, we must alter our commonsense notion of space and time and be prepared for some surprising consequences. As you read the pages ahead, keep in mind that our commonsense ideas are based on a lifetime of everyday experiences and not on observations of objects moving at hundreds of thousands of kilometers per second. Therefore, these results may seem strange, but that is only because we have no experience with them.

39.4 Consequences of the Special Theory of Relativity

As we examine some of the consequences of relativity in this section, we restrict our discussion to the concepts of simultaneity, time intervals, and lengths, all three of which are quite different in relativistic mechanics from what they are in Newtonian mechanics. In relativistic mechanics, for example, the distance between two points and the time interval between two events depend on the frame of reference in which they are measured.

Simultaneity and the Relativity of Time

A basic premise of Newtonian mechanics is that a universal time scale exists that is the same for all observers. Newton and his followers took simultaneity for granted. In his special theory of relativity, Einstein abandoned this assumption.

Einstein devised the following thought experiment to illustrate this point. A boxcar moves with uniform velocity, and two bolts of lightning strike its ends as illustrated in Figure 39.5a, leaving marks on the boxcar and on the ground. The marks on the boxcar are labeled A' and B', and those on the ground are labeled A and B. An observer O' moving with the boxcar is midway between A' and B', and a ground observer O is midway between A and B. The events recorded by the observers are the striking of the boxcar by the two lightning bolts.

The light signals emitted from A and B at the instant at which the two bolts strike later reach observer O at the same time as indicated in Figure 39.5b. This observer

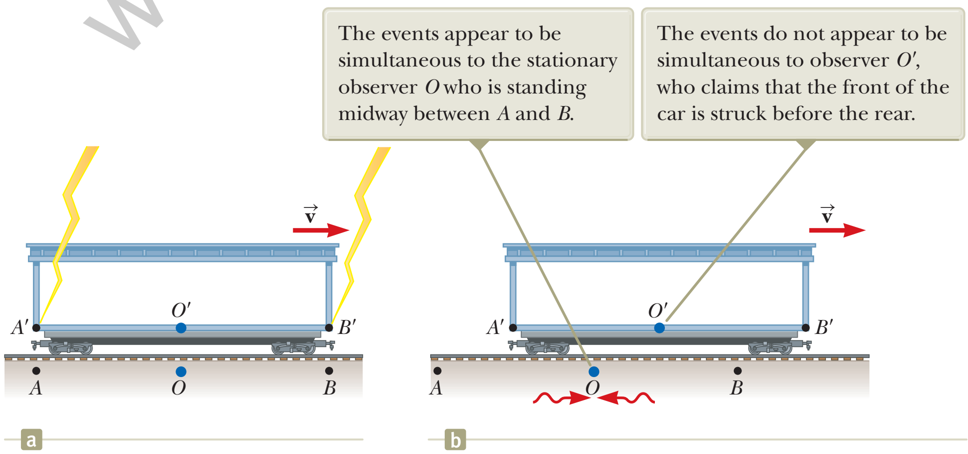


Figure 39.5 (a) Two lightning bolts strike the ends of a moving boxcar. (b) The leftward-traveling light signal has already passed O', but the rightward-traveling signal has not yet reached O'.

Pitfall Prevention 39.2

Who's Right? You might wonder which observer in Figure 39.5 is correct concerning the two lightning strikes. *Both are correct* because the principle of relativity states that *there is no preferred inertial frame of reference.* Although the two observers reach different conclusions, both are correct in their own reference frame because the concept of simultaneity is not absolute. That, in fact, is the central point of relativity: any uniformly moving frame of reference can be used to describe events and do physics.

realizes that the signals traveled at the same speed over equal distances and so concludes that the events at A and B occurred simultaneously. Now consider the same events as viewed by observer O'. By the time the signals have reached observer O, observer O' has moved as indicated in Figure 39.5b. Therefore, the signal from B' has already swept past O', but the signal from A' has not yet reached O'. In other words, O' sees the signal from B' before seeing the signal from A'. According to Einstein, *the two observers must find that light travels at the same speed*. Therefore, observer O' concludes that one lightning bolt strikes the front of the boxcar *before* the other one strikes the back.

This thought experiment clearly demonstrates that the two events that appear to be simultaneous to observer O do *not* appear to be simultaneous to observer O'. Simultaneity is not an absolute concept but rather one that depends on the state of motion of the observer. Einstein's thought experiment demonstrates that two observers can disagree on the simultaneity of two events. This disagreement, however, depends on the transit time of light to the observers and therefore does *not* demonstrate the deeper meaning of relativity. In relativistic analyses of high-speed situations, simultaneity is relative even when the transit time is subtracted out. In fact, in all the relativistic effects we discuss, we ignore differences caused by the transit time of light to the observers.

Time Dilation

To illustrate that observers in different inertial frames can measure different time intervals between a pair of events, consider a vehicle moving to the right with a speed v such as the boxcar shown in Figure 39.6a. A mirror is fixed to the ceiling of the vehicle, and observer O' at rest in the frame attached to the vehicle holds a flashlight a distance d below the mirror. At some instant, the flashlight emits a pulse of light directed toward the mirror (event 1), and at some later time after reflecting from the mirror, the pulse arrives back at the flashlight (event 2). Observer O' carries a clock and uses it to measure the time interval Δt_p between these two events. (The subscript p stands for *proper,* as we shall see in a moment.) We model the pulse of light as a particle under constant speed. Because the light pulse has a speed c, the time interval required for the pulse to travel from O' to the mirror and back is

$$\Delta t_p = \frac{\text{distance traveled}}{\text{speed}} = \frac{2d}{c} \tag{39.5}$$

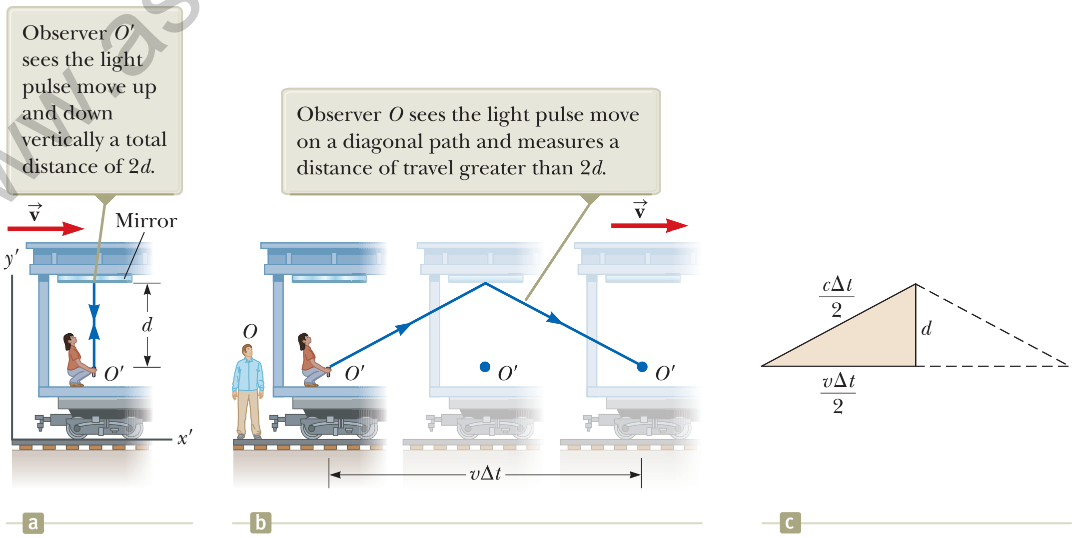


Figure 39.6 (a) A mirror is fixed to a moving vehicle, and a light pulse is sent out by observer O' at rest in the vehicle. (b) Relative to a stationary observer O standing alongside the vehicle, the mirror and O' move with a speed v. (c) The right triangle for calculating the relationship between Δt and Δt_p.

Now consider the same pair of events as viewed by observer O in a second frame at rest with respect to the ground as shown in Figure 39.6b. According to this observer, the mirror and the flashlight are moving to the right with a speed v, and as a result, the sequence of events appears entirely different. By the time the light from the flashlight reaches the mirror, the mirror has moved to the right a distance $v\,\Delta t/2$, where Δt is the time interval required for the light to travel from O' to the mirror and back to O' as measured by O. Observer O concludes that because of the motion of the vehicle, if the light is to hit the mirror, it must leave the flashlight at an angle with respect to the vertical direction. Comparing Figure 39.6a with Figure 39.6b, we see that the light must travel farther in part (b) than in part (a). (Notice that neither observer "knows" that he or she is moving. Each is at rest in his or her own inertial frame.)

According to the second postulate of the special theory of relativity, both observers must measure c for the speed of light. Because the light travels farther according to O, the time interval Δt measured by O is longer than the time interval Δt_p measured by O'. To obtain a relationship between these two time intervals, let's use the right triangle shown in Figure 39.6c. The Pythagorean theorem gives

$$\left(\frac{c\,\Delta t}{2}\right)^2 = \left(\frac{v\,\Delta t}{2}\right)^2 + d^2$$

Solving for Δt gives

$$\Delta t = \frac{2d}{\sqrt{c^2 - v^2}} = \frac{2d}{c\sqrt{1 - \dfrac{v^2}{c^2}}} \quad \textbf{(39.6)}$$

Because $\Delta t_p = 2d/c$, we can express this result as

$$\boxed{\Delta t = \frac{\Delta t_p}{\sqrt{1 - \dfrac{v^2}{c^2}}} = \gamma\,\Delta t_p} \quad \textbf{(39.7)}$$

◀ **Time dilation**

where

$$\boxed{\gamma = \frac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}}} \quad \textbf{(39.8)}$$

Because γ is always greater than unity, Equation 39.7 shows that the time interval Δt measured by an observer moving with respect to a clock is longer than the time interval Δt_p measured by an observer at rest with respect to the clock. This effect is known as **time dilation.**

Time dilation is not observed in our everyday lives, which can be understood by considering the factor γ. This factor deviates significantly from a value of 1 only for very high speeds as shown in Figure 39.7 and Table 39.1. For example, for a speed of $0.1c$, the value of γ is 1.005. Therefore, there is a time dilation of only 0.5% at

Table 39.1 Approximate Values for γ at Various Speeds

v/c	γ
0	1
0.001 0	1.000 000 5
0.010	1.000 05
0.10	1.005
0.20	1.021
0.30	1.048
0.40	1.091
0.50	1.155
0.60	1.250
0.70	1.400
0.80	1.667
0.90	2.294
0.92	2.552
0.94	2.931
0.96	3.571
0.98	5.025
0.99	7.089
0.995	10.01
0.999	22.37

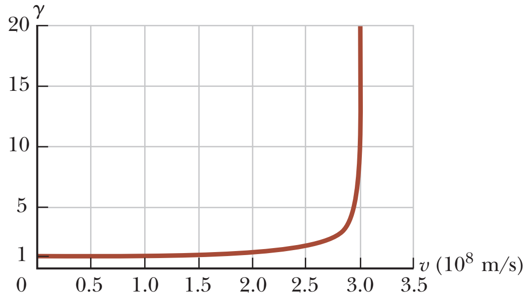


Figure 39.7 Graph of γ versus v. As the speed approaches that of light, γ increases rapidly.

one-tenth the speed of light. Speeds encountered on an everyday basis are far slower than $0.1c$, so we do not experience time dilation in normal situations.

The time interval Δt_p in Equations 39.5 and 39.7 is called the **proper time interval.** (Einstein used the German term *Eigenzeit,* which means "own-time.") In general, the proper time interval is the time interval between two events measured by an observer *who sees the events occur at the same point in space.*

If a clock is moving with respect to you, the time interval between ticks of the moving clock is observed to be longer than the time interval between ticks of an identical clock in your reference frame. Therefore, it is often said that a moving clock is measured to run more slowly than a clock in your reference frame by a factor γ. We can generalize this result by stating that all physical processes, including mechanical, chemical, and biological ones, are measured to slow down when those processes occur in a frame moving with respect to the observer. For example, the heartbeat of an astronaut moving through space keeps time with a clock inside the spacecraft. Both the astronaut's clock and heartbeat are measured to slow down relative to a clock back on the Earth (although the astronaut would have no sensation of life slowing down in the spacecraft).

Quick Quiz 39.3 Suppose the observer O' on the train in Figure 39.6 aims her flashlight at the far wall of the boxcar and turns it on and off, sending a pulse of light toward the far wall. Both O' and O measure the time interval between when the pulse leaves the flashlight and when it hits the far wall. Which observer measures the proper time interval between these two events? **(a)** O' **(b)** O **(c)** both observers **(d)** neither observer

Quick Quiz 39.4 A crew on a spacecraft watches a movie that is two hours long. The spacecraft is moving at high speed through space. Does an Earth-based observer watching the movie screen on the spacecraft through a powerful telescope measure the duration of the movie to be **(a)** longer than, **(b)** shorter than, or **(c)** equal to two hours?

Pitfall Prevention 39.3

The Proper Time Interval It is *very* important in relativistic calculations to correctly identify the observer who measures the proper time interval. The proper time interval between two events is always the time interval measured by an observer for whom the two events take place at the same position.

Time dilation is a very real phenomenon that has been verified by various experiments involving natural clocks. One experiment reported by J. C. Hafele and R. E. Keating provided direct evidence of time dilation.[4] Time intervals measured with four cesium atomic clocks in jet flight were compared with time intervals measured by Earth-based reference atomic clocks. To compare these results with theory, many factors had to be considered, including periods of speeding up and slowing down relative to the Earth, variations in direction of travel, and the weaker gravitational field experienced by the flying clocks than that experienced by the Earth-based clock. The results were in good agreement with the predictions of the special theory of relativity and were explained in terms of the relative motion between the Earth and the jet aircraft. In their paper, Hafele and Keating stated that "relative to the atomic time scale of the U.S. Naval Observatory, the flying clocks lost 59 ± 10 ns during the eastward trip and gained 273 ± 7 ns during the westward trip."

Another interesting example of time dilation involves the observation of *muons,* unstable elementary particles that have a charge equal to that of the electron and a mass 207 times that of the electron. Muons can be produced by the collision of cosmic radiation with atoms high in the atmosphere. Slow-moving muons in the laboratory have a lifetime that is measured to be the proper time interval $\Delta t_p = 2.2\ \mu$s. If we take $2.2\ \mu$s as the average lifetime of a muon and assume that muons created by cosmic radiation have a speed close to the speed of light, we find that these particles can travel a distance of approximately $(3.0 \times 10^8\ \text{m/s})(2.2 \times 10^{-6}\ \text{s}) \approx 6.6 \times 10^2$ m before they decay (Fig. 39.8a). Hence, they are unlikely to reach the

[4] J. C. Hafele and R. E. Keating, "Around the World Atomic Clocks: Relativistic Time Gains Observed," *Science* **177:**168, 1972.

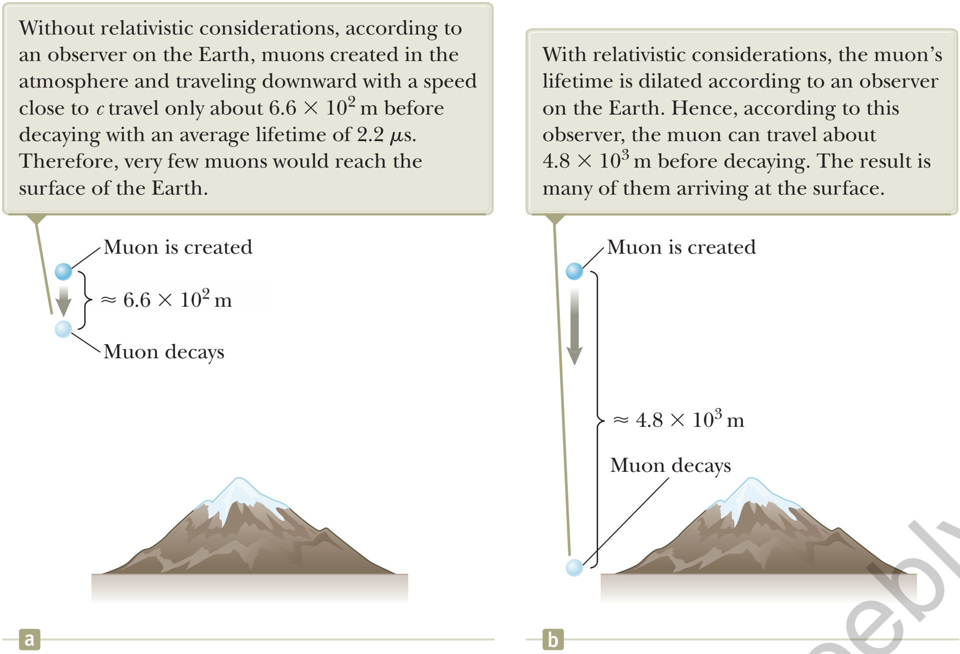


Figure 39.8 Travel of muons according to an Earth-based observer.

surface of the Earth from high in the atmosphere where they are produced. Experiments show, however, that a large number of muons *do* reach the surface. The phenomenon of time dilation explains this effect. As measured by an observer on the Earth, the muons have a dilated lifetime equal to $\gamma \Delta t_p$. For example, for $v = 0.99c$, $\gamma \approx 7.1$, and $\gamma \Delta t_p \approx 16$ μs. Hence, the average distance traveled by the muons in this time interval as measured by an observer on the Earth is approximately $(0.99)(3.0 \times 10^8 \text{ m/s})(16 \times 10^{-6} \text{ s}) \approx 4.8 \times 10^3$ m as indicated in Figure 39.8b.

In 1976, at the laboratory of the European Council for Nuclear Research (CERN) in Geneva, muons injected into a large storage ring reached speeds of approximately 0.999 4c. Electrons produced by the decaying muons were detected by counters around the ring, enabling scientists to measure the decay rate and hence the muon lifetime. The lifetime of the moving muons was measured to be approximately 30 times as long as that of the stationary muon, in agreement with the prediction of relativity to within two parts in a thousand.

Example 39.1 What Is the Period of the Pendulum?

The period of a pendulum is measured to be 3.00 s in the reference frame of the pendulum. What is the period when measured by an observer moving at a speed of 0.960c relative to the pendulum?

SOLUTION

Conceptualize Let's change frames of reference. Instead of the observer moving at 0.960c, we can take the equivalent point of view that the observer is at rest and the pendulum is moving at 0.960c past the stationary observer. Hence, the pendulum is an example of a clock moving at high speed with respect to an observer.

Categorize Based on the Conceptualize step, we can categorize this example as a substitution problem involving relativistic time dilation.

The proper time interval, measured in the rest frame of the pendulum, is $\Delta t_p = 3.00$ s.

Use Equation 39.7 to find the dilated time interval:

$$\Delta t = \gamma \Delta t_p = \frac{1}{\sqrt{1 - \dfrac{(0.960c)^2}{c^2}}} \Delta t_p = \frac{1}{\sqrt{1 - 0.921\,6}} \Delta t_p$$

$$= 3.57(3.00 \text{ s}) = \boxed{10.7 \text{ s}}$$

continued

39.1 continued

This result shows that a moving pendulum is indeed measured to take longer to complete a period than a pendulum at rest does. The period increases by a factor of $\gamma = 3.57$.

WHAT IF? What if the speed of the observer increases by 4.00%? Does the dilated time interval increase by 4.00%?

Answer Based on the highly nonlinear behavior of γ as a function of v in Figure 39.7, we would guess that the increase in Δt would be different from 4.00%.

Find the new speed if it increases by 4.00%:

$$v_{\text{new}} = (1.040\,0)(0.960c) = 0.998\,4c$$

Perform the time dilation calculation again:

$$\Delta t = \gamma\,\Delta t_p = \frac{1}{\sqrt{1 - \dfrac{(0.998\,4c)^2}{c^2}}}\,\Delta t_p = \frac{1}{\sqrt{1 - 0.996\,8}}\,\Delta t_p$$
$$= 17.68(3.00\text{ s}) = 53.1\text{ s}$$

Therefore, the 4.00% increase in speed results in almost a 400% increase in the dilated time!

Example 39.2 How Long Was Your Trip?

Suppose you are driving your car on a business trip and are traveling at 30 m/s. Your boss, who is waiting at your destination, expects the trip to take 5.0 h. When you arrive late, your excuse is that the clock in your car registered the passage of 5.0 h but that you were driving fast and so your clock ran more slowly than the clock in your boss's office. If your car clock actually did indicate a 5.0-h trip, how much time passed on your boss's clock, which was at rest on the Earth?

SOLUTION

Conceptualize The observer is your boss standing stationary on the Earth. The clock is in your car, moving at 30 m/s with respect to your boss.

Categorize The low speed of 30 m/s suggests we might categorize this problem as one in which we use classical concepts and equations. Based on the problem statement that the moving clock runs more slowly than a stationary clock, however, we categorize this problem as one involving time dilation.

Analyze The proper time interval, measured in the rest frame of the car, is $\Delta t_p = 5.0$ h.

Use Equation 39.8 to evaluate γ:

$$\gamma = \frac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}} = \frac{1}{\sqrt{1 - \dfrac{(3.0 \times 10^1\text{ m/s})^2}{(3.0 \times 10^8\text{ m/s})^2}}} = \frac{1}{\sqrt{1 - 10^{-14}}}$$

If you try to determine this value on your calculator, you will probably obtain $\gamma = 1$. Instead, perform a binomial expansion:

$$\gamma = (1 - 10^{-14})^{-1/2} \approx 1 + \tfrac{1}{2}(10^{-14}) = 1 + 5.0 \times 10^{-15}$$

Use Equation 39.7 to find the dilated time interval measured by your boss:

$$\Delta t = \gamma\,\Delta t_p = (1 + 5.0 \times 10^{-15})(5.0\text{ h})$$
$$= 5.0\text{ h} + 2.5 \times 10^{-14}\text{ h} = \boxed{5.0\text{ h} + 0.090\text{ ns}}$$

Finalize Your boss's clock would be only 0.090 ns ahead of your car clock. You might want to think of another excuse!

The Twin Paradox

An intriguing consequence of time dilation is the *twin paradox* (Fig. 39.9). Consider an experiment involving a set of twins named Speedo and Goslo. When they are 20 years old, Speedo, the more adventuresome of the two, sets out on an epic journey from the Earth to Planet X, located 20 light-years away. One light-year (ly) is the distance light travels through free space in 1 year. Furthermore, Speedo's

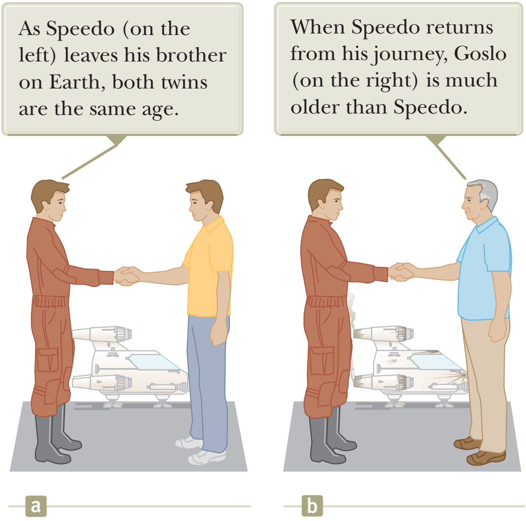


Figure 39.9 The twin paradox. Speedo takes a journey to a star 20 light-years away and returns to the Earth.

spacecraft is capable of reaching a speed of $0.95c$ relative to the inertial frame of his twin brother back home on the Earth. After reaching Planet X, Speedo becomes homesick and immediately returns to the Earth at the same speed $0.95c$. Upon his return, Speedo is shocked to discover that Goslo has aged 42 years and is now 62 years old. Speedo, on the other hand, has aged only 13 years.

The paradox is *not* that the twins have aged at different rates. Here is the apparent paradox. From Goslo's frame of reference, he was at rest while his brother traveled at a high speed away from him and then came back. According to Speedo, however, he himself remained stationary while Goslo and the Earth raced away from him and then headed back. Therefore, we might expect Speedo to claim that Goslo ages more slowly than himself. The situation appears to be symmetrical from either twin's point of view. Which twin *actually* ages more slowly?

The situation is actually not symmetrical. Consider a third observer moving at a constant speed relative to Goslo. According to the third observer, Goslo never changes inertial frames. Goslo's speed relative to the third observer is always the same. The third observer notes, however, that Speedo accelerates during his journey when he slows down and starts moving back toward the Earth, *changing reference frames in the process.* From the third observer's perspective, there is something very different about the motion of Goslo when compared to Speedo. Therefore, there is no paradox: only Goslo, who is always in a single inertial frame, can make correct predictions based on special relativity. Goslo finds that instead of aging 42 years, Speedo ages only $(1 - v^2/c^2)^{1/2}(42 \text{ years}) = 13$ years. Of these 13 years, Speedo spends 6.5 years traveling to Planet X and 6.5 years returning.

Quick Quiz 39.5 Suppose astronauts are paid according to the amount of time they spend traveling in space. After a long voyage traveling at a speed approaching c, would a crew rather be paid according to **(a)** an Earth-based clock, **(b)** their spacecraft's clock, or **(c)** either clock?

Length Contraction

The measured distance between two points in space also depends on the frame of reference of the observer. The **proper length** L_p of an object is the length measured by an observer *at rest relative to the object.* The length of an object measured by someone in a reference frame that is moving with respect to the object is always less than the proper length. This effect is known as **length contraction.**

To understand length contraction, consider a spacecraft traveling with a speed v from one star to another. There are two observers: one on the Earth and the other in the spacecraft. The observer at rest on the Earth (and also assumed to be at rest with respect to the two stars) measures the distance between the stars to be the proper length L_p. According to this observer, the time interval required for the spacecraft to complete the voyage is given by the particle under constant velocity model as $\Delta t = L_p/v$. The passages of the two stars by the spacecraft occur at the same position for the space traveler. Therefore, the space traveler measures the proper time interval Δt_p. Because of time dilation, the proper time interval is related to the Earth-measured time interval by $\Delta t_p = \Delta t/\gamma$. Because the space traveler reaches the second star in the time Δt_p, he or she concludes that the distance L between the stars is

$$L = v\,\Delta t_p = v\,\frac{\Delta t}{\gamma}$$

Because the proper length is $L_p = v\,\Delta t$, we see that

Length contraction

$$L = \frac{L_p}{\gamma} = L_p\sqrt{1 - \frac{v^2}{c^2}} \qquad \textbf{(39.9)}$$

where $\sqrt{1 - v^2/c^2}$ is a factor less than unity. If an object has a proper length L_p when it is measured by an observer at rest with respect to the object, its length L when it moves with speed v in a direction parallel to its length is measured to be shorter according to Equation 39.9.

For example, suppose a meterstick moves past a stationary Earth-based observer with speed v as in Figure 39.10. The length of the meterstick as measured by an observer in a frame attached to the stick is the proper length L_p shown in Figure 39.10a. The length of the stick L measured by the Earth observer is shorter than L_p by the factor $(1 - v^2/c^2)^{1/2}$ as suggested in Figure 39.10b. Notice that length contraction takes place only along the direction of motion.

The proper length and the proper time interval are defined differently. The proper length is measured by an observer for whom the endpoints of the length remain fixed in space. The proper time interval is measured by someone for whom the two events take place at the same position in space. As an example of this point, let's return to the decaying muons moving at speeds close to the speed of light. An observer in the muon's reference frame measures the proper lifetime, whereas an Earth-based observer measures the proper length (the distance between the creation point and the decay point in Fig. 39.8b). In the muon's reference frame, there is no time dilation, but the distance of travel to the surface is shorter when measured in this frame. Likewise, in the Earth observer's reference frame, there is time dilation, but the distance of travel is measured to be the proper length. Therefore, when calculations on the muon are performed in both frames, the outcome of the experiment in one frame is the same as the outcome in the other frame: more muons reach the surface than would be predicted without relativistic effects.

Quick Quiz 39.6 You are packing for a trip to another star. During the journey, you will be traveling at 0.99c. You are trying to decide whether you should buy smaller sizes of your clothing because you will be thinner on your trip due to length contraction. You also plan to save money by reserving a smaller cabin to sleep in because you will be shorter when you lie down. Should you **(a)** buy smaller sizes of clothing, **(b)** reserve a smaller cabin, **(c)** do neither of these things, or **(d)** do both of these things?

Quick Quiz 39.7 You are observing a spacecraft moving away from you. You measure it to be shorter than when it was at rest on the ground next to you. You also see a clock through the spacecraft window, and you observe that the passage of time on the clock is measured to be slower than that of the watch on your wrist. Compared with when the spacecraft was on the ground, what do you measure if the spacecraft turns around and comes *toward* you at the same speed? **(a)** The spacecraft is measured to be longer, and the clock runs faster. **(b)** The spacecraft is measured to be longer, and the clock runs slower. **(c)** The spacecraft is

Pitfall Prevention 39.4

The Proper Length As with the proper time interval, it is *very* important in relativistic calculations to correctly identify the observer who measures the proper length. The proper length between two points in space is always the length measured by an observer at rest with respect to the points. Often, the proper time interval and the proper length are *not* measured by the same observer.

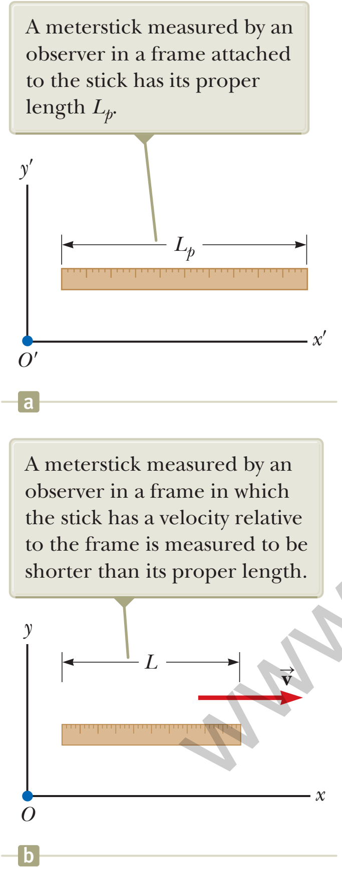


Figure 39.10 The length of a meterstick is measured by two observers.

measured to be shorter, and the clock runs faster. **(d)** The spacecraft is measured to be shorter, and the clock runs slower.

Space–Time Graphs

It is sometimes helpful to represent a physical situation with a **space–time graph,** in which ct is the ordinate and position x is the abscissa. The twin paradox is displayed in such a graph in Figure 39.11 from Goslo's point of view. A path through space–time is called a **world-line.** At the origin, the world-lines of Speedo (blue) and Goslo (green) coincide because the twins are in the same location at the same time. After Speedo leaves on his trip, his world-line diverges from that of his brother. Goslo's world-line is vertical because he remains fixed in location. At Goslo and Speedo's reunion, the two world-lines again come together. It would be impossible for Speedo to have a world-line that crossed the path of a light beam that left the Earth when he did. To do so would require him to have a speed greater than c (which, as shown in Sections 39.6 and 39.7, is not possible).

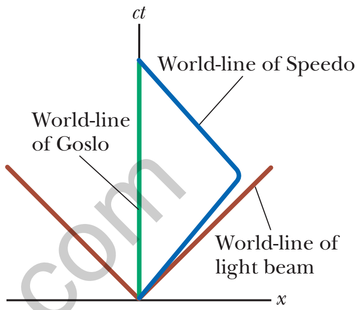


Figure 39.11 The twin paradox on a space–time graph. The twin who stays on the Earth has a world-line along the ct axis (green). The path of the traveling twin through space–time is represented by a world-line that changes direction (blue). The red-brown lines are world-lines for light beams traveling in the positive x direction (on the right) or the negative x direction (on the left).

World-lines for light beams are diagonal lines on space–time graphs, typically drawn at 45° to the right or left of vertical (assuming the x and ct axes have the same scales), depending on whether the light beam is traveling in the direction of increasing or decreasing x. All possible future events for Goslo and Speedo lie above the x axis and between the red-brown lines in Figure 39.11 because neither twin can travel faster than light. The only past events that Goslo and Speedo could have experienced occur between two similar 45° world-lines that approach the origin from below the x axis.

If Figure 39.11 is rotated about the ct axis, the red-brown lines sweep out a cone, called the *light cone,* which generalizes Figure 39.11 to two space dimensions. The y axis can be imagined coming out of the page. All future events for an observer at the origin must lie within the light cone. We can imagine another rotation that would generalize the light cone to three space dimensions to include z, but because of the requirement for four dimensions (three space dimensions and time), we cannot represent this situation in a two-dimensional drawing on paper.

Example 39.3 A Voyage to Sirius AM

An astronaut takes a trip to Sirius, which is located a distance of 8 light-years from the Earth. The astronaut measures the time of the one-way journey to be 6 years. If the spaceship moves at a constant speed of $0.8c$, how can the 8-ly distance be reconciled with the 6-year trip time measured by the astronaut?

SOLUTION

Conceptualize An observer on the Earth measures light to require 8 years to travel between Sirius and the Earth. The astronaut measures a time interval for his travel of only 6 years. Is the astronaut traveling faster than light?

Categorize Because the astronaut is measuring a length of space between the Earth and Sirius that is in motion with respect to her, we categorize this example as a length contraction problem. We also model the astronaut as a *particle under constant velocity.*

Analyze The distance of 8 ly represents the proper length from the Earth to Sirius measured by an observer on the Earth seeing both objects nearly at rest.

Calculate the contracted length measured by the astronaut using Equation 39.9:

$$L = \frac{8\text{ ly}}{\gamma} = (8\text{ ly})\sqrt{1 - \frac{v^2}{c^2}} = (8\text{ ly})\sqrt{1 - \frac{(0.8c)^2}{c^2}} = 5\text{ ly}$$

Use the particle under constant velocity model to find the travel time measured on the astronaut's clock:

$$\Delta t = \frac{L}{v} = \frac{5\text{ ly}}{0.8c} = \frac{5\text{ ly}}{0.8(1\text{ ly/yr})} = 6\text{ yr}$$

continued

39.3 continued

Finalize Notice that we have used the value for the speed of light as $c = 1$ ly/yr. The trip takes a time interval shorter than 8 years for the astronaut because, to her, the distance between the Earth and Sirius is measured to be shorter.

WHAT IF? What if this trip is observed with a very powerful telescope by a technician in Mission Control on the Earth? At what time will this technician *see* that the astronaut has arrived at Sirius?

Answer The time interval the technician measures for the astronaut to arrive is

$$\Delta t = \frac{L_p}{v} = \frac{8 \text{ ly}}{0.8c} = 10 \text{ yr}$$

For the technician to *see* the arrival, the light from the scene of the arrival must travel back to the Earth and enter the telescope. This travel requires a time interval of

$$\Delta t = \frac{L_p}{v} = \frac{8 \text{ ly}}{c} = 8 \text{ yr}$$

Therefore, the technician sees the arrival after 10 yr + 8 yr = 18 yr. If the astronaut immediately turns around and comes back home, she arrives, according to the technician, 20 years after leaving, only 2 years *after the technician saw her arrive!* In addition, the astronaut would have aged by only 12 years.

Example 39.4 The Pole-in-the-Barn Paradox AM

The twin paradox, discussed earlier, is a classic "paradox" in relativity. Another classic "paradox" is as follows. Suppose a runner moving at $0.75c$ carries a horizontal pole 15 m long toward a barn that is 10 m long. The barn has front and rear doors that are initially open. An observer on the ground can instantly and simultaneously close and open the two doors by remote control. When the runner and the pole are inside the barn, the ground observer closes and then opens both doors so that the runner and pole are momentarily captured inside the barn and then proceed to exit the barn from the back doorway. Do both the runner and the ground observer agree that the runner makes it safely through the barn?

SOLUTION

Conceptualize From your everyday experience, you would be surprised to see a 15-m pole fit inside a 10-m barn, but we are becoming used to surprising results in relativistic situations.

Categorize The pole is in motion with respect to the ground observer so that the observer measures its length to be contracted, whereas the stationary barn has a proper length of 10 m. We categorize this example as a length contraction problem. The runner carrying the pole is modeled as a *particle under constant velocity*.

Analyze Use Equation 39.9 to find the contracted length of the pole according to the ground observer:

$$L_{\text{pole}} = L_p\sqrt{1 - \frac{v^2}{c^2}} = (15 \text{ m})\sqrt{1 - (0.75)^2} = 9.9 \text{ m}$$

Therefore, the ground observer measures the pole to be slightly shorter than the barn and there is no problem with momentarily capturing the pole inside it. The "paradox" arises when we consider the runner's point of view.

Use Equation 39.9 to find the contracted length of the barn according to the running observer:

$$L_{\text{barn}} = L_p\sqrt{1 - \frac{v^2}{c^2}} = (10 \text{ m})\sqrt{1 - (0.75)^2} = 6.6 \text{ m}$$

Because the pole is in the rest frame of the runner, the runner measures it to have its proper length of 15 m. Now the situation looks even worse: How can a 15-m pole fit inside a *6.6-m* barn? Although this question is the classic one that is often asked, it is not the question we have asked because it is not the important one. We asked, "*Does the runner make it safely through the barn?*"

The resolution of the "paradox" lies in the relativity of simultaneity. The closing of the two doors is measured to be simultaneous by the ground observer. Because the doors are at different positions, however, they do not close simulta-

39.4 continued

neously as measured by the runner. The rear door closes and then opens first, allowing the leading end of the pole to exit. The front door of the barn does not close until the trailing end of the pole passes by.

We can analyze this "paradox" using a space–time graph. Figure 39.12a is a space–time graph from the ground observer's point of view. We choose $x = 0$ as the position of the front doorway of the barn and $t = 0$ as the instant at which the leading end of the pole is located at the front doorway of the barn. The world-lines for the two doorways of the barn are separated by 10 m and are vertical because the barn is not moving relative to this observer. For the pole, we follow two tilted world-lines, one for each end of the moving pole. These world-lines are 9.9 m apart horizontally, which is the contracted length seen by the ground observer. As seen in Figure 39.12a, the pole is entirely within the barn at some time.

Figure 39.12b shows the space–time graph according to the runner. Here, the world-lines for the pole are separated by 15 m and are vertical because the pole is at rest in the runner's frame of reference. The barn is hurtling *toward* the runner, so the world-lines for the front and rear doorways of the barn are tilted to the left. The world-lines for the barn are separated by 6.6 m, the contracted length as seen by the runner. The leading end of the pole leaves the rear doorway of the barn long before the trailing end of the pole enters the barn. Therefore, the opening of the rear door occurs before the closing of the front door.

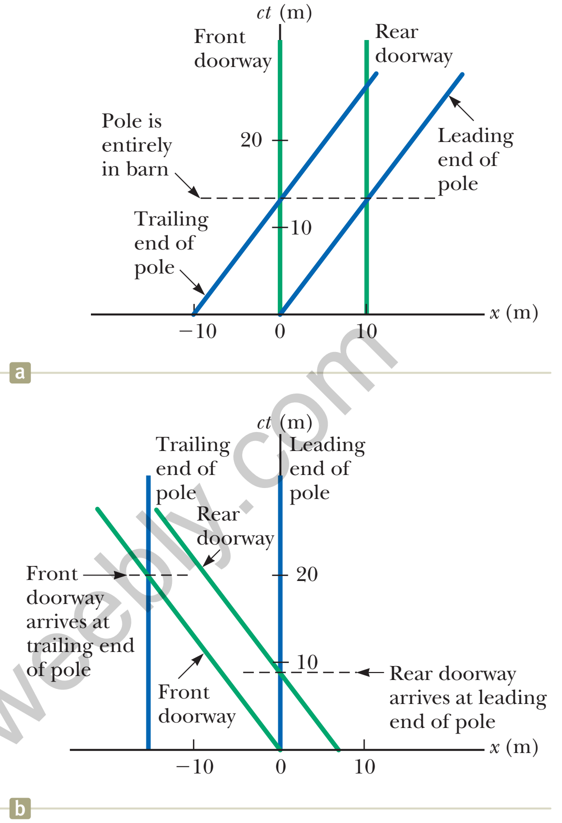


Figure 39.12 (Example 39.4) Space–time graphs for the pole-in-the-barn paradox (a) from the ground observer's point of view and (b) from the runner's point of view.

From the ground observer's point of view, use the particle under constant velocity model to find the time after $t = 0$ at which the trailing end of the pole enters the barn:

$$(1)\quad t = \frac{\Delta x}{v} = \frac{9.9\text{ m}}{0.75c} = \frac{13.2\text{ m}}{c}$$

From the runner's point of view, use the particle under constant velocity model to find the time at which the leading end of the pole leaves the barn:

$$(2)\quad t = \frac{\Delta x}{v} = \frac{6.6\text{ m}}{0.75c} = \frac{8.8\text{ m}}{c}$$

Find the time at which the trailing end of the pole enters the front door of the barn:

$$(3)\quad t = \frac{\Delta x}{v} = \frac{15\text{ m}}{0.75c} = \frac{20\text{ m}}{c}$$

Finalize From Equation (1), the pole should be completely inside the barn at a time corresponding to $ct = 13.2$ m. This situation is consistent with the point on the ct axis in Figure 39.12a where the pole is inside the barn. From Equation (2), the leading end of the pole leaves the barn at $ct = 8.8$ m. This situation is consistent with the point on the ct axis in Figure 39.12b where the rear doorway of the barn arrives at the leading end of the pole. Equation (3) gives $ct =$ 20 m, which agrees with the instant shown in Figure 39.12b at which the front doorway of the barn arrives at the trailing end of the pole.

The Relativistic Doppler Effect

Another important consequence of time dilation is the shift in frequency observed for light emitted by atoms in motion as opposed to light emitted by atoms at rest. This phenomenon, known as the Doppler effect, was introduced in Chapter 17 as it pertains to sound waves. In the case of sound, the velocity v_S of the source with respect to the medium of propagation can be distinguished from the velocity v_O of the observer with respect to the medium (the air). Light waves must be analyzed differently, however, because *they require no medium of propagation,* and no method exists for distinguishing the velocity of a light source from the velocity of the observer. The only measurable velocity is the *relative velocity* v between the source and the observer.

If a light source and an observer approach each other with a relative speed v, the frequency f' measured by the observer is

$$f' = \frac{\sqrt{1 + v/c}}{\sqrt{1 - v/c}}\, f \tag{39.10}$$

where f is the frequency of the source measured in its rest frame. This relativistic Doppler shift equation, unlike the Doppler shift equation for sound, depends only on the relative speed v of the source and observer and holds for relative speeds as great as c. As you might expect, the equation predicts that $f' > f$ when the source and observer approach each other. We obtain the expression for the case in which the source and observer recede from each other by substituting negative values for v in Equation 39.10.

The most spectacular and dramatic use of the relativistic Doppler effect is the measurement of shifts in the frequency of light emitted by a moving astronomical object such as a galaxy. Light emitted by atoms and normally found in the extreme violet region of the spectrum is shifted toward the red end of the spectrum for atoms in other galaxies, indicating that these galaxies are *receding* from us. American astronomer Edwin Hubble (1889–1953) performed extensive measurements of this *red shift* to confirm that most galaxies are moving away from us, indicating that the Universe is expanding.

39.5 The Lorentz Transformation Equations

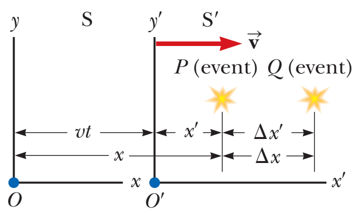


Figure 39.13 Events occur at points P and Q and are observed by an observer at rest in the S frame and another in the S′ frame, which is moving to the right with a speed v.

Suppose two events occur at points P and Q and are reported by two observers, one at rest in a frame S and another in a frame S′ that is moving to the right with speed v as in Figure 39.13. The observer in S reports the events with space–time coordinates (x, y, z, t), and the observer in S′ reports the same events using the coordinates (x', y', z', t'). Equation 39.1 predicts that the distance between the two points in space at which the events occur does not depend on motion of the observer: $\Delta x = \Delta x'$. Because this prediction is contradictory to the notion of length contraction, the Galilean transformation is not valid when v approaches the speed of light. In this section, we present the correct transformation equations that apply for all speeds in the range $0 < v < c$.

The equations that are valid for all speeds and that enable us to transform coordinates from S to S′ are the **Lorentz transformation equations:**

Lorentz transformation for S → S′

$$x' = \gamma(x - vt) \qquad y' = y \qquad z' = z \qquad t' = \gamma\left(t - \frac{v}{c^2}\,x\right) \tag{39.11}$$

These transformation equations were developed by Hendrik A. Lorentz (1853–1928) in 1890 in connection with electromagnetism. It was Einstein, however, who recognized their physical significance and took the bold step of interpreting them within the framework of the special theory of relativity.

Notice the difference between the Galilean and Lorentz time equations. In the Galilean case, $t = t'$. In the Lorentz case, however, the value for t' assigned to an event by an observer O' in the S′ frame in Figure 39.13 depends both on the time t and on the coordinate x as measured by an observer O in the S frame, which is consistent with the notion that an event is characterized by four space–time coordinates (x, y, z, t). In other words, in relativity, space and time are *not* separate concepts but rather are closely interwoven with each other.

If you wish to transform coordinates in the S′ frame to coordinates in the S frame, simply replace v by $-v$ and interchange the primed and unprimed coordinates in Equations 39.11:

$$x = \gamma(x' + vt') \qquad y = y' \qquad z = z' \qquad t = \gamma\left(t' + \frac{v}{c^2}x'\right) \tag{39.12}$$

◀ **Inverse Lorentz transformation for S′ → S**

When $v << c$, the Lorentz transformation equations should reduce to the Galilean equations. As v approaches zero, $v/c << 1$; therefore, $\gamma \rightarrow 1$ and Equations 39.11 indeed reduce to the Galilean space–time transformation equations in Equation 39.1.

In many situations, we would like to know the difference in coordinates between two events or the time interval between two events as seen by observers O and O'. From Equations 39.11 and 39.12, we can express the differences between the four variables x, x', t, and t' in the form

$$\left.\begin{aligned}\Delta x' &= \gamma(\Delta x - v\,\Delta t)\\ \Delta t' &= \gamma\left(\Delta t - \frac{v}{c^2}\Delta x\right)\end{aligned}\right\}\text{S} \rightarrow \text{S}' \tag{39.13}$$

$$\left.\begin{aligned}\Delta x &= \gamma(\Delta x' + v\,\Delta t')\\ \Delta t &= \gamma\left(\Delta t' + \frac{v}{c^2}\Delta x'\right)\end{aligned}\right\}\text{S}' \rightarrow \text{S} \tag{39.14}$$

where $\Delta x' = x'_2 - x'_1$ and $\Delta t' = t'_2 - t'_1$ are the differences measured by observer O' and $\Delta x = x_2 - x_1$ and $\Delta t = t_2 - t_1$ are the differences measured by observer O. (We have not included the expressions for relating the y and z coordinates because they are unaffected by motion along the x direction.[5])

Example 39.5 Simultaneity and Time Dilation Revisited

(A) Use the Lorentz transformation equations in difference form to show that simultaneity is not an absolute concept.

SOLUTION

Conceptualize Imagine two events that are simultaneous and separated in space as measured in the S′ frame such that $\Delta t' = 0$ and $\Delta x' \neq 0$. These measurements are made by an observer O' who is moving with speed v relative to O.

Categorize The statement of the problem tells us to categorize this example as one involving the use of the Lorentz transformation.

Analyze From the expression for Δt given in Equation 39.14, find the time interval Δt measured by observer O:

$$\Delta t = \gamma\left(\Delta t' + \frac{v}{c^2}\Delta x'\right) = \gamma\left(0 + \frac{v}{c^2}\Delta x'\right) = \gamma\frac{v}{c^2}\Delta x'$$

Finalize The time interval for the same two events as measured by O is nonzero, so the events do not appear to be simultaneous to O.

(B) Use the Lorentz transformation equations in difference form to show that a moving clock is measured to run more slowly than a clock that is at rest with respect to an observer.

SOLUTION

Conceptualize Imagine that observer O' carries a clock that he uses to measure a time interval $\Delta t'$. He finds that two events occur at the same place in his reference frame ($\Delta x' = 0$) but at different times ($\Delta t' \neq 0$). Observer O' is moving with speed v relative to O.

continued

[5]Although relative motion of the two frames along the x axis does not change the y and z coordinates of an object, it does change the y and z velocity components of an object moving in either frame as noted in Section 39.6.

▶ 39.5 continued

Categorize The statement of the problem tells us to categorize this example as one involving the use of the Lorentz transformation.

Analyze From the expression for Δt given in Equation 39.14, find the time interval Δt measured by observer O:

$$\Delta t = \gamma\left(\Delta t' + \frac{v}{c^2}\Delta x'\right) = \gamma\left[\Delta t' + \frac{v}{c^2}(0)\right] = \gamma\,\Delta t'$$

Finalize This result is the equation for time dilation found earlier (Eq. 39.7), where $\Delta t' = \Delta t_p$ is the proper time interval measured by the clock carried by observer O'. Therefore, O measures the moving clock to run slow.

39.6 The Lorentz Velocity Transformation Equations

Suppose two observers in relative motion with respect to each other are both observing an object's motion. Previously, we defined an event as occurring at an instant of time. Now let's interpret the "event" as the object's motion. We know that the Galilean velocity transformation (Eq. 39.2) is valid for low speeds. How do the observers' measurements of the velocity of the object relate to each other if the speed of the object or the relative speed of the observers is close to that of light? Once again, S′ is our frame moving at a speed v relative to S. Suppose an object has a velocity component u'_x measured in the S′ frame, where

$$u'_x = \frac{dx'}{dt'} \qquad \textbf{(39.15)}$$

Using Equation 39.11, we have

$$dx' = \gamma(dx - v\,dt)$$

$$dt' = \gamma\left(dt - \frac{v}{c^2}\,dx\right)$$

Substituting these values into Equation 39.15 gives

$$u'_x = \frac{dx - v\,dt}{dt - \dfrac{v}{c^2}\,dx} = \frac{\dfrac{dx}{dt} - v}{1 - \dfrac{v}{c^2}\dfrac{dx}{dt}}$$

The term dx/dt, however, is simply the velocity component u_x of the object measured by an observer in S, so this expression becomes

Lorentz velocity transformation for S → S′ ▶

$$\boxed{u'_x = \frac{u_x - v}{1 - \dfrac{u_x v}{c^2}}} \qquad \textbf{(39.16)}$$

If the object has velocity components along the y and z axes, the components as measured by an observer in S′ are

$$u'_y = \frac{u_y}{\gamma\left(1 - \dfrac{u_x v}{c^2}\right)} \quad \text{and} \quad u'_z = \frac{u_z}{\gamma\left(1 - \dfrac{u_x v}{c^2}\right)} \qquad \textbf{(39.17)}$$

Notice that u'_y and u'_z do not contain the parameter v in the numerator because the relative velocity is along the x axis.

When v is much smaller than c (the nonrelativistic case), the denominator of Equation 39.16 approaches unity and so $u'_x \approx u_x - v$, which is the Galilean velocity transformation equation. In another extreme, when $u_x = c$, Equation 39.16 becomes

$$u'_x = \frac{c - v}{1 - \dfrac{cv}{c^2}} = \frac{c\left(1 - \dfrac{v}{c}\right)}{1 - \dfrac{v}{c}} = c$$

This result shows that a speed measured as c by an observer in S is also measured as c by an observer in S′, independent of the relative motion of S and S′. This conclusion is consistent with Einstein's second postulate: the speed of light must be c relative to all inertial reference frames. Furthermore, we find that the speed of an object can never be measured as larger than c. That is, the speed of light is the ultimate speed. We shall return to this point later.

To obtain u_x in terms of u'_x, we replace v by $-v$ in Equation 39.16 and interchange the roles of u_x and u'_x:

$$u_x = \frac{u'_x + v}{1 + \dfrac{u'_x v}{c^2}} \qquad \textbf{(39.18)}$$

Quick Quiz 39.8 You are driving on a freeway at a relativistic speed. **(i)** Straight ahead of you, a technician standing on the ground turns on a searchlight and a beam of light moves exactly vertically upward as seen by the technician. As you observe the beam of light, do you measure the magnitude of the vertical component of its velocity as (a) equal to c, (b) greater than c, or (c) less than c? **(ii)** If the technician aims the searchlight directly at you instead of upward, do you measure the magnitude of the horizontal component of its velocity as (a) equal to c, (b) greater than c, or (c) less than c?

Pitfall Prevention 39.5

What Can the Observers Agree On? We have seen several measurements that the two observers O and O' do *not* agree on: (1) the time interval between events that take place in the same position in one of their frames, (2) the distance between two points that remain fixed in one of their frames, (3) the velocity components of a moving particle, and (4) whether two events occurring at different locations in both frames are simultaneous or not. The two observers *can* agree on (1) their relative speed of motion v with respect to each other, (2) the speed c of any ray of light, and (3) the simultaneity of two events that take place at the same position *and* time in some frame.

Example 39.6 Relative Velocity of Two Spacecraft

Two spacecraft A and B are moving in opposite directions as shown in Figure 39.14. An observer on the Earth measures the speed of spacecraft A to be $0.750c$ and the speed of spacecraft B to be $0.850c$. Find the velocity of spacecraft B as observed by the crew on spacecraft A.

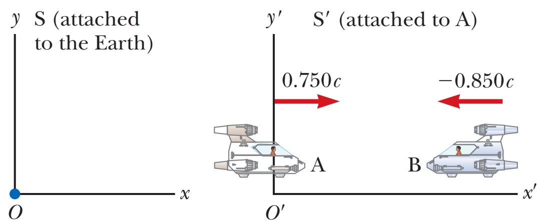


Figure 39.14 (Example 39.6) Two spacecraft A and B move in opposite directions. The speed of spacecraft B relative to spacecraft A is *less* than c and is obtained from the relativistic velocity transformation equation.

SOLUTION

Conceptualize There are two observers, one (O) on the Earth and one (O') on spacecraft A. The event is the motion of spacecraft B.

Categorize Because the problem asks to find an observed velocity, we categorize this example as one requiring the Lorentz velocity transformation.

Analyze The Earth-based observer at rest in the S frame makes two measurements, one of each spacecraft. We want to find the velocity of spacecraft B as measured by the crew on spacecraft A. Therefore, $u_x = -0.850c$. The velocity of spacecraft A is also the velocity of the observer at rest in spacecraft A (the S′ frame) relative to the observer at rest on the Earth. Therefore, $v = 0.750c$.

Obtain the velocity u'_x of spacecraft B relative to spacecraft A using Equation 39.16:

$$u'_x = \frac{u_x - v}{1 - \dfrac{u_x v}{c^2}} = \frac{-0.850c - 0.750c}{1 - \dfrac{(-0.850c)(0.750c)}{c^2}} = \boxed{-0.977c}$$

Finalize The negative sign indicates that spacecraft B is moving in the negative x direction as observed by the crew on spacecraft A. Is that consistent with your expectation from Figure 39.14? Notice that the speed is less than c. That is, an

continued

39.6 continued

object whose speed is less than c in one frame of reference must have a speed less than c in any other frame. (Had you used the Galilean velocity transformation equation in this example, you would have found that $u'_x = u_x - v = -0.850c - 0.750c = -1.60c$, which is impossible. The Galilean transformation equation does not work in relativistic situations.)

WHAT IF? What if the two spacecraft pass each other? What is their relative speed now?

Answer The calculation using Equation 39.16 involves only the velocities of the two spacecraft and does not depend on their locations. After they pass each other, they have the same velocities, so the velocity of spacecraft B as observed by the crew on spacecraft A is the same, $-0.977c$. The only difference after they pass is that spacecraft B is receding from spacecraft A, whereas it was approaching spacecraft A before it passed.

Example 39.7 Relativistic Leaders of the Pack

Two motorcycle pack leaders named David and Emily are racing at relativistic speeds along perpendicular paths as shown in Figure 39.15. How fast does Emily recede as seen by David over his right shoulder?

SOLUTION

Conceptualize The two observers are David and the police officer in Figure 39.15. The event is the motion of Emily. Figure 39.15 represents the situation as seen by the police officer at rest in frame S. Frame S′ moves along with David.

Categorize Because the problem asks to find an observed velocity, we categorize this problem as one requiring the Lorentz velocity transformation. The motion takes place in two dimensions.

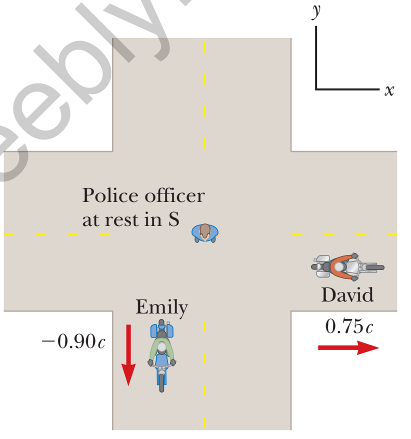


Figure 39.15 (Example 39.7) David moves east with a speed 0.75c relative to the police officer, and Emily travels south at a speed 0.90c relative to the officer.

Analyze Identify the velocity components for David and Emily according to the police officer:

David: $v_x = v = 0.75c \quad v_y = 0$

Emily: $u_x = 0 \quad u_y = -0.90c$

Using Equations 39.16 and 39.17, calculate u'_x and u'_y for Emily as measured by David:

$$u'_x = \frac{u_x - v}{1 - \dfrac{u_x v}{c^2}} = \frac{0 - 0.75c}{1 - \dfrac{(0)(0.75c)}{c^2}} = -0.75c$$

$$u'_y = \frac{u_y}{\gamma\left(1 - \dfrac{u_x v}{c^2}\right)} = \frac{\sqrt{1 - \dfrac{(0.75c)^2}{c^2}}\,(-0.90c)}{1 - \dfrac{(0)(0.75c)}{c^2}} = -0.60c$$

Using the Pythagorean theorem, find the speed of Emily as measured by David:

$$u' = \sqrt{(u'_x)^2 + (u'_y)^2} = \sqrt{(-0.75c)^2 + (-0.60c)^2} = \boxed{0.96c}$$

Finalize This speed is less than c, as required by the special theory of relativity.

39.7 Relativistic Linear Momentum

To describe the motion of particles within the framework of the special theory of relativity properly, you must replace the Galilean transformation equations by the Lorentz transformation equations. Because the laws of physics must remain unchanged under the Lorentz transformation, we must generalize Newton's laws and the definitions of linear momentum and energy to conform to the Lorentz transformation equations and the principle of relativity. These generalized definitions should reduce to the classical (nonrelativistic) definitions for $v << c$.

First, recall from the isolated system model that when two particles (or objects that can be modeled as particles) collide, the total momentum of the isolated system of the two particles remains constant. Suppose we observe this collision in a reference frame S and confirm that the momentum of the system is conserved. Now imagine that the momenta of the particles are measured by an observer in a second reference frame S′ moving with velocity $\vec{\mathbf{v}}$ relative to the first frame. Using the Lorentz velocity transformation equation and the classical definition of linear momentum, $\vec{\mathbf{p}} = m\vec{\mathbf{u}}$ (where $\vec{\mathbf{u}}$ is the velocity of a particle), we find that linear momentum of the system is *not* measured to be conserved by the observer in S′. Because the laws of physics are the same in all inertial frames, however, linear momentum of the system must be conserved in all frames. We have a contradiction. In view of this contradiction and assuming the Lorentz velocity transformation equation is correct, we must modify the definition of linear momentum so that the momentum of an isolated system is conserved for all observers. For any particle, the correct relativistic equation for linear momentum that satisfies this condition is

$$\vec{\mathbf{p}} \equiv \frac{m\vec{\mathbf{u}}}{\sqrt{1 - \dfrac{u^2}{c^2}}} = \gamma m\vec{\mathbf{u}} \tag{39.19}$$

◀ **Definition of relativistic linear momentum**

where m is the mass of the particle and $\vec{\mathbf{u}}$ is the velocity of the particle. When u is much less than c, $\gamma = (1 - u^2/c^2)^{-1/2}$ approaches unity and $\vec{\mathbf{p}}$ approaches $m\vec{\mathbf{u}}$. Therefore, the relativistic equation for $\vec{\mathbf{p}}$ reduces to the classical expression when u is much smaller than c, as it should.

The relativistic force $\vec{\mathbf{F}}$ acting on a particle whose linear momentum is $\vec{\mathbf{p}}$ is defined as

$$\vec{\mathbf{F}} \equiv \frac{d\vec{\mathbf{p}}}{dt} \tag{39.20}$$

where $\vec{\mathbf{p}}$ is given by Equation 39.19. This expression, which is the relativistic form of Newton's second law, is reasonable because it preserves classical mechanics in the limit of low velocities and is consistent with conservation of linear momentum for an isolated system ($\vec{\mathbf{F}}_{\text{ext}} = 0$) both relativistically and classically.

It is left as an end-of-chapter problem (Problem 88) to show that under relativistic conditions, the acceleration $\vec{\mathbf{a}}$ of a particle decreases under the action of a constant force, in which case $a \propto (1 - u^2/c^2)^{3/2}$. This proportionality shows that as the particle's speed approaches c, the acceleration caused by any finite force approaches zero. Hence, it is impossible to accelerate a particle from rest to a speed $u \geq c$. This argument reinforces that the speed of light is the ultimate speed, the speed limit of the Universe. It is the maximum possible speed for energy transfer and for information transfer. Any object with mass must move at a lower speed.

Pitfall Prevention 39.6

Watch Out for "Relativistic Mass" Some older treatments of relativity maintained the conservation of momentum principle at high speeds by using a model in which a particle's mass increases with speed. You might still encounter this notion of "relativistic mass" in your outside reading, especially in older books. Be aware that this notion is no longer widely accepted; today, mass is considered as *invariant,* independent of speed. The mass of an object in all frames is considered to be the mass as measured by an observer at rest with respect to the object.

Example 39.8 Linear Momentum of an Electron

An electron, which has a mass of 9.11×10^{-31} kg, moves with a speed of $0.750c$. Find the magnitude of its relativistic momentum and compare this value with the momentum calculated from the classical expression.

SOLUTION

Conceptualize Imagine an electron moving with high speed. The electron carries momentum, but the magnitude of its momentum is not given by $p = mu$ because the speed is relativistic.

Categorize We categorize this example as a substitution problem involving a relativistic equation.

continued

39.8 continued

Use Equation 39.19 with $u = 0.750c$ to find the magnitude of the momentum:

$$p = \frac{m_e u}{\sqrt{1 - \dfrac{u^2}{c^2}}}$$

$$p = \frac{(9.11 \times 10^{-31}\ \text{kg})(0.750)(3.00 \times 10^8\ \text{m/s})}{\sqrt{1 - \dfrac{(0.750c)^2}{c^2}}}$$

$$= \boxed{3.10 \times 10^{-22}\ \text{kg} \cdot \text{m/s}}$$

The classical expression (used incorrectly here) gives $p_{\text{classical}} = m_e u = 2.05 \times 10^{-22}\ \text{kg} \cdot \text{m/s}$. Hence, the correct relativistic result is 50% greater than the classical result!

39.8 Relativistic Energy

We have seen that the definition of linear momentum requires generalization to make it compatible with Einstein's postulates. This conclusion implies that the definition of kinetic energy must most likely be modified also.

To derive the relativistic form of the work–kinetic energy theorem, imagine a particle moving in one dimension along the x axis. A force in the x direction causes the momentum of the particle to change according to Equation 39.20. In what follows, we assume the particle is accelerated from rest to some final speed u. The work done by the force F on the particle is

$$W = \int_{x_1}^{x_2} F\,dx = \int_{x_1}^{x_2} \frac{dp}{dt}\,dx \qquad \textbf{(39.21)}$$

To perform this integration and find the work done on the particle and the relativistic kinetic energy as a function of u, we first evaluate dp/dt:

$$\frac{dp}{dt} = \frac{d}{dt}\frac{mu}{\sqrt{1 - \dfrac{u^2}{c^2}}} = \frac{m}{\left(1 - \dfrac{u^2}{c^2}\right)^{3/2}}\frac{du}{dt}$$

Substituting this expression for dp/dt and $dx = u\,dt$ into Equation 39.21 gives

$$W = \int_0^t \frac{m}{\left(1 - \dfrac{u^2}{c^2}\right)^{3/2}}\frac{du}{dt}(u\,dt) = m\int_0^u \frac{u}{\left(1 - \dfrac{u^2}{c^2}\right)^{3/2}}\,du$$

where we use the limits 0 and u in the integral because the integration variable has been changed from t to u. Evaluating the integral gives

$$W = \frac{mc^2}{\sqrt{1 - \dfrac{u^2}{c^2}}} - mc^2 \qquad \textbf{(39.22)}$$

Recall from Chapter 7 that the work done by a force acting on a system consisting of a single particle equals the change in kinetic energy of the particle: $W = \Delta K$. Because we assumed the initial speed of the particle is zero, its initial kinetic energy is zero, so $W = K - K_i = K - 0 = K$. Therefore, the work W in Equation 39.22 is equivalent to the relativistic kinetic energy K:

Relativistic kinetic energy ▶

$$K = \frac{mc^2}{\sqrt{1 - \dfrac{u^2}{c^2}}} - mc^2 = \gamma mc^2 - mc^2 = (\gamma - 1)mc^2 \qquad \textbf{(39.23)}$$

This equation is routinely confirmed by experiments using high-energy particle accelerators.

At low speeds, where $u/c << 1$, Equation 39.23 should reduce to the classical expression $K = \frac{1}{2}mu^2$. We can check that by using the binomial expansion $(1 - \beta^2)^{-1/2} \approx 1 + \frac{1}{2}\beta^2 + \cdots$ for $\beta << 1$, where the higher-order powers of β are neglected in the expansion. (In treatments of relativity, β is a common symbol used to represent u/c or v/c.) In our case, $\beta = u/c$, so

$$\gamma = \frac{1}{\sqrt{1 - \dfrac{u^2}{c^2}}} = \left(1 - \frac{u^2}{c^2}\right)^{-1/2} \approx 1 + \frac{1}{2}\frac{u^2}{c^2}$$

Substituting this result into Equation 39.23 gives

$$K \approx \left[\left(1 + \frac{1}{2}\frac{u^2}{c^2}\right) - 1\right]mc^2 = \frac{1}{2}mu^2 \quad (\text{for } u/c << 1)$$

which is the classical expression for kinetic energy. A graph comparing the relativistic and nonrelativistic expressions is given in Figure 39.16. In the relativistic case, the particle speed never exceeds c, regardless of the kinetic energy. The two curves are in good agreement when $u << c$.

The constant term mc^2 in Equation 39.23, which is independent of the speed of the particle, is called the **rest energy** E_R of the particle:

$$E_R = mc^2 \qquad \textbf{(39.24)}$$

Equation 39.24 shows that **mass is a form of energy,** where c^2 is simply a constant conversion factor. This expression also shows that a small mass corresponds to an enormous amount of energy, a concept fundamental to nuclear and elementary-particle physics.

The term γmc^2 in Equation 39.23, which depends on the particle speed, is the sum of the kinetic and rest energies. It is called the **total energy** E:

Total energy = kinetic energy + rest energy

$$E = K + mc^2 \qquad \textbf{(39.25)}$$

or

$$E = \frac{mc^2}{\sqrt{1 - \dfrac{u^2}{c^2}}} = \gamma mc^2 \qquad \textbf{(39.26)}$$

◀ **Total energy of a relativistic particle**

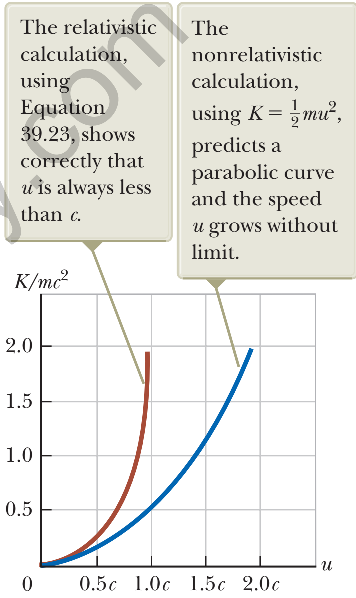


Figure 39.16 A graph comparing relativistic and nonrelativistic kinetic energy of a moving particle. The energies are plotted as a function of particle speed u.

In many situations, the linear momentum or energy of a particle rather than its speed is measured. It is therefore useful to have an expression relating the total energy E to the relativistic linear momentum p, which is accomplished by using the expressions $E = \gamma mc^2$ and $p = \gamma mu$. By squaring these equations and subtracting, we can eliminate u (Problem 58). The result, after some algebra, is[6]

$$E^2 = p^2c^2 + (mc^2)^2 \qquad \textbf{(39.27)}$$

◀ **Energy–momentum relationship for a relativistic particle**

When the particle is at rest, $p = 0$, so $E = E_R = mc^2$.

In Section 35.1, we introduced the concept of a particle of light, called a **photon.** For particles that have zero mass, such as photons, we set $m = 0$ in Equation 39.27 and find that

$$E = pc \qquad \textbf{(39.28)}$$

[6]One way to remember this relationship is to draw a right triangle having a hypotenuse of length E and legs of lengths pc and mc^2.

This equation is an exact expression relating total energy and linear momentum for photons, which always travel at the speed of light (in vacuum).

Finally, because the mass m of a particle is independent of its motion, m must have the same value in all reference frames. For this reason, m is often called the **invariant mass.** On the other hand, because the total energy and linear momentum of a particle both depend on velocity, these quantities depend on the reference frame in which they are measured.

When dealing with subatomic particles, it is convenient to express their energy in electron volts (Section 25.1) because the particles are usually given this energy by acceleration through a potential difference. The conversion factor, as you recall from Equation 25.5, is

$$1 \text{ eV} = 1.602 \times 10^{-19} \text{ J}$$

For example, the mass of an electron is 9.109×10^{-31} kg. Hence, the rest energy of the electron is

$$m_e c^2 = (9.109 \times 10^{-31} \text{ kg})(2.998 \times 10^8 \text{ m/s})^2 = 8.187 \times 10^{-14} \text{ J}$$

$$= (8.187 \times 10^{-14} \text{ J})(1 \text{ eV}/1.602 \times 10^{-19} \text{ J}) = 0.511 \text{ MeV}$$

Quick Quiz 39.9 The following *pairs* of energies—particle 1: E, $2E$; particle 2: E, $3E$; particle 3: $2E$, $4E$—represent the rest energy and total energy of three different particles. Rank the particles from greatest to least according to their **(a)** mass, **(b)** kinetic energy, and **(c)** speed.

Example 39.9 The Energy of a Speedy Proton

(A) Find the rest energy of a proton in units of electron volts.

SOLUTION

Conceptualize Even if the proton is not moving, it has energy associated with its mass. If it moves, the proton possesses more energy, with the total energy being the sum of its rest energy and its kinetic energy.

Categorize The phrase "rest energy" suggests we must take a relativistic rather than a classical approach to this problem.

Analyze Use Equation 39.24 to find the rest energy:

$$E_R = m_p c^2 = (1.672\,6 \times 10^{-27} \text{ kg})(2.998 \times 10^8 \text{ m/s})^2$$

$$= (1.504 \times 10^{-10} \text{ J})\left(\frac{1.00 \text{ eV}}{1.602 \times 10^{-19} \text{ J}}\right) = \boxed{938 \text{ MeV}}$$

(B) If the total energy of a proton is three times its rest energy, what is the speed of the proton?

SOLUTION

Use Equation 39.26 to relate the total energy of the proton to the rest energy:

$$E = 3m_p c^2 = \frac{m_p c^2}{\sqrt{1 - \frac{u^2}{c^2}}} \rightarrow 3 = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}}$$

Solve for u:

$$1 - \frac{u^2}{c^2} = \frac{1}{9} \rightarrow \frac{u^2}{c^2} = \frac{8}{9}$$

$$u = \frac{\sqrt{8}}{3} c = 0.943c = \boxed{2.83 \times 10^8 \text{ m/s}}$$

(C) Determine the kinetic energy of the proton in units of electron volts.

39.9 continued

SOLUTION

Use Equation 39.25 to find the kinetic energy of the proton:

$$K = E - m_pc^2 = 3m_pc^2 - m_pc^2 = 2m_pc^2$$
$$= 2(938 \text{ MeV}) = \boxed{1.88 \times 10^3 \text{ MeV}}$$

(D) What is the proton's momentum?

SOLUTION

Use Equation 39.27 to calculate the momentum:

$$E^2 = p^2c^2 + (m_pc^2)^2 = (3m_pc^2)^2$$
$$p^2c^2 = 9(m_pc^2)^2 - (m_pc^2)^2 = 8(m_pc^2)^2$$
$$p = \sqrt{8}\,\frac{m_pc^2}{c} = \sqrt{8}\,\frac{938 \text{ MeV}}{c} = \boxed{2.65 \times 10^3 \text{ MeV}/c}$$

Finalize The unit of momentum in part (D) is written MeV/c, which is a common unit in particle physics. For comparison, you might want to solve this example using classical equations.

WHAT IF? In classical physics, if the momentum of a particle doubles, the kinetic energy increases by a factor of 4. What happens to the kinetic energy of the proton in this example if its momentum doubles?

Answer Based on what we have seen so far in relativity, it is likely you would predict that its kinetic energy does not increase by a factor of 4.

Find the new doubled momentum:

$$p_{\text{new}} = 2\left(\sqrt{8}\,\frac{m_pc^2}{c}\right) = 4\sqrt{2}\,\frac{m_pc^2}{c}$$

Use this result in Equation 39.27 to find the new total energy:

$$E_{\text{new}}^2 = p_{\text{new}}^2c^2 + (m_pc^2)^2$$
$$E_{\text{new}}^2 = \left(4\sqrt{2}\,\frac{m_pc^2}{c}\right)^2c^2 + (m_pc^2)^2 = 33(m_pc^2)^2$$
$$E_{\text{new}} = \sqrt{33}\,m_pc^2 = 5.7m_pc^2$$

Use Equation 39.25 to find the new kinetic energy:

$$K_{\text{new}} = E_{\text{new}} - m_pc^2 = 5.7m_pc^2 - m_pc^2 = 4.7m_pc^2$$

This value is a little more than twice the kinetic energy found in part (C), not four times. In general, the factor by which the kinetic energy increases if the momentum doubles depends on the initial momentum, but it approaches 4 as the momentum approaches zero. In this latter situation, classical physics correctly describes the situation.

Equation 39.26, $E = \gamma mc^2$, represents the total energy of a particle. This important equation suggests that even when a particle is at rest ($\gamma = 1$), it still possesses enormous energy through its mass. The clearest experimental proof of the equivalence of mass and energy occurs in nuclear and elementary-particle interactions in which the conversion of mass into kinetic energy takes place. Consequently, we cannot use the principle of conservation of energy in relativistic situations as it was outlined in Chapter 8. We must modify the principle by including rest energy as another form of energy storage.

This concept is important in atomic and nuclear processes, in which the change in mass is a relatively large fraction of the initial mass. In a conventional nuclear reactor, for example, the uranium nucleus undergoes *fission,* a reaction that results in several lighter fragments having considerable kinetic energy. In the case of ^{235}U, which is used as fuel in nuclear power plants, the fragments are two lighter nuclei and a few neutrons. The total mass of the fragments is less than that of the ^{235}U by an amount Δm. The corresponding energy Δmc^2 associated with this mass difference is exactly equal to the sum of the kinetic energies of the fragments. The kinetic energy is absorbed as the fragments move through water, raising the internal energy of the water. This internal energy is used to produce steam for the generation of electricity.

Next, consider a basic *fusion* reaction in which two deuterium atoms combine to form one helium atom. The decrease in mass that results from the creation of one helium atom from two deuterium atoms is $\Delta m = 4.25 \times 10^{-29}$ kg. Hence, the corresponding energy that results from one fusion reaction is $\Delta mc^2 = 3.83 \times 10^{-12}$ J = 23.9 MeV. To appreciate the magnitude of this result, consider that if only 1 g of deuterium were converted to helium, the energy released would be on the order of 10^{12} J! In 2013's cost of electrical energy, this energy would be worth approximately \$35 000. We shall present more details of these nuclear processes in Chapter 45 of the extended version of this textbook.

Example 39.10 Mass Change in a Radioactive Decay

The ^{216}Po nucleus is unstable and exhibits radioactivity (Chapter 44). It decays to ^{212}Pb by emitting an alpha particle, which is a helium nucleus, ^{4}He. The relevant masses, in atomic mass units (see Table A.1 in Appendix A), are $m_i = m(^{216}\text{Po}) = 216.001\ 915$ u and $m_f = m(^{212}\text{Pb}) + m(^4\text{He}) = 211.991\ 898\ \text{u} + 4.002\ 603$ u.

(A) Find the mass change of the system in this decay.

SOLUTION

Conceptualize The initial system is the ^{216}Po nucleus. Imagine the mass of the system decreasing during the decay and transforming to kinetic energy of the alpha particle and the ^{212}Pb nucleus after the decay.

Categorize We use concepts discussed in this section, so we categorize this example as a substitution problem.

Calculate the change in mass using the mass values given in the problem statement.

$$\Delta m = 216.001\ 915\ \text{u} - (211.991\ 898\ \text{u} + 4.002\ 603\ \text{u})$$
$$= 0.007\ 414\ \text{u} = \boxed{1.23 \times 10^{-29}\ \text{kg}}$$

(B) Find the energy this mass change represents.

SOLUTION

Use Equation 39.24 to find the energy associated with this mass change:

$$E = \Delta mc^2 = (1.23 \times 10^{-29}\ \text{kg})(3.00 \times 10^8\ \text{m/s})^2$$
$$= 1.11 \times 10^{-12}\ \text{J} = \boxed{6.92\ \text{MeV}}$$

39.9 The General Theory of Relativity

Up to this point, we have sidestepped a curious puzzle. Mass has two seemingly different properties: a *gravitational attraction* for other masses and an *inertial* property that represents a resistance to acceleration. We first discussed these two attributes for mass in Section 5.5. To designate these two attributes, we use the subscripts g and i and write

$$\text{Gravitational property:} \qquad F_g = m_g g$$

$$\text{Inertial property:} \qquad \sum F = m_i a$$

The value for the gravitational constant G was chosen to make the magnitudes of m_g and m_i numerically equal. Regardless of how G is chosen, however, the strict proportionality of m_g and m_i has been established experimentally to an extremely high degree: a few parts in 10^{12}. Therefore, it appears that gravitational mass and inertial mass may indeed be exactly proportional.

Why, though? They seem to involve two entirely different concepts: a force of mutual gravitational attraction between two masses and the resistance of a single mass to being accelerated. This question, which puzzled Newton and many other physicists over the years, was answered by Einstein in 1916 when he published his theory of gravitation, known as the *general theory of relativity.* Because it is a mathematically complex theory, we offer merely a hint of its elegance and insight.

In Einstein's view, the dual behavior of mass was evidence for a very intimate and basic connection between the two behaviors. He pointed out that no mechanical experiment (such as dropping an object) could distinguish between the two situations illustrated in Figures 39.17a and 39.17b. In Figure 39.17a, a person standing in an elevator on the surface of a planet feels pressed into the floor due to the gravitational force. If he releases his briefcase, he observes it moving toward the floor with acceleration $\vec{\mathbf{g}} = -g\hat{\mathbf{j}}$. In Figure 39.17b, the person is in an elevator in empty space accelerating upward with $\vec{\mathbf{a}}_{el} = +g\hat{\mathbf{j}}$. The person feels pressed into the floor with the same force as in Figure 39.17a. If he releases his briefcase, he observes it moving toward the floor with acceleration g, exactly as in the previous situation. In each situation, an object released by the observer undergoes a downward acceleration of magnitude g relative to the floor. In Figure 39.17a, the person is at rest in an inertial frame in a gravitational field due to the planet. In Figure 39.17b, the person is in a noninertial frame accelerating in gravity-free space. Einstein's claim is that these two situations are completely equivalent.

Einstein carried this idea further and proposed that *no* experiment, mechanical or otherwise, could distinguish between the two situations. This extension to include all phenomena (not just mechanical ones) has interesting consequences. For example, suppose a light pulse is sent horizontally across the elevator as in Figure 39.17c, in which the elevator is accelerating upward in empty space. From the point of view of an observer in an inertial frame outside the elevator, the light travels in a straight line while the floor of the elevator accelerates upward. According to the observer on the elevator, however, the trajectory of the light pulse bends downward as the floor of the elevator (and the observer) accelerates upward. Therefore, based on the equality of parts (a) and (b) of the figure, Einstein proposed that a

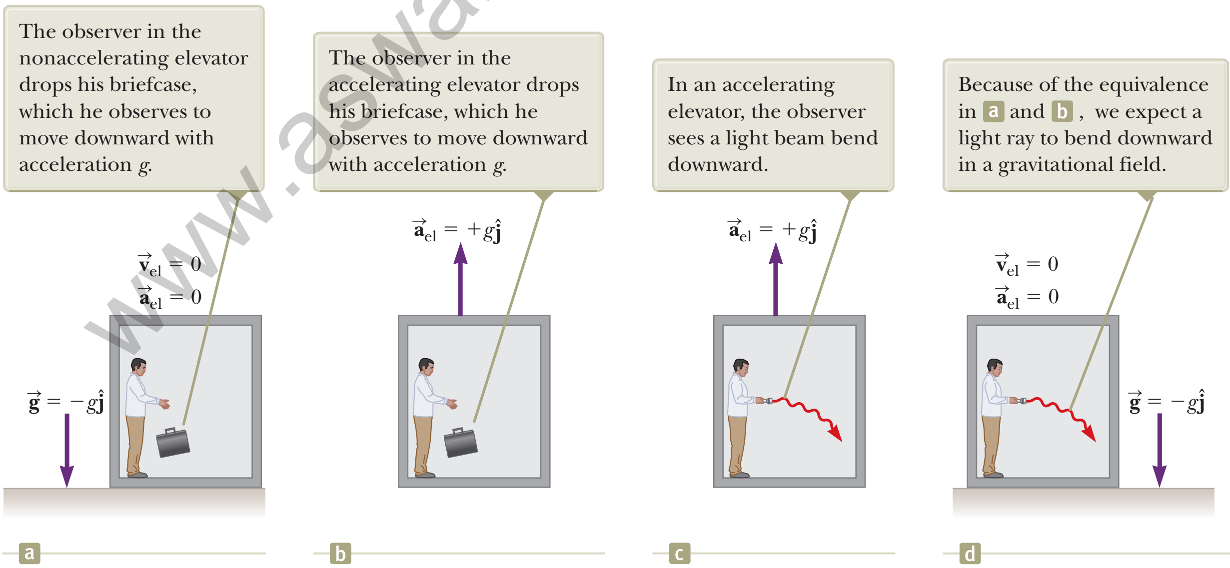


Figure 39.17 (a) The observer is at rest in an elevator in a uniform gravitational field $\vec{\mathbf{g}} = -g\hat{\mathbf{j}}$, directed downward. (b) The observer is in a region where gravity is negligible, but the elevator moves upward with an acceleration $\vec{\mathbf{a}}_{el} = +g\hat{\mathbf{j}}$. According to Einstein, the frames of reference in (a) and (b) are equivalent in every way. No local experiment can distinguish any difference between the two frames. (c) An observer watches a beam of light in an accelerating elevator. (d) Einstein's prediction of the behavior of a beam of light in a gravitational field.

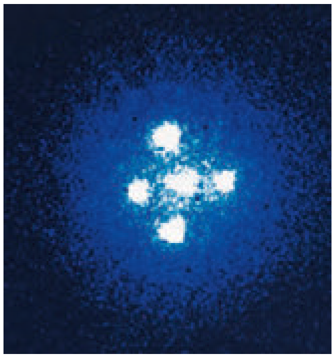
beam of light should also be bent downward by a gravitational field as in Figure 39.17d. Experiments have verified the effect, although the bending is small. A laser aimed at the horizon falls less than 1 cm after traveling 6 000 km. (No such bending is predicted in Newton's theory of gravitation.)

Einstein's **general theory of relativity** has two postulates:

- All the laws of nature have the same form for observers in any frame of reference, whether accelerated or not.
- In the vicinity of any point, a gravitational field is equivalent to an accelerated frame of reference in gravity-free space (the **principle of equivalence**).

One interesting effect predicted by the general theory is that time is altered by gravity. A clock in the presence of gravity runs slower than one located where gravity is negligible. Consequently, the frequencies of radiation emitted by atoms in the presence of a strong gravitational field are *redshifted* to lower frequencies when compared with the same emissions in the presence of a weak field. This gravitational redshift has been detected in spectral lines emitted by atoms in massive stars. It has also been verified on the Earth by comparing the frequencies of gamma rays emitted from nuclei separated vertically by about 20 m.

The second postulate suggests a gravitational field may be "transformed away" at any point if we choose an appropriate accelerated frame of reference, a freely falling one. Einstein developed an ingenious method of describing the acceleration necessary to make the gravitational field "disappear." He specified a concept, the *curvature of space–time,* that describes the gravitational effect at every point. In fact, the curvature of space–time completely replaces Newton's gravitational theory. According to Einstein, there is no such thing as a gravitational force. Rather, the presence of a mass causes a curvature of space–time in the vicinity of the mass, and this curvature dictates the space–time path that all freely moving objects must follow.



Einstein's cross. The four outer bright spots are images of the same galaxy that have been bent around a massive object located between the galaxy and the Earth. The massive object acts like a lens, causing the rays of light that were diverging from the distant galaxy to converge on the Earth. (If the intervening massive object had a uniform mass distribution, we would see a bright ring instead of four spots.)

As an example of the effects of curved space–time, imagine two travelers moving on parallel paths a few meters apart on the surface of the Earth and maintaining an exact northward heading along two longitude lines. As they observe each other near the equator, they will claim that their paths are exactly parallel. As they approach the North Pole, however, they notice that they are moving closer together and will meet at the North Pole. Therefore, they claim that they moved along parallel paths, but moved toward each other, *as if there were an attractive force between them.* The travelers make this conclusion based on their everyday experience of moving on flat surfaces. From our mental representation, however, we realize they are walking on a curved surface, and it is the geometry of the curved surface, rather than an attractive force, that causes them to converge. In a similar way, general relativity replaces the notion of forces with the movement of objects through curved space–time.

One prediction of the general theory of relativity is that a light ray passing near the Sun should be deflected in the curved space–time created by the Sun's mass. This prediction was confirmed when astronomers detected the bending of starlight near the Sun during a total solar eclipse that occurred shortly after World War I (Fig. 39.18). When this discovery was announced, Einstein became an international celebrity.

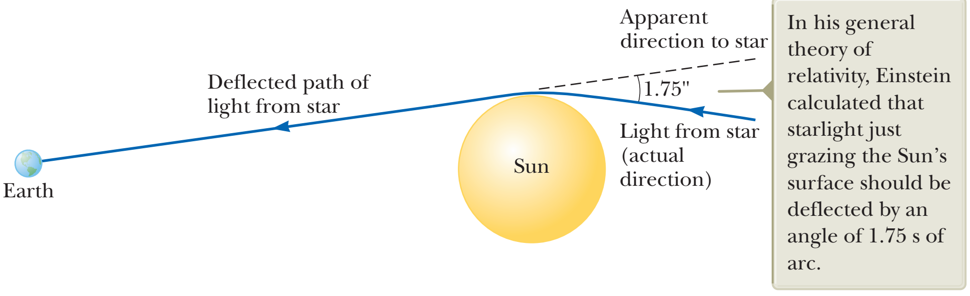


Figure 39.18 Deflection of starlight passing near the Sun. Because of this effect, the Sun or some other remote object can act as a *gravitational lens.*

If the concentration of mass becomes very great as is believed to occur when a large star exhausts its nuclear fuel and collapses to a very small volume, a **black hole** may form as discussed in Chapter 13. Here, the curvature of space–time is so extreme that within a certain distance from the center of the black hole all matter and light become trapped as discussed in Section 13.6.

Summary

Definitions

The relativistic expression for the **linear momentum** of a particle moving with a velocity $\vec{\mathbf{u}}$ is

$$\vec{\mathbf{p}} \equiv \frac{m\vec{\mathbf{u}}}{\sqrt{1-\frac{u^2}{c^2}}} = \gamma m\vec{\mathbf{u}} \qquad \textbf{(39.19)}$$

The relativistic force $\vec{\mathbf{F}}$ acting on a particle whose linear momentum is $\vec{\mathbf{p}}$ is defined as

$$\vec{\mathbf{F}} \equiv \frac{d\vec{\mathbf{p}}}{dt} \qquad \textbf{(39.20)}$$

Concepts and Principles

The two basic postulates of the special theory of relativity are as follows:

- The laws of physics must be the same in all inertial reference frames.
- The speed of light in vacuum has the same value, $c = 3.00 \times 10^8$ m/s, in all inertial frames, regardless of the velocity of the observer or the velocity of the source emitting the light.

Three consequences of the special theory of relativity are as follows:

- Events that are measured to be simultaneous for one observer are not necessarily measured to be simultaneous for another observer who is in motion relative to the first.
- Clocks in motion relative to an observer are measured to run slower by a factor $\gamma = (1 - v^2/c^2)^{-1/2}$. This phenomenon is known as **time dilation.**
- The lengths of objects in motion are measured to be shorter in the direction of motion by a factor $1/\gamma = (1 - v^2/c^2)^{1/2}$. This phenomenon is known as **length contraction.**

To satisfy the postulates of special relativity, the Galilean transformation equations must be replaced by the **Lorentz transformation equations:**

$$x' = \gamma(x - vt) \quad y' = y \quad z' = z \quad t' = \gamma\left(t - \frac{v}{c^2}x\right) \qquad \textbf{(39.11)}$$

where $\gamma = (1 - v^2/c^2)^{-1/2}$ and the S′ frame moves in the x direction at speed v relative to the S frame.

The relativistic form of the **Lorentz velocity transformation equation** is

$$u'_x = \frac{u_x - v}{1 - \frac{u_x v}{c^2}} \qquad \textbf{(39.16)}$$

where u'_x is the x component of the velocity of an object as measured in the S′ frame and u_x is its component as measured in the S frame.

The relativistic expression for the **kinetic energy** of a particle is

$$K = \frac{mc^2}{\sqrt{1-\frac{u^2}{c^2}}} - mc^2 = (\gamma - 1)mc^2 \qquad \textbf{(39.23)}$$

The constant term mc^2 in Equation 39.23 is called the **rest energy** E_R of the particle:

$$E_R = mc^2 \qquad \textbf{(39.24)}$$

The **total energy** E of a particle is given by

$$E = \frac{mc^2}{\sqrt{1-\frac{u^2}{c^2}}} = \gamma mc^2 \qquad \textbf{(39.26)}$$

The relativistic linear momentum of a particle is related to its total energy through the equation

$$E^2 = p^2c^2 + (mc^2)^2 \qquad \textbf{(39.27)}$$

Objective Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. **(i)** Does the speed of an electron have an upper limit? (a) yes, the speed of light c (b) yes, with another value (c) no **(ii)** Does the magnitude of an electron's momentum have an upper limit? (a) yes, m_ec (b) yes, with another value (c) no **(iii)** Does the electron's kinetic energy have an upper limit? (a) yes, m_ec^2 (b) yes, $\frac{1}{2}m_ec^2$ (c) yes, with another value (d) no

2. A spacecraft zooms past the Earth with a constant velocity. An observer on the Earth measures that an undamaged clock on the spacecraft is ticking at one-third the rate of an identical clock on the Earth. What does an observer on the spacecraft measure about the Earth-based clock's ticking rate? (a) It runs more than three times faster than his own clock. (b) It runs three times faster than his own. (c) It runs at the same rate as his own. (d) It runs at one-third the rate of his own. (e) It runs at less than one-third the rate of his own.

3. As a car heads down a highway traveling at a speed v away from a ground observer, which of the following statements are true about the measured speed of the light beam from the car's headlights? More than one statement may be correct. (a) The ground observer measures the light speed to be $c + v$. (b) The driver measures the light speed to be c. (c) The ground observer measures the light speed to be c. (d) The driver measures the light speed to be $c - v$. (e) The ground observer measures the light speed to be $c - v$.

4. A spacecraft built in the shape of a sphere moves past an observer on the Earth with a speed of $0.500c$. What shape does the observer measure for the spacecraft as it goes by? (a) a sphere (b) a cigar shape, elongated along the direction of motion (c) a round pillow shape, flattened along the direction of motion (d) a conical shape, pointing in the direction of motion

5. An astronaut is traveling in a spacecraft in outer space in a straight line at a constant speed of $0.500c$. Which of the following effects would she experience? (a) She would feel heavier. (b) She would find it harder to breathe. (c) Her heart rate would change. (d) Some of the dimensions of her spacecraft would be shorter. (e) None of those answers is correct.

6. You measure the volume of a cube at rest to be V_0. You then measure the volume of the same cube as it passes you in a direction parallel to one side of the cube. The speed of the cube is $0.980c$, so $\gamma \approx 5$. Is the volume you measure close to (a) $V_0/25$, (b) $V_0/5$, (c) V_0, (d) $5V_0$, or (e) $25V_0$?

7. Two identical clocks are set side by side and synchronized. One remains on the Earth. The other is put into orbit around the Earth moving rapidly toward the east. **(i)** As measured by an observer on the Earth, does the orbiting clock (a) run faster than the Earth-based clock, (b) run at the same rate, or (c) run slower? **(ii)** The orbiting clock is returned to its original location and brought to rest relative to the Earth-based clock. Thereafter, what happens? (a) Its reading lags farther and farther behind the Earth-based clock. (b) It lags behind the Earth-based clock by a constant amount. (c) It is synchronous with the Earth-based clock. (d) It is ahead of the Earth-based clock by a constant amount. (e) It gets farther and farther ahead of the Earth-based clock.

8. The following three particles all have the same total energy E: (a) a photon, (b) a proton, and (c) an electron. Rank the magnitudes of the particles' momenta from greatest to smallest.

9. Which of the following statements are fundamental postulates of the special theory of relativity? More than one statement may be correct. (a) Light moves through a substance called the ether. (b) The speed of light depends on the inertial reference frame in which it is measured. (c) The laws of physics depend on the inertial reference frame in which they are used. (d) The laws of physics are the same in all inertial reference frames. (e) The speed of light is independent of the inertial reference frame in which it is measured.

10. A distant astronomical object (a quasar) is moving away from us at half the speed of light. What is the speed of the light we receive from this quasar? (a) greater than c (b) c (c) between $c/2$ and c (d) $c/2$ (e) between 0 and $c/2$

Conceptual Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. In several cases, a nearby star has been found to have a large planet orbiting about it, although light from the planet could not be seen separately from the starlight. Using the ideas of a system rotating about its center of mass and of the Doppler shift for light, explain how an astronomer could determine the presence of the invisible planet.

2. Explain why, when defining the length of a rod, it is necessary to specify that the positions of the ends of the rod are to be measured simultaneously.

3. A train is approaching you at very high speed as you stand next to the tracks. Just as an observer on the train passes you, you both begin to play the same recorded version of a Beethoven symphony on identical iPods. (a) According to you, whose iPod finishes the symphony first? (b) **What If?** According to the observer on the train, whose iPod finishes the symphony first? (c) Whose iPod actually finishes the symphony first?

4. List three ways our day-to-day lives would change if the speed of light were only 50 m/s.

5. How is acceleration indicated on a space–time graph?

6. (a) "Newtonian mechanics correctly describes objects moving at ordinary speeds, and relativistic mechanics correctly describes objects moving very fast." (b) "Relativistic mechanics must make a smooth transition as

it reduces to Newtonian mechanics in a case in which the speed of an object becomes small compared with the speed of light." Argue for or against statements (a) and (b).

7. The speed of light in water is 230 Mm/s. Suppose an electron is moving through water at 250 Mm/s. Does that violate the principle of relativity? Explain.

8. A particle is moving at a speed less than $c/2$. If the speed of the particle is doubled, what happens to its momentum?

9. Give a physical argument that shows it is impossible to accelerate an object of mass m to the speed of light, even with a continuous force acting on it.

10. Explain how the Doppler effect with microwaves is used to determine the speed of an automobile.

11. It is said that Einstein, in his teenage years, asked the question, "What would I see in a mirror if I carried it in my hands and ran at a speed near that of light?" How would you answer this question?

12. **(i)** An object is placed at a position $p > f$ from a concave mirror as shown in Figure CQ39.12a, where f is the focal length of the mirror. In a finite time interval, the object is moved to the right to a position at the focal point F of the mirror. Show that the image of the object moves at a speed greater than the speed of light. **(ii)** A laser pointer is suspended in a horizontal plane and set into rapid rotation as shown in Figure CQ39.12b. Show that the spot of light it produces on a distant screen can move across the screen at a speed greater than the speed of light. (If you carry out this experiment, make sure the direct laser light cannot enter a person's eyes.) **(iii)** Argue that the experiments in parts (i) and (ii) do not invalidate the principle that no material, no energy, and no information can move faster than light moves in a vacuum.

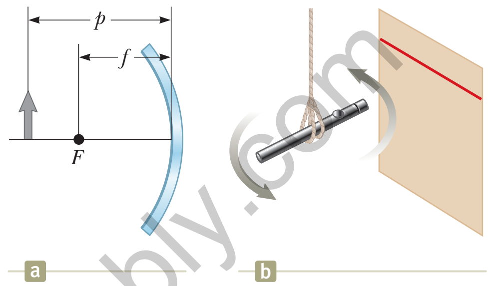


Figure CQ39.12

13. With regard to reference frames, how does general relativity differ from special relativity?

14. Two identical clocks are in the same house, one upstairs in a bedroom and the other downstairs in the kitchen. Which clock runs slower? Explain.

Problems

The problems found in this chapter may be assigned online in Enhanced WebAssign

1. straightforward; **2.** intermediate; **3.** challenging

1. full solution available in the *Student Solutions Manual/Study Guide*

AMT Analysis Model tutorial available in Enhanced WebAssign

GP Guided Problem

M Master It tutorial available in Enhanced WebAssign

W Watch It video solution available in Enhanced WebAssign

Section 39.1 The Principle of Galilean Relativity

Problems 46–48, 50, 51, 53–54, and 79 in Chapter 4 can be assigned with this section.

1. The truck in Figure P39.1 is moving at a speed of 10.0 m/s relative to the ground. The person on the truck throws a baseball in the backward direction at a speed of 20.0 m/s relative to the truck. What is the velocity of the baseball as measured by the observer on the ground?

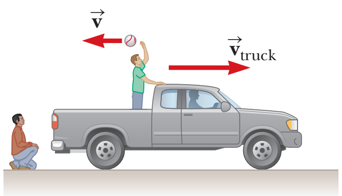


Figure P39.1

2. In a laboratory frame of reference, an observer notes that Newton's second law is valid. Assume forces and masses are measured to be the same in any reference frame for speeds small compared with the speed of light. (a) Show that Newton's second law is also valid for an observer moving at a constant speed, small compared with the speed of light, relative to the laboratory frame. (b) Show that Newton's second law is *not* valid in a reference frame moving past the laboratory frame with a constant acceleration.

3. The speed of the Earth in its orbit is 29.8 km/s. If that is the magnitude of the velocity $\vec{v}$ of the ether wind in Figure P39.3, find the angle ϕ between the velocity of light $\vec{c}$ in vacuum and the resultant velocity of light if there were an ether.

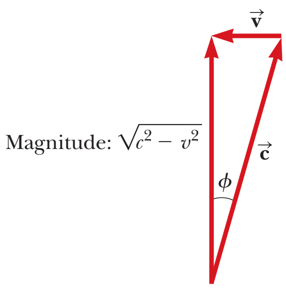


Figure P39.3

4. AMT A car of mass 2 000 kg moving with a speed of 20.0 m/s collides and locks together with a 1 500-kg car at rest at a stop sign. Show that momentum is conserved in a reference frame moving at 10.0 m/s in the direction of the moving car.

Section 39.2 The Michelson–Morley Experiment

Section 39.3 Einstein's Principle of Relativity

Section 39.4 Consequences of the Special Theory of Relativity

Problem 82 in Chapter 4 can be assigned with this section.

5. A star is 5.00 ly from the Earth. At what speed must a spacecraft travel on its journey to the star such that the Earth–star distance measured in the frame of the spacecraft is 2.00 ly?

6. A meterstick moving at $0.900c$ relative to the Earth's surface approaches an observer at rest with respect to the Earth's surface. (a) What is the meterstick's length as measured by the observer? (b) Qualitatively, how would the answer to part (a) change if the observer started running toward the meterstick?

7. W At what speed does a clock move if it is measured to run at a rate one-half the rate of a clock at rest with respect to an observer?

8. A muon formed high in the Earth's atmosphere is measured by an observer on the Earth's surface to travel at speed $v = 0.990c$ for a distance of 4.60 km before it decays into an electron, a neutrino, and an antineutrino ($\mu^- \rightarrow e^- + \nu + \bar{\nu}$). (a) For what time interval does the muon live as measured in its reference frame? (b) How far does the Earth travel as measured in the frame of the muon?

9. W How fast must a meterstick be moving if its length is measured to shrink to 0.500 m?

10. An astronaut is traveling in a space vehicle moving at $0.500c$ relative to the Earth. The astronaut measures her pulse rate at 75.0 beats per minute. Signals generated by the astronaut's pulse are radioed to the Earth when the vehicle is moving in a direction perpendicular to the line that connects the vehicle with an observer on the Earth. (a) What pulse rate does the Earth-based observer measure? (b) **What If?** What would be the pulse rate if the speed of the space vehicle were increased to $0.990c$?

11. A physicist drives through a stop light. When he is pulled over, he tells the police officer that the Doppler shift made the red light of wavelength 650 nm appear green to him, with a wavelength of 520 nm. The police officer writes out a traffic citation for speeding. How fast was the physicist traveling, according to his own testimony?

12. W A fellow astronaut passes by you in a spacecraft traveling at a high speed. The astronaut tells you that his craft is 20.0 m long and that the identical craft you are sitting in is 19.0 m long. According to your observations, (a) how long is your craft, (b) how long is the astronaut's craft, and (c) what is the speed of the astronaut's craft relative to your craft?

13. A deep-space vehicle moves away from the Earth with a speed of $0.800c$. An astronaut on the vehicle measures a time interval of 3.00 s to rotate her body through 1.00 rev as she floats in the vehicle. What time interval is required for this rotation according to an observer on the Earth?

14. For what value of v does $\gamma = 1.010\,0$? Observe that for speeds lower than this value, time dilation and length contraction are effects amounting to less than 1%.

15. A supertrain with a proper length of 100 m travels at a speed of $0.950c$ as it passes through a tunnel having a proper length of 50.0 m. As seen by a trackside observer, is the train ever completely within the tunnel? If so, by how much do the train's ends clear the ends of the tunnel?

16. M The average lifetime of a pi meson in its own frame of reference (i.e., the proper lifetime) is 2.6×10^{-8} s. If the meson moves with a speed of $0.98c$, what is (a) its mean lifetime as measured by an observer on Earth, and (b) the average distance it travels before decaying, as measured by an observer on Earth? (c) What distance would it travel if time dilation did not occur?

17. An astronomer on the Earth observes a meteoroid in the southern sky approaching the Earth at a speed of $0.800c$. At the time of its discovery the meteoroid is 20.0 ly from the Earth. Calculate (a) the time interval required for the meteoroid to reach the Earth as measured by the Earthbound astronomer, (b) this time interval as measured by a tourist on the meteoroid, and (c) the distance to the Earth as measured by the tourist.

18. A cube of steel has a volume of 1.00 cm^3 and a mass of 8.00 g when at rest on the Earth. If this cube is now given a speed $u = 0.900c$, what is its density as measured by a stationary observer? Note that relativistic density is defined as E_R/c^2V.

19. AMT M A spacecraft with a proper length of 300 m passes by an observer on the Earth. According to this observer, it takes 0.750 μs for the spacecraft to pass a fixed point. Determine the speed of the spacecraft as measured by the Earth-based observer.

20. A spacecraft with a proper length of L_p passes by an observer on the Earth. According to this observer, it takes a time interval Δt for the spacecraft to pass a fixed point. Determine the speed of the object as measured by the Earth-based observer.

21. A light source recedes from an observer with a speed v_S that is small compared with c. (a) Show that the fractional shift in the measured wavelength is given by the approximate expression

$$\frac{\Delta\lambda}{\lambda} \approx \frac{v_S}{c}$$

This phenomenon is known as the *redshift* because the visible light is shifted toward the red. (b) Spectroscopic measurements of light at $\lambda = 397$ nm coming from a galaxy in Ursa Major reveal a redshift of 20.0 nm. What is the recessional speed of the galaxy?

22. **Review.** In 1963, astronaut Gordon Cooper orbited the Earth 22 times. The press stated that for each orbit, he aged two-millionths of a second less than he would have had he remained on the Earth. (a) Assuming Cooper was 160 km above the Earth in a circular orbit, determine the difference in elapsed time between someone on the Earth and the orbiting astronaut for the 22 orbits. You may use the approximation

$$\frac{1}{\sqrt{1-x}} \approx 1 + \frac{x}{2}$$

for small x. (b) Did the press report accurate information? Explain.

23. Police radar detects the speed of a car (Fig. P39.23) as follows. Microwaves of a precisely known frequency are broadcast toward the car. The moving car reflects the microwaves with a Doppler shift. The reflected waves are received and combined with an attenuated version of the transmitted wave. Beats occur between the two microwave signals. The beat frequency is measured. (a) For an electromagnetic wave reflected back to its source from a mirror approaching at speed v, show that the reflected wave has frequency

$$f' = \frac{c+v}{c-v} f$$

where f is the source frequency. (b) Noting that v is much less than c, show that the beat frequency can be written as $f_{\text{beat}} = 2v/\lambda$. (c) What beat frequency is measured for a car speed of 30.0 m/s if the microwaves have frequency 10.0 GHz? (d) If the beat frequency measurement in part (c) is accurate to ±5.0 Hz, how accurate is the speed measurement?

© Cengage Learning/Ed Dodd

Figure P39.23

24. The identical twins Speedo and Goslo join a migration from the Earth to Planet X, 20.0 ly away in a reference frame in which both planets are at rest. The twins, of the same age, depart at the same moment on different spacecraft. Speedo's spacecraft travels steadily at $0.950c$ and Goslo's at $0.750c$. (a) Calculate the age difference between the twins after Goslo's spacecraft lands on Planet X. (b) Which twin is older?

25. An atomic clock moves at 1 000 km/h for 1.00 h as measured by an identical clock on the Earth. At the end of the 1.00-h interval, how many nanoseconds slow will the moving clock be compared with the Earth-based clock?

26. **Review.** An alien civilization occupies a planet circling a brown dwarf, several light-years away. The plane of the planet's orbit is perpendicular to a line from the brown dwarf to the Sun, so the planet is at nearly a fixed position relative to the Sun. The extraterrestrials have come to love broadcasts of *MacGyver*, on television channel 2, at carrier frequency 57.0 MHz. Their line of sight to us is in the plane of the Earth's orbit. Find the difference between the highest and lowest frequencies they receive due to the Earth's orbital motion around the Sun.

Section 39.5 The Lorentz Transformation Equations

27. A red light flashes at position $x_R = 3.00$ m and time $t_R = 1.00 \times 10^{-9}$ s, and a blue light flashes at $x_B = 5.00$ m and $t_B = 9.00 \times 10^{-9}$ s, all measured in the S reference frame. Reference frame S′ moves uniformly to the right and has its origin at the same point as S at $t = t' = 0$. Both flashes are observed to occur at the same place in S′. (a) Find the relative speed between S and S′. (b) Find the location of the two flashes in frame S′. (c) At what time does the red flash occur in the S′ frame?

28. Shannon observes two light pulses to be emitted from the same location, but separated in time by 3.00 μs. Kimmie observes the emission of the same two pulses to be separated in time by 9.00 μs. (a) How fast is Kimmie moving relative to Shannon? (b) According to Kimmie, what is the separation in space of the two pulses?

29. A moving rod is observed to have a length of $\ell = 2.00$ m and to be oriented at an angle of $\theta = 30.0°$ with respect to the direction of motion as shown in Figure P39.29. The rod has a speed of $0.995c$. (a) What is the proper length of the rod? (b) What is the orientation angle in the proper frame?

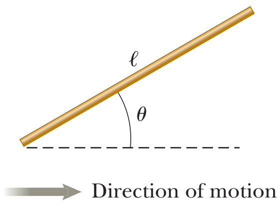


Figure P39.29

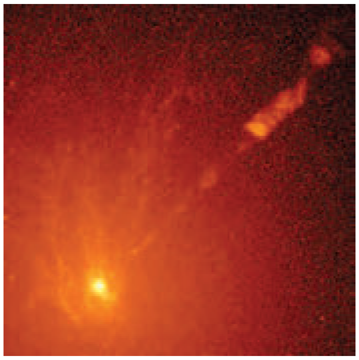
30. A rod moving with a speed v along the horizontal direction is observed to have length ℓ and to make an angle θ with respect to the direction of motion as shown in Figure P39.29. (a) Show that the length of the rod as measured by an observer at rest with respect to the rod is $\ell_p = \ell[1 - (v^2/c^2)\cos^2\theta]^{1/2}$. (b) Show that the angle θ_p that the rod makes with the x axis according to an observer at rest with respect to the rod can be found from $\tan\theta_p = \gamma \tan\theta$. These results show that the rod is observed to be both contracted and rotated. (Take the lower end of the rod to be at the origin of the coordinate system in which the rod is at rest.)

31. Keilah, in reference frame S, measures two events to be simultaneous. Event A occurs at the point (50.0 m, 0, 0) at the instant 9:00:00 Universal time on January 15,

2013. Event B occurs at the point (150 m, 0, 0) at the same moment. Torrey, moving past with a velocity of $0.800c\hat{\mathbf{i}}$, also observes the two events. In her reference frame S′, which event occurred first and what time interval elapsed between the events?

Section 39.6 The Lorentz Velocity Transformation Equations

32. Figure P39.32 shows a jet of material (at the upper right) being ejected by galaxy M87 (at the lower left). Such jets are believed to be evidence of supermassive black holes at the center of a galaxy. Suppose two jets of material from the center of a galaxy are ejected in opposite directions. Both jets move at $0.750c$ relative to the galaxy center. Determine the speed of one jet relative to the other.



NASA/STSci

Figure P39.32

33. An enemy spacecraft moves away from the Earth at a speed of $v = 0.800c$ (Fig. P39.33). A galactic patrol spacecraft pursues at a speed of $u = 0.900c$ relative to the Earth. Observers on the Earth measure the patrol craft to be overtaking the enemy craft at a relative speed of $0.100c$. With what speed is the patrol craft overtaking the enemy craft as measured by the patrol craft's crew?

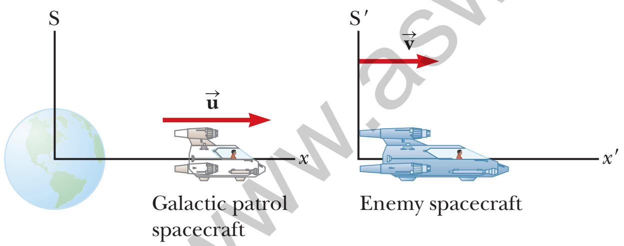


Figure P39.33

34. A spacecraft is launched from the surface of the Earth with a velocity of $0.600c$ at an angle of 50.0° above the horizontal positive x axis. Another spacecraft is moving past with a velocity of $0.700c$ in the negative x direction. Determine the magnitude and direction of the velocity of the first spacecraft as measured by the pilot of the second spacecraft.

35. A rocket moves with a velocity of $0.92c$ to the right with respect to a stationary observer A. An observer B moving relative to observer A finds that the rocket is moving with a velocity of $0.95c$ to the left. What is the velocity of observer B relative to observer A? (*Hint:* Consider observer B's velocity in the frame of reference of the rocket.)

Section 39.7 Relativistic Linear Momentum

36. Calculate the momentum of an electron moving with a speed of (a) $0.010\,0c$, (b) $0.500c$, and (c) $0.900c$.

37. An electron has a momentum that is three times larger than its classical momentum. (a) Find the speed of the electron. (b) **What If?** How would your result change if the particle were a proton?

38. Show that the speed of an object having momentum of magnitude p and mass m is

$$u = \frac{c}{\sqrt{1 + (mc/p)^2}}$$

39. (a) Calculate the classical momentum of a proton traveling at $0.990c$, neglecting relativistic effects. (b) Repeat the calculation while including relativistic effects. (c) Does it make sense to neglect relativity at such speeds?

40. The speed limit on a certain roadway is 90.0 km/h. Suppose speeding fines are made proportional to the amount by which a vehicle's momentum exceeds the momentum it would have when traveling at the speed limit. The fine for driving at 190 km/h (that is, 100 km/h over the speed limit) is \$80.0. What, then, is the fine for traveling (a) at 1 090 km/h? (b) At 1 000 000 090 km/h?

41. A golf ball travels with a speed of 90.0 m/s. By what fraction does its relativistic momentum magnitude p differ from its classical value mu? That is, find the ratio $(p - mu)/mu$.

42. The nonrelativistic expression for the momentum of a particle, $p = mu$, agrees with experiment if $u << c$. For what speed does the use of this equation give an error in the measured momentum of (a) 1.00% and (b) 10.0%?

43. An unstable particle at rest spontaneously breaks into two fragments of unequal mass. The mass of the first fragment is 2.50×10^{-28} kg, and that of the other is 1.67×10^{-27} kg. If the lighter fragment has a speed of $0.893c$ after the breakup, what is the speed of the heavier fragment?

Section 39.8 Relativistic Energy

44. Determine the energy required to accelerate an electron from (a) $0.500c$ to $0.900c$ and (b) $0.900c$ to $0.990c$.

45. An electron has a kinetic energy five times greater than its rest energy. Find (a) its total energy and (b) its speed.

46. Protons in an accelerator at the Fermi National Laboratory near Chicago are accelerated to a total energy that is 400 times their rest energy. (a) What is the speed of these protons in terms of c? (b) What is their kinetic energy in MeV?

47. A proton moves at $0.950c$. Calculate its (a) rest energy, (b) total energy, and (c) kinetic energy.

48. (a) Find the kinetic energy of a 78.0-kg spacecraft launched out of the solar system with speed 106 km/s by using the classical equation $K = \frac{1}{2}mu^2$. (b) **What If?** Calculate its kinetic energy using the relativistic equation. (c) Explain the result of comparing the answers of parts (a) and (b).

49. AMT M A proton in a high-energy accelerator moves with a speed of $c/2$. Use the work–kinetic energy theorem to find the work required to increase its speed to (a) $0.750c$ and (b) $0.995c$.

50. Show that for any object moving at less than one-tenth the speed of light, the relativistic kinetic energy agrees with the result of the classical equation $K = \frac{1}{2}mu^2$ to within less than 1%. Therefore, for most purposes, the classical equation is sufficient to describe these objects.

51. The total energy of a proton is twice its rest energy. Find the momentum of the proton in MeV/c units.

52. Consider electrons accelerated to a total energy of 20.0 GeV in the 3.00-km-long Stanford Linear Accelerator. (a) What is the factor γ for the electrons? (b) What is the electrons' speed at the given energy? (c) What is the length of the accelerator in the electrons' frame of reference when they are moving at their highest speed?

53. W When 1.00 g of hydrogen combines with 8.00 g of oxygen, 9.00 g of water is formed. During this chemical reaction, 2.86×10^5 J of energy is released. (a) Is the mass of the water larger or smaller than the mass of the reactants? (b) What is the difference in mass? (c) Explain whether the change in mass is likely to be detectable.

54. W In a nuclear power plant, the fuel rods last 3 yr before they are replaced. The plant can transform energy at a maximum possible rate of 1.00 GW. Supposing it operates at 80.0% capacity for 3.00 yr, what is the loss of mass of the fuel?

55. The power output of the Sun is 3.85×10^{26} W. By how much does the mass of the Sun decrease each second?

56. A gamma ray (a high-energy photon) can produce an electron (e^-) and a positron (e^+) of equal mass when it enters the electric field of a heavy nucleus: $\gamma \rightarrow e^+ + e^-$. What minimum gamma-ray energy is required to accomplish this task?

57. A spaceship of mass 2.40×10^6 kg is to be accelerated to a speed of $0.700c$. (a) What minimum amount of energy does this acceleration require from the spaceship's fuel, assuming perfect efficiency? (b) How much fuel would it take to provide this much energy if all the rest energy of the fuel could be transformed to kinetic energy of the spaceship?

58. Show that the energy–momentum relationship in Equation 39.27, $E^2 = p^2c^2 + (mc^2)^2$, follows from the expressions $E = \gamma mc^2$ and $p = \gamma mu$.

59. The rest energy of an electron is 0.511 MeV. The rest energy of a proton is 938 MeV. Assume both particles have kinetic energies of 2.00 MeV. Find the speed of (a) the electron and (b) the proton. (c) By what factor does the speed of the electron exceed that of the proton? (d) Repeat the calculations in parts (a) through (c) assuming both particles have kinetic energies of 2 000 MeV.

60. Consider a car moving at highway speed u. Is its actual kinetic energy larger or smaller than $\frac{1}{2}mu^2$? Make an order-of-magnitude estimate of the amount by which its actual kinetic energy differs from $\frac{1}{2}mu^2$. In your solution, state the quantities you take as data and the values you measure or estimate for them. You may find Appendix B.5 useful.

61. M A pion at rest ($m_\pi = 273m_e$) decays to a muon ($m_\mu = 207m_e$) and an antineutrino ($m_{\bar{\nu}} \approx 0$). The reaction is written $\pi^- \rightarrow \mu^- + \bar{\nu}$. Find (a) the kinetic energy of the muon and (b) the energy of the antineutrino in electron volts.

62. GP An unstable particle with mass $m = 3.34 \times 10^{-27}$ kg is initially at rest. The particle decays into two fragments that fly off along the x axis with velocity components $u_1 = 0.987c$ and $u_2 = -0.868c$. From this information, we wish to determine the masses of fragments 1 and 2. (a) Is the initial system of the unstable particle, which becomes the system of the two fragments, isolated or nonisolated? (b) Based on your answer to part (a), what two analysis models are appropriate for this situation? (c) Find the values of γ for the two fragments after the decay. (d) Using one of the analysis models in part (b), find a relationship between the masses m_1 and m_2 of the fragments. (e) Using the second analysis model in part (b), find a second relationship between the masses m_1 and m_2. (f) Solve the relationships in parts (d) and (e) simultaneously for the masses m_1 and m_2.

63. Massive stars ending their lives in supernova explosions produce the nuclei of all the atoms in the bottom half of the periodic table by fusion of smaller nuclei. This problem roughly models that process. A particle of mass $m = 1.99 \times 10^{-26}$ kg moving with a velocity $\vec{\mathbf{u}} = 0.500c\hat{\mathbf{i}}$ collides head-on and sticks to a particle of mass $m' = m/3$ moving with the velocity $\vec{\mathbf{u}} = -0.500c\hat{\mathbf{i}}$. What is the mass of the resulting particle?

64. Massive stars ending their lives in supernova explosions produce the nuclei of all the atoms in the bottom half of the periodic table by fusion of smaller nuclei. This problem roughly models that process. A particle of mass m moving along the x axis with a velocity component $+u$ collides head-on and sticks to a particle of mass $m/3$ moving along the x axis with the velocity component $-u$. (a) What is the mass M of the resulting particle? (b) Evaluate the expression from part (a) in the limit $u \rightarrow 0$. (c) Explain whether the result agrees with what you should expect from nonrelativistic physics.

Section 39.9 The General Theory of Relativity

65. **Review.** A global positioning system (GPS) satellite moves in a circular orbit with period 11 h 58 min. (a) Determine the radius of its orbit. (b) Determine its speed. (c) The nonmilitary GPS signal is broadcast at a frequency of 1 575.42 MHz in the reference frame of the satellite. When it is received on the Earth's surface by a GPS receiver (Fig. P39.65 on page 1230), what is the fractional change in this frequency due to time dilation as described by special relativity? (d) The gravitational "blueshift" of the frequency according to general relativity is a separate effect. It is called a blueshift to indicate a change to a higher frequency. The magnitude of that fractional change is given by

$$\frac{\Delta f}{f} = \frac{\Delta U_g}{mc^2}$$

where U_g is the change in gravitational potential energy of an object–Earth system when the object of mass m is moved between the two points where the signal is observed. Calculate this fractional change in frequency due to the change in position of the satellite from the Earth's surface to its orbital position. (e) What is the overall fractional change in frequency due to both time dilation and gravitational blueshift?

Figure P39.65

Additional Problems

66. An electron has a speed of $0.750c$. (a) Find the speed of a proton that has the same kinetic energy as the electron. (b) **What If?** Find the speed of a proton that has the same momentum as the electron.

67. M The net nuclear fusion reaction inside the Sun can be written as $4\,{}^1\text{H} \rightarrow {}^4\text{He} + E$. The rest energy of each hydrogen atom is 938.78 MeV, and the rest energy of the helium-4 atom is 3 728.4 MeV. Calculate the percentage of the starting mass that is transformed to other forms of energy.

68. *Why is the following situation impossible?* On their 40th birthday, twins Speedo and Goslo say good-bye as Speedo takes off for a planet that is 50 ly away. He travels at a constant speed of $0.85c$ and immediately turns around and comes back to the Earth after arriving at the planet. Upon arriving back at the Earth, Speedo has a joyous reunion with Goslo.

69. A Doppler weather radar station broadcasts a pulse of radio waves at frequency 2.85 GHz. From a relatively small batch of raindrops at bearing 38.6° east of north, the station receives a reflected pulse after 180 μs with a frequency shifted upward by 254 Hz. From a similar batch of raindrops at bearing 39.6° east of north, the station receives a reflected pulse after the same time delay, with a frequency shifted downward by 254 Hz. These pulses have the highest and lowest frequencies the station receives. (a) Calculate the radial velocity components of both batches of raindrops. (b) Assume that these raindrops are swirling in a uniformly rotating vortex. Find the angular speed of their rotation.

70. An object having mass 900 kg and traveling at speed $0.850c$ collides with a stationary object having mass 1 400 kg. The two objects stick together. Find (a) the speed and (b) the mass of the composite object.

71. M An astronaut wishes to visit the Andromeda galaxy, making a one-way trip that will take 30.0 years in the spaceship's frame of reference. Assume the galaxy is 2.00 million light-years away and his speed is constant. (a) How fast must he travel relative to Earth? (b) What will be the kinetic energy of his spacecraft, which has mass of 1.00×10^6 kg? (c) What is the cost of this energy if it is purchased at a typical consumer price for electric energy, 13.0¢ per kWh? The following approximation will prove useful:

$$\frac{1}{\sqrt{1+x}} \approx 1 - \frac{x}{2} \quad \text{for } x << 1$$

72. A physics professor on the Earth gives an exam to her students, who are in a spacecraft traveling at speed v relative to the Earth. The moment the craft passes the professor, she signals the start of the exam. She wishes her students to have a time interval T_0 (spacecraft time) to complete the exam. Show that she should wait a time interval (Earth time) of

$$T = T_0\sqrt{\frac{1 - v/c}{1 + v/c}}$$

before sending a light signal telling them to stop. (*Suggestion:* Remember that it takes some time for the second light signal to travel from the professor to the students.)

73. M An interstellar space probe is launched from Earth. After a brief period of acceleration, it moves with a constant velocity, 70.0% of the speed of light. Its nuclear-powered batteries supply the energy to keep its data transmitter active continuously. The batteries have a lifetime of 15.0 years as measured in a rest frame. (a) How long do the batteries on the space probe last as measured by mission control on Earth? (b) How far is the probe from Earth when its batteries fail as measured by mission control? (c) How far is the probe from Earth as measured by its built-in trip odometer when its batteries fail? (d) For what total time after launch are data received from the probe by mission control? Note that radio waves travel at the speed of light and fill the space between the probe and Earth at the time the battery fails.

74. The equation

$$K = \left(\frac{1}{\sqrt{1 - u^2/c^2}} - 1\right)mc^2$$

gives the kinetic energy of a particle moving at speed u. (a) Solve the equation for u. (b) From the equation for u, identify the minimum possible value of speed and the corresponding kinetic energy. (c) Identify the maximum possible speed and the corresponding kinetic energy. (d) Differentiate the equation for u with respect to time to obtain an equation describing the acceleration of a particle as a function of its kinetic energy and the power input to the particle. (e) Observe that for a nonrelativistic particle we have $u = (2K/m)^{1/2}$ and that differentiating this equation with respect to time gives $a = P/(2mK)^{1/2}$. State the limiting form of the expression in part (d) at low energy. State how it compares with the nonrelativistic expression. (f) State the limiting form of the expression in part (d) at high energy. (g) Consider a particle with constant input power. Explain how the answer to part (f) helps account for the answer to part (c).

75. Consider the astronaut planning the trip to Andromeda in Problem 71. (a) To three significant figures, what is the value for γ for the speed found in part (a) of Problem 71? (b) Just as the astronaut leaves on his constant-speed trip, a light beam is also sent in the direction of Andromeda. According to the Earth observer, how much later does the astronaut arrive at Andromeda after the arrival of the light beam?

76. An object disintegrates into two fragments. One fragment has mass 1.00 MeV/c^2 and momentum 1.75 MeV/c in the positive x direction, and the other has mass 1.50 MeV/c^2 and momentum 2.00 MeV/c in the positive y direction. Find (a) the mass and (b) the speed of the original object.

77. M The cosmic rays of highest energy are protons that have kinetic energy on the order of 10^{13} MeV. (a) As measured in the proton's frame, what time interval would a proton of this energy require to travel across the Milky Way galaxy, which has a proper diameter $\sim 10^5$ ly? (b) From the point of view of the proton, how many kilometers across is the galaxy?

78. M Spacecraft I, containing students taking a physics exam, approaches the Earth with a speed of 0.600c (relative to the Earth), while spacecraft II, containing professors proctoring the exam, moves at 0.280c (relative to the Earth) directly toward the students. If the professors stop the exam after 50.0 min have passed on their clock, for what time interval does the exam last as measured by (a) the students and (b) an observer on the Earth?

79. **Review.** Around the core of a nuclear reactor shielded by a large pool of water, Cerenkov radiation appears as a blue glow. (See Fig. P17.38 on page 528.) Cerenkov radiation occurs when a particle travels faster through a medium than the speed of light in that medium. It is the electromagnetic equivalent of a bow wave or a sonic boom. An electron is traveling through water at a speed 10.0% faster than the speed of light in water. Determine the electron's (a) total energy, (b) kinetic energy, and (c) momentum. (d) Find the angle between the shock wave and the electron's direction of motion.

80. The motion of a transparent medium influences the speed of light. This effect was first observed by Fizeau in 1851. Consider a light beam in water. The water moves with speed v in a horizontal pipe. Assume the light travels in the same direction as the water moves. The speed of light with respect to the water is c/n, where $n = 1.33$ is the index of refraction of water. (a) Use the velocity transformation equation to show that the speed of the light measured in the laboratory frame is

$$u = \frac{c}{n}\left(\frac{1 + nv/c}{1 + v/nc}\right)$$

(b) Show that for $v << c$, the expression from part (a) becomes, to a good approximation,

$$u \approx \frac{c}{n} + v - \frac{v}{n^2}$$

(c) Argue for or against the view that we should expect the result to be $u = (c/n) + v$ according to the Galilean transformation and that the presence of the term $-v/n^2$ represents a relativistic effect appearing even at "nonrelativistic" speeds. (d) Evaluate u in the limit as the speed of the water approaches c.

81. Imagine that the entire Sun, of mass M_S, collapses to a sphere of radius R_g such that the work required to remove a small mass m from the surface would be equal to its rest energy mc^2. This radius is called the *gravitational radius* for the Sun. (a) Use this approach to show that $R_g = GM_S/c^2$. (b) Find a numerical value for R_g.

82. *Why is the following situation impossible?* An experimenter is accelerating electrons for use in probing a material. She finds that when she accelerates them through a potential difference of 84.0 kV, the electrons have half the speed she wishes. She quadruples the potential difference to 336 kV, and the electrons accelerated through this potential difference have her desired speed.

83. An alien spaceship traveling at 0.600c toward the Earth launches a landing craft. The landing craft travels in the same direction with a speed of 0.800c relative to the mother ship. As measured on the Earth, the spaceship is 0.200 ly from the Earth when the landing craft is launched. (a) What speed do the Earth-based observers measure for the approaching landing craft? (b) What is the distance to the Earth at the moment of the landing craft's launch as measured by the aliens? (c) What travel time is required for the landing craft to reach the Earth as measured by the aliens on the mother ship? (d) If the landing craft has a mass of 4.00×10^5 kg, what is its kinetic energy as measured in the Earth reference frame?

84. (a) Prepare a graph of the relativistic kinetic energy and the classical kinetic energy, both as a function of speed, for an object with a mass of your choice. (b) At what speed does the classical kinetic energy underestimate the experimental value by 1%? (c) By 5%? (d) By 50%?

85. AMT An observer in a coasting spacecraft moves toward a mirror at speed $v = 0.650c$ relative to the reference frame labeled S in Figure P39.85 (page 1232). The mirror is stationary with respect to S. A light pulse emitted by the spacecraft travels toward the mirror and is reflected back to the spacecraft. The spacecraft is a distance $d = 5.66 \times 10^{10}$ m from the mirror (as measured by observers in S) at the moment the light pulse leaves the spacecraft. What is the total travel time of the pulse as measured by observers in (a) the S frame and (b) the spacecraft?

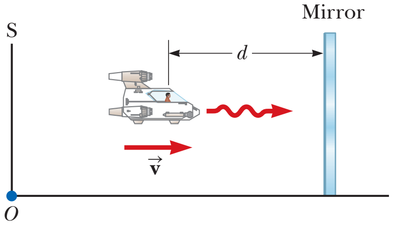


Figure P39.85 Problems 85 and 86.

86. An observer in a coasting spacecraft moves toward a mirror at speed v relative to the reference frame labeled S in Figure P39.85. The mirror is stationary with respect to S. A light pulse emitted by the spacecraft travels toward the mirror and is reflected back to the spacecraft. The spacecraft is a distance d from the mirror (as measured by observers in S) at the moment the light pulse leaves the spacecraft. What is the total travel time of the pulse as measured by observers in (a) the S frame and (b) the spacecraft?

87. A ^{57}Fe nucleus at rest emits a 14.0-keV photon. Use conservation of energy and momentum to find the kinetic energy of the recoiling nucleus in electron volts. Use $Mc^2 = 8.60 \times 10^{-9}$ J for the final state of the ^{57}Fe nucleus.

Challenge Problems

88. A particle with electric charge q moves along a straight line in a uniform electric field $\vec{\mathbf{E}}$ with speed u. The electric force exerted on the charge is $q\vec{\mathbf{E}}$. The velocity of the particle and the electric field are both in the x direction. (a) Show that the acceleration of the particle in the x direction is given by

$$a = \frac{du}{dt} = \frac{qE}{m}\left(1 - \frac{u^2}{c^2}\right)^{3/2}$$

(b) Discuss the significance of the dependence of the acceleration on the speed. (c) **What If?** If the particle starts from rest at $x = 0$ at $t = 0$, how would you proceed to find the speed of the particle and its position at time t?

89. The creation and study of new and very massive elementary particles is an important part of contemporary physics. To create a particle of mass M requires an energy Mc^2. With enough energy, an exotic particle can be created by allowing a fast-moving proton to collide with a similar target particle. Consider a perfectly inelastic collision between two protons: an incident proton with mass m_p, kinetic energy K, and momentum magnitude p joins with an originally stationary target proton to form a single product particle of mass M. Not all the kinetic energy of the incoming proton is available to create the product particle because conservation of momentum requires that the system as a whole still must have some kinetic energy after the collision. Therefore, only a fraction of the energy of the incident particle is available to create a new particle. (a) Show that the energy available to create a product particle is given by

$$Mc^2 = 2m_pc^2\sqrt{1 + \frac{K}{2m_pc^2}}$$

This result shows that when the kinetic energy K of the incident proton is large compared with its rest energy m_pc^2, then M approaches $(2m_pK)^{1/2}/c$. Therefore, if the energy of the incoming proton is increased by a factor of 9, the mass you can create increases only by a factor of 3, not by a factor of 9 as would be expected. (b) This problem can be alleviated by using *colliding beams* as is the case in most modern accelerators. Here the total momentum of a pair of interacting particles can be zero. The center of mass can be at rest after the collision, so, in principle, all the initial kinetic energy can be used for particle creation. Show that

$$Mc^2 = 2mc^2\left(1 + \frac{K}{mc^2}\right)$$

where K is the kinetic energy of each of the two identical colliding particles. Here, if $K >> mc^2$, we have M directly proportional to K as we would desire.

90. Suppose our Sun is about to explode. In an effort to escape, we depart in a spacecraft at $v = 0.800c$ and head toward the star Tau Ceti, 12.0 ly away. When we reach the midpoint of our journey from the Earth, we see our Sun explode, and, unfortunately, at the same instant, we see Tau Ceti explode as well. (a) In the spacecraft's frame of reference, should we conclude that the two explosions occurred simultaneously? If not, which occurred first? (b) **What If?** In a frame of reference in which the Sun and Tau Ceti are at rest, did they explode simultaneously? If not, which exploded first?

91. Owen and Dina are at rest in frame S′, which is moving at $0.600c$ with respect to frame S. They play a game of catch while Ed, at rest in frame S, watches the action (Fig. P39.91). Owen throws the ball to Dina at $0.800c$ (according to Owen), and their separation (measured in S′) is equal to 1.80×10^{12} m. (a) According to Dina, how fast is the ball moving? (b) According to Dina, what time interval is required for the ball to reach her? According to Ed, (c) how far apart are Owen and Dina, (d) how fast is the ball moving, and (e) what time interval is required for the ball to reach Dina?

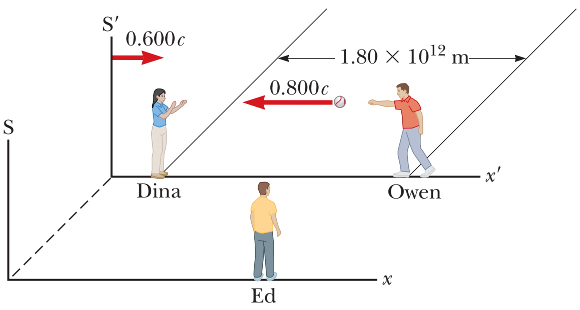
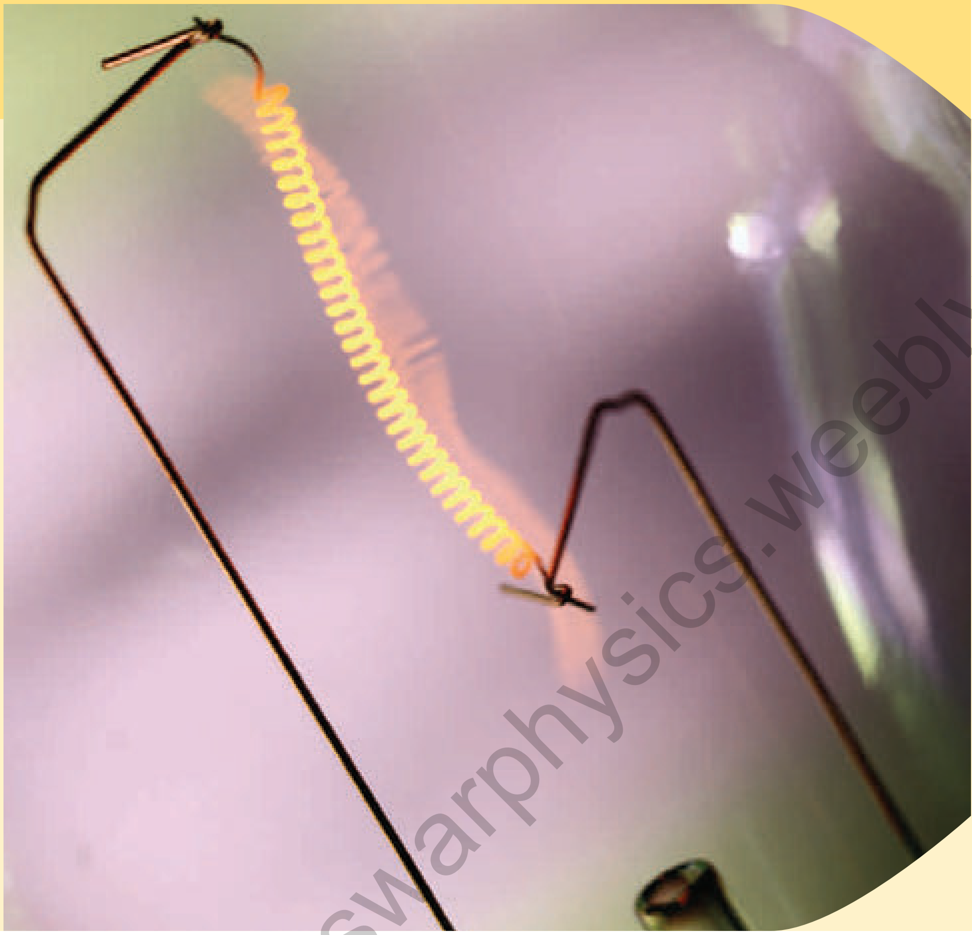


Figure P39.91

Introduction to Quantum Physics

CHAPTER

40



This lightbulb filament glows with an orange color. Why? Classical physics is unable to explain the experimentally observed wavelength distribution of electromagnetic radiation from a hot object. A theory proposed in 1900 and describing the radiation from such objects represents the dawn of quantum physics. *(Steve Cole/Getty Images)*

In Chapter 39, we discussed that Newtonian mechanics must be replaced by Einstein's special theory of relativity when dealing with particle speeds comparable to the speed of light. As the 20th century progressed, many experimental and theoretical problems were resolved by the special theory of relativity. For many other problems, however, neither relativity nor classical physics could provide a theoretical answer. Attempts to apply the laws of classical physics to explain the behavior of matter on the atomic scale were consistently unsuccessful. For example, the emission of discrete wavelengths of light from atoms in a high-temperature gas could not be explained within the framework of classical physics.

As physicists sought new ways to solve these puzzles, another revolution took place in physics between 1900 and 1930. A new theory called *quantum mechanics* was highly successful in explaining the behavior of particles of microscopic size. Like the special theory of relativity, the quantum theory requires a modification of our ideas concerning the physical world.

The first explanation of a phenomenon using quantum theory was introduced by Max Planck. Many subsequent mathematical developments and interpretations were made by a number of distinguished physicists, including Einstein, Bohr, de Broglie, Schrödinger, and

Pitfall Prevention 40.1
Expect to Be Challenged If the discussions of quantum physics in this and subsequent chapters seem strange and confusing to you, it's because your whole life experience has taken place in the macroscopic world, where quantum effects are not evident.

Heisenberg. Despite the great success of the quantum theory, Einstein frequently played the role of its critic, especially with regard to the manner in which the theory was interpreted.

Because an extensive study of quantum theory is beyond the scope of this book, this chapter is simply an introduction to its underlying principles.

40.1 Blackbody Radiation and Planck's Hypothesis

An object at any temperature emits electromagnetic waves in the form of **thermal radiation** from its surface as discussed in Section 20.7. The characteristics of this radiation depend on the temperature and properties of the object's surface. Careful study shows that the radiation consists of a continuous distribution of wavelengths from all portions of the electromagnetic spectrum. If the object is at room temperature, the wavelengths of thermal radiation are mainly in the infrared region and hence the radiation is not detected by the human eye. As the surface temperature of the object increases, the object eventually begins to glow visibly red, like the coils of a toaster. At sufficiently high temperatures, the glowing object appears white, as in the hot tungsten filament of an incandescent lightbulb.

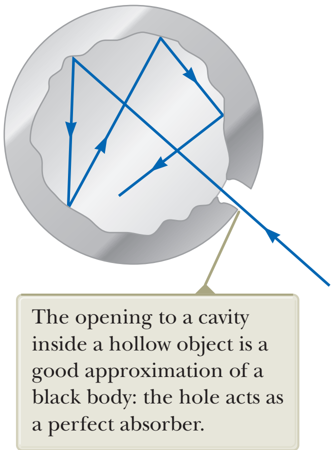


Figure 40.1 A physical model of a black body.

From a classical viewpoint, thermal radiation originates from accelerated charged particles in the atoms near the surface of the object; those charged particles emit radiation much as small antennas do. The thermally agitated particles can have a distribution of energies, which accounts for the continuous spectrum of radiation emitted by the object. By the end of the 19th century, however, it became apparent that the classical theory of thermal radiation was inadequate. The basic problem was in understanding the observed distribution of wavelengths in the radiation emitted by a black body. As defined in Section 20.7, a **black body** is an ideal system that absorbs all radiation incident on it. The electromagnetic radiation emitted by the black body is called **blackbody radiation.**

A good approximation of a black body is a small hole leading to the inside of a hollow object as shown in Figure 40.1. Any radiation incident on the hole from outside the cavity enters the hole and is reflected a number of times on the interior walls of the cavity; hence, the hole acts as a perfect absorber. The nature of the radiation leaving the cavity through the hole depends only on the temperature of the cavity walls and not on the material of which the walls are made. The spaces between lumps of hot charcoal (Fig. 40.2) emit light that is very much like blackbody radiation.

SOMMAI/Shutterstock.com

Figure 40.2 The glow emanating from the spaces between these hot charcoal briquettes is, to a close approximation, blackbody radiation. The color of the light depends only on the temperature of the briquettes.

The radiation emitted by oscillators in the cavity walls in Figure 40.1 experiences boundary conditions and can be analyzed using the waves under boundary conditions analysis model. As the radiation reflects from the cavity's walls, standing electromagnetic waves are established within the three-dimensional interior of the cavity. Many standing-wave modes are possible, and the distribution of the energy in the cavity among these modes determines the wavelength distribution of the radiation leaving the cavity through the hole.

The wavelength distribution of radiation from cavities was studied experimentally in the late 19th century. Figure 40.3 shows how the intensity of blackbody radiation varies with temperature and wavelength. The following two consistent experimental findings were seen as especially significant:

1. **The total power of the emitted radiation increases with temperature.** We discussed this behavior briefly in Chapter 20, where we introduced **Stefan's law:**

Stefan's law ▶

$$P = \sigma A e T^4 \tag{40.1}$$

where P is the power in watts radiated at all wavelengths from the surface of an object, $\sigma = 5.670 \times 10^{-8}\ \mathrm{W/m^2 \cdot K^4}$ is the Stefan–Boltzmann constant, A is the surface area of the object in square meters, e is the emissivity of the surface, and T is the surface temperature in kelvins. For a black body, the emissivity is $e = 1$ exactly.

2. **The peak of the wavelength distribution shifts to shorter wavelengths as the temperature increases.** This behavior is described by the following relationship, called **Wien's displacement law:**

$$\lambda_{\max} T = 2.898 \times 10^{-3} \text{ m} \cdot \text{K} \tag{40.2}$$

◀ **Wien's displacement law**

where $\lambda_{\max}$ is the wavelength at which the curve peaks and T is the absolute temperature of the surface of the object emitting the radiation. The wavelength at the curve's peak is inversely proportional to the absolute temperature; that is, as the temperature increases, the peak is "displaced" to shorter wavelengths (Fig. 40.3).

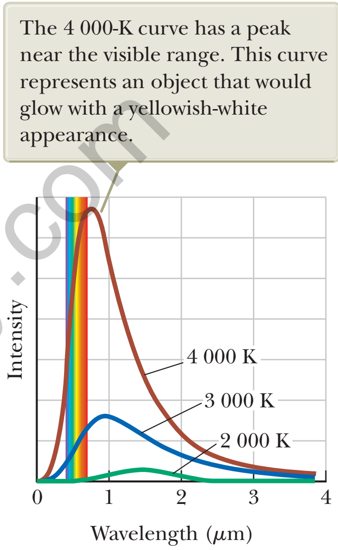


Figure 40.3 Intensity of blackbody radiation versus wavelength at three temperatures. The visible range of wavelengths is between 0.4 μm and 0.7 μm. At approximately 6 000 K, the peak is in the center of the visible wavelengths and the object appears white.

Wien's displacement law is consistent with the behavior of the object mentioned at the beginning of this section. At room temperature, the object does not appear to glow because the peak is in the infrared region of the electromagnetic spectrum. At higher temperatures, it glows red because the peak is in the near infrared with some radiation at the red end of the visible spectrum, and at still higher temperatures, it glows white because the peak is in the visible so that all colors are emitted.

Quick Quiz 40.1 Figure 40.4 shows two stars in the constellation Orion. Betelgeuse appears to glow red, whereas Rigel looks blue in color. Which star has a higher surface temperature? **(a)** Betelgeuse **(b)** Rigel **(c)** both the same **(d)** impossible to determine

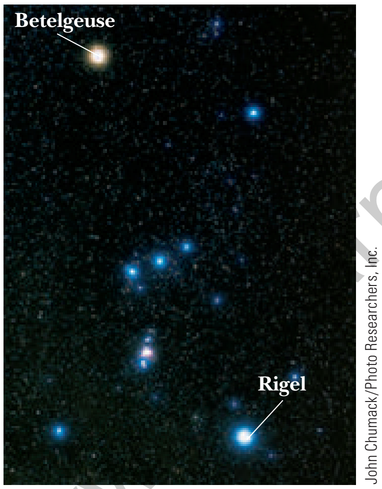


Figure 40.4 (Quick Quiz 40.1) Which star is hotter, Betelgeuse or Rigel?

A successful theory for blackbody radiation must predict the shape of the curves in Figure 40.3, the temperature dependence expressed in Stefan's law, and the shift of the peak with temperature described by Wien's displacement law. Early attempts to use classical ideas to explain the shapes of the curves in Figure 40.3 failed.

Let's consider one of these early attempts. To describe the distribution of energy from a black body, we define $I(\lambda, T)\, d\lambda$ to be the intensity, or power per unit area, emitted in the wavelength interval $d\lambda$. The result of a calculation based on a classical theory of blackbody radiation known as the **Rayleigh–Jeans law** is

$$I(\lambda, T) = \frac{2\pi c k_B T}{\lambda^4} \tag{40.3}$$

◀ **Rayleigh–Jeans law**

where k_B is Boltzmann's constant. The black body is modeled as the hole leading into a cavity (Fig. 40.1), resulting in many modes of oscillation of the electromagnetic field caused by accelerated charges in the cavity walls and the emission of electromagnetic waves at all wavelengths. In the classical theory used to derive

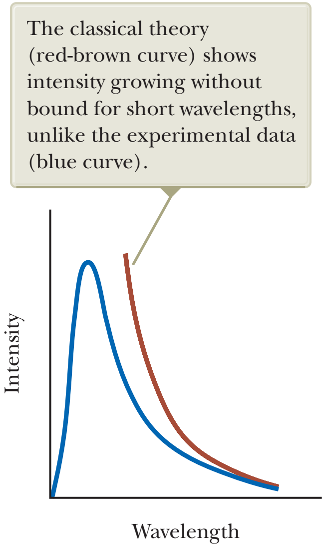


Figure 40.5 Comparison of experimental results and the curve predicted by the Rayleigh–Jeans law for the distribution of blackbody radiation.

Equation 40.3, the average energy for each wavelength of the standing-wave modes is assumed to be proportional to $k_B T$, based on the theorem of equipartition of energy discussed in Section 21.1.

An experimental plot of the blackbody radiation spectrum, together with the theoretical prediction of the Rayleigh–Jeans law, is shown in Figure 40.5. At long wavelengths, the Rayleigh–Jeans law is in reasonable agreement with experimental data, but at short wavelengths, major disagreement is apparent.

As λ approaches zero, the function $I(\lambda, T)$ given by Equation 40.3 approaches infinity. Hence, according to classical theory, not only should short wavelengths predominate in a blackbody spectrum, but also the energy emitted by any black body should become infinite in the limit of zero wavelength. In contrast to this prediction, the experimental data plotted in Figure 40.5 show that as λ approaches zero, $I(\lambda, T)$ also approaches zero. This mismatch of theory and experiment was so disconcerting that scientists called it the *ultraviolet catastrophe.* (This "catastrophe"—infinite energy—occurs as the wavelength approaches zero; the word *ultraviolet* was applied because ultraviolet wavelengths are short.)

In 1900, Max Planck developed a theory of blackbody radiation that leads to an equation for $I(\lambda, T)$ that is in complete agreement with experimental results at all wavelengths. In discussing this theory, we use the outline of properties of structural models introduced in Chapter 21:

1. *Physical components:*
   Planck assumed the cavity radiation came from atomic oscillators in the cavity walls in Figure 40.1.
2. *Behavior of the components:*
   (a) The energy of an oscillator can have only certain *discrete* values E_n:

$$E_n = nhf \tag{40.4}$$

   where n is a positive integer called a **quantum number,**[1] f is the oscillator's frequency, and h is a parameter Planck introduced that is now called **Planck's constant.** Because the energy of each oscillator can have only discrete values given by Equation 40.4, we say the energy is **quantized.** Each discrete energy value corresponds to a different **quantum state,** represented by the quantum number n. When the oscillator is in the $n = 1$ quantum state, its energy is hf; when it is in the $n = 2$ quantum state, its energy is $2hf$; and so on.

   (b) The oscillators emit or absorb energy when making a transition from one quantum state to another. The entire energy difference between the initial and final states in the transition is emitted or absorbed as a single quantum of radiation. If the transition is from one state to a lower adjacent state—say, from the $n = 3$ state to the $n = 2$ state—Equation 40.4 shows that the amount of energy emitted by the oscillator and carried by the quantum of radiation is

$$E = hf \tag{40.5}$$

According to property 2(b), an oscillator emits or absorbs energy only when it changes quantum states. If it remains in one quantum state, no energy is absorbed or emitted. Figure 40.6 is an **energy-level diagram** showing the quantized energy levels and allowed transitions proposed by Planck. This important semigraphical representation is used often in quantum physics.[2] The vertical axis is linear in energy, and the allowed energy levels are represented as horizontal lines. The quantized system can have only the energies represented by the horizontal lines.

Max Planck
German Physicist (1858–1947)
Planck introduced the concept of "quantum of action" (Planck's constant, *h*) in an attempt to explain the spectral distribution of blackbody radiation, which laid the foundations for quantum theory. In 1918, he was awarded the Nobel Prize in Physics for this discovery of the quantized nature of energy.

[1]A quantum number is generally an integer (although half-integer quantum numbers can occur) that describes an allowed state of a system, such as the values of n describing the normal modes of oscillation of a string fixed at both ends, as discussed in Section 18.3.

[2]We first saw an energy-level diagram in Section 21.3.

The key point in Planck's theory is the radical assumption of quantized energy states. This development—a clear deviation from classical physics—marked the birth of the quantum theory.

In the Rayleigh–Jeans model, the average energy associated with a particular wavelength of standing waves in the cavity is the same for all wavelengths and is equal to k_BT. Planck used the same classical ideas as in the Rayleigh–Jeans model to arrive at the energy density as a product of constants and the average energy for a given wavelength, but the average energy is not given by the equipartition theorem. A wave's average energy is the average energy difference between levels of the oscillator, *weighted according to the probability of the wave being emitted*. This weighting is based on the occupation of higher-energy states as described by the Boltzmann distribution law, which was discussed in Section 21.5. According to this law, the probability of a state being occupied is proportional to the factor e^{-E/k_BT}, where E is the energy of the state.

At low frequencies (long wavelengths), according to property 2(a), the energy levels are close together as on the right in Figure 40.7, and many of the energy states are excited because the Boltzmann factor e^{-E/k_BT} is relatively large for these states. Therefore, there are many contributions to the outgoing radiation, although each contribution has very low energy. Now, consider high-frequency radiation, that is, radiation with short wavelength. To obtain this radiation, the allowed energies are very far apart as on the left in Figure 40.7. The probability of thermal agitation exciting these high energy levels is small because of the small value of the Boltzmann factor for large values of E. At high frequencies, the low probability of excitation results in very little contribution to the total energy, even though each quantum is of large energy. This low probability "turns the curve over" and brings it down to zero again at short wavelengths.

Using this approach, Planck generated a theoretical expression for the wavelength distribution that agreed remarkably well with the experimental curves in Figure 40.3:

$$I(\lambda, T) = \frac{2\pi hc^2}{\lambda^5\left(e^{hc/\lambda k_BT} - 1\right)} \tag{40.6}$$

◀ **Planck's wavelength distribution function**

Pitfall Prevention 40.2

***n* Is Again an Integer** In the preceding chapters on optics, we used the symbol n for the index of refraction, which was not an integer. Here we are again using n as we did in Chapter 18 to indicate the standing-wave mode on a string or in an air column. In quantum physics, n is often used as an integer quantum number to identify a particular quantum state of a system.

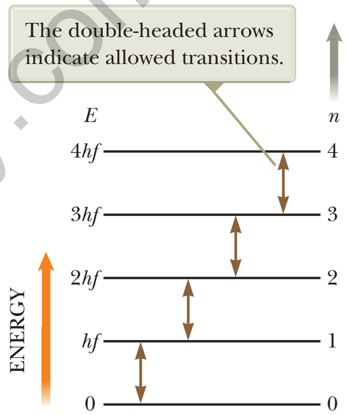


Figure 40.6 Allowed energy levels for an oscillator with frequency f.

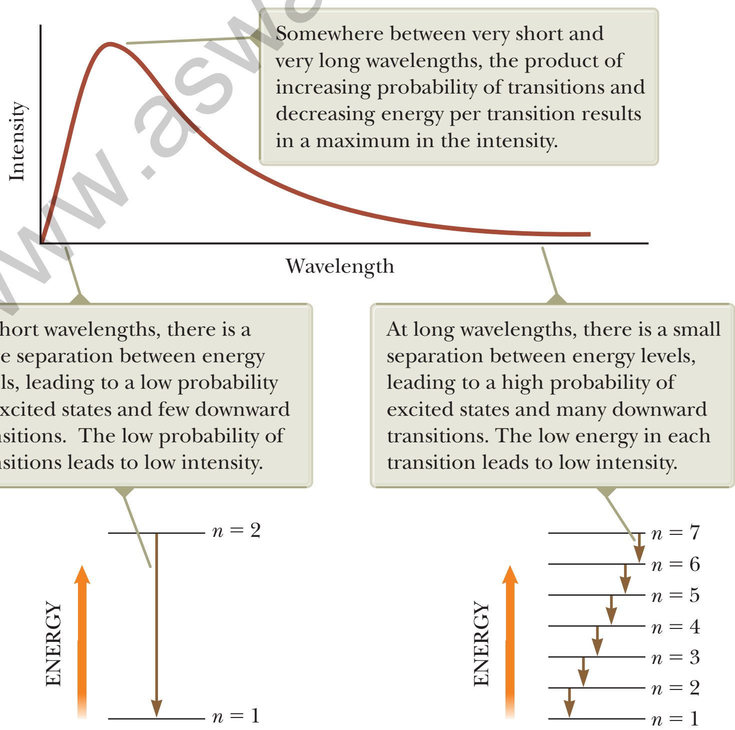


Figure 40.7 In Planck's model, the average energy associated with a given wavelength is the product of the energy of a transition and a factor related to the probability of the transition occurring.

This function includes the parameter h, which Planck adjusted so that his curve matched the experimental data at all wavelengths. The value of this parameter is found to be independent of the material of which the black body is made and independent of the temperature; it is a fundamental constant of nature. The value of h, Planck's constant, which was first introduced in Chapter 35, is

Planck's constant

$$h = 6.626 \times 10^{-34}\ \mathrm{J \cdot s} \tag{40.7}$$

At long wavelengths, Equation 40.6 reduces to the Rayleigh–Jeans expression, Equation 40.3 (see Problem 14), and at short wavelengths, it predicts an exponential decrease in $I(\lambda, T)$ with decreasing wavelength, in agreement with experimental results.

When Planck presented his theory, most scientists (including Planck!) did not consider the quantum concept to be realistic. They believed it was a mathematical trick that happened to predict the correct results. Hence, Planck and others continued to search for a more "rational" explanation of blackbody radiation. Subsequent developments, however, showed that a theory based on the quantum concept (rather than on classical concepts) had to be used to explain not only blackbody radiation but also a number of other phenomena at the atomic level.

In 1905, Einstein rederived Planck's results by assuming the oscillations of the electromagnetic field were themselves quantized. In other words, he proposed that quantization is a fundamental property of light and other electromagnetic radiation, which led to the concept of photons as shall be discussed in Section 40.2. Critical to the success of the quantum or photon theory was the relation between energy and frequency, which classical theory completely failed to predict.

You may have had your body temperature measured at the doctor's office by an *ear thermometer,* which can read your temperature very quickly (Fig. 40.8). In a fraction of a second, this type of thermometer measures the amount of infrared radiation emitted by the eardrum. It then converts the amount of radiation into a temperature reading. This thermometer is very sensitive because temperature is raised to the fourth power in Stefan's law. Suppose you have a fever 1°C above normal. Because absolute temperatures are found by adding 273 to Celsius temperatures, the ratio of your fever temperature to normal body temperature of 37°C is

Figure 40.8 An ear thermometer measures a patient's temperature by detecting the intensity of infrared radiation leaving the eardrum.

$$\frac{T_{\text{fever}}}{T_{\text{normal}}} = \frac{38°\text{C} + 273°\text{C}}{37°\text{C} + 273°\text{C}} = 1.003\,2$$

which is only a 0.32% increase in temperature. The increase in radiated power, however, is proportional to the fourth power of temperature, so

$$\frac{P_{\text{fever}}}{P_{\text{normal}}} = \left(\frac{38°\text{C} + 273°\text{C}}{37°\text{C} + 273°\text{C}}\right)^4 = 1.013$$

The result is a 1.3% increase in radiated power, which is easily measured by modern infrared radiation sensors.

Example 40.1 Thermal Radiation from Different Objects

(A) Find the peak wavelength of the blackbody radiation emitted by the human body when the skin temperature is 35°C.

SOLUTION

Conceptualize Thermal radiation is emitted from the surface of any object. The peak wavelength is related to the surface temperature through Wien's displacement law (Eq. 40.2).

Categorize We evaluate results using an equation developed in this section, so we categorize this example as a substitution problem.

40.1 continued

Solve Equation 40.2 for λ_{max}: (1) $\lambda_{max} = \frac{2.898 \times 10^{-3}\text{ m}\cdot\text{K}}{T}$

Substitute the surface temperature: $\lambda_{max} = \frac{2.898 \times 10^{-3}\text{ m}\cdot\text{K}}{308\text{ K}} = \boxed{9.41\ \mu\text{m}}$

This radiation is in the infrared region of the spectrum and is invisible to the human eye. Some animals (pit vipers, for instance) are able to detect radiation of this wavelength and therefore can locate warm-blooded prey even in the dark.

(B) Find the peak wavelength of the blackbody radiation emitted by the tungsten filament of a lightbulb, which operates at 2 000 K.

SOLUTION

Substitute the filament temperature into Equation (1): $\lambda_{max} = \frac{2.898 \times 10^{-3}\text{ m}\cdot\text{K}}{2\,000\text{ K}} = \boxed{1.45\ \mu\text{m}}$

This radiation is also in the infrared, meaning that most of the energy emitted by a lightbulb is not visible to us.

(C) Find the peak wavelength of the blackbody radiation emitted by the Sun, which has a surface temperature of approximately 5 800 K.

SOLUTION

Substitute the surface temperature into Equation (1): $\lambda_{max} = \frac{2.898 \times 10^{-3}\text{ m}\cdot\text{K}}{5\,800\text{ K}} = \boxed{0.500\ \mu\text{m}}$

This radiation is near the center of the visible spectrum, near the color of a yellow-green tennis ball. Because it is the most prevalent color in sunlight, our eyes have evolved to be most sensitive to light of approximately this wavelength.

Example 40.2 The Quantized Oscillator AM

A 2.00-kg block is attached to a massless spring that has a force constant of $k = 25.0$ N/m. The spring is stretched 0.400 m from its equilibrium position and released from rest.

(A) Find the total energy of the system and the frequency of oscillation according to classical calculations.

SOLUTION

Conceptualize We understand the details of the block's motion from our study of simple harmonic motion in Chapter 15. Review that material if you need to.

Categorize The phrase "according to classical calculations" tells us to categorize this part of the problem as a classical analysis of the oscillator. We model the block as a *particle in simple harmonic motion*.

Analyze Based on the way the block is set into motion, its amplitude is 0.400 m.

Evaluate the total energy of the block–spring system using Equation 15.21: $E = \frac{1}{2}kA^2 = \frac{1}{2}(25.0\text{ N/m})(0.400\text{ m})^2 = \boxed{2.00\text{ J}}$

Evaluate the frequency of oscillation from Equation 15.14: $f = \frac{1}{2\pi}\sqrt{\frac{k}{m}} = \frac{1}{2\pi}\sqrt{\frac{25.0\text{ N/m}}{2.00\text{ kg}}} = \boxed{0.563\text{ Hz}}$

(B) Assuming the energy of the oscillator is quantized, find the quantum number n for the system oscillating with this amplitude.

continued

40.2 continued

SOLUTION

Categorize This part of the problem is categorized as a quantum analysis of the oscillator. We model the block–spring system as a Planck oscillator.

Analyze Solve Equation 40.4 for the quantum number n:

$$n = \frac{E_n}{hf}$$

Substitute numerical values:

$$n = \frac{2.00\text{ J}}{(6.626 \times 10^{-34}\text{ J}\cdot\text{s})(0.563\text{ Hz})} = \boxed{5.36 \times 10^{33}}$$

Finalize Notice that 5.36×10^{33} is a very large quantum number, which is typical for macroscopic systems. Changes between quantum states for the oscillator are explored next.

WHAT IF? Suppose the oscillator makes a transition from the $n = 5.36 \times 10^{33}$ state to the state corresponding to $n = 5.36 \times 10^{33} - 1$. By how much does the energy of the oscillator change in this one-quantum change?

Answer From Equation 40.5 and the result to part (A), the energy carried away due to the transition between states differing in n by 1 is

$$E = hf = (6.626 \times 10^{-34}\text{ J}\cdot\text{s})(0.563\text{ Hz}) = 3.73 \times 10^{-34}\text{ J}$$

This energy change due to a one-quantum change is fractionally equal to $3.73 \times 10^{-34}\text{ J}/2.00\text{ J}$, or on the order of one part in 10^{34}! It is such a small fraction of the total energy of the oscillator that it cannot be detected. Therefore, even though the energy of a macroscopic block–spring system is quantized and does indeed decrease by small quantum jumps, our senses perceive the decrease as continuous. Quantum effects become important and detectable only on the submicroscopic level of atoms and molecules.

40.2 The Photoelectric Effect

Blackbody radiation was the first phenomenon to be explained with a quantum model. In the latter part of the 19th century, at the same time that data were taken on thermal radiation, experiments showed that light incident on certain metallic surfaces causes electrons to be emitted from those surfaces. This phenomenon, which was first discussed in Section 35.1, is known as the **photoelectric effect,** and the emitted electrons are called **photoelectrons.**[3]

Figure 40.9 is a diagram of an apparatus for studying the photoelectric effect. An evacuated glass or quartz tube contains a metallic plate E (the emitter) connected to the negative terminal of a battery and another metallic plate C (the collector) that is connected to the positive terminal of the battery. When the tube is kept in the dark, the ammeter reads zero, indicating no current in the circuit. However, when plate E is illuminated by light having an appropriate wavelength, a current is detected by the ammeter, indicating a flow of charges across the gap between plates E and C. This current arises from photoelectrons emitted from plate E and collected at plate C.

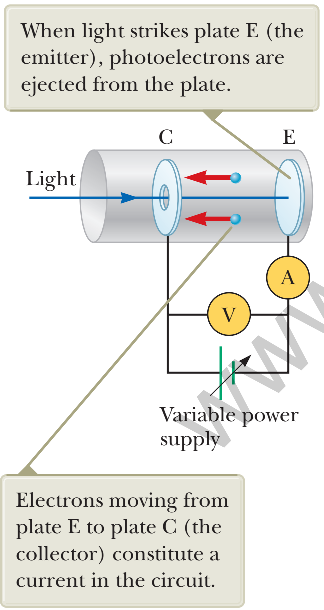


Figure 40.9 A circuit diagram for studying the photoelectric effect.

Figure 40.10 is a plot of photoelectric current versus potential difference ΔV applied between plates E and C for two light intensities. At large values of ΔV, the current reaches a maximum value; all the electrons emitted from E are collected at C, and the current cannot increase further. In addition, the maximum current increases as the intensity of the incident light increases, as you might expect,

[3]Photoelectrons are not different from other electrons. They are given this name solely because of their ejection from a metal by light in the photoelectric effect.

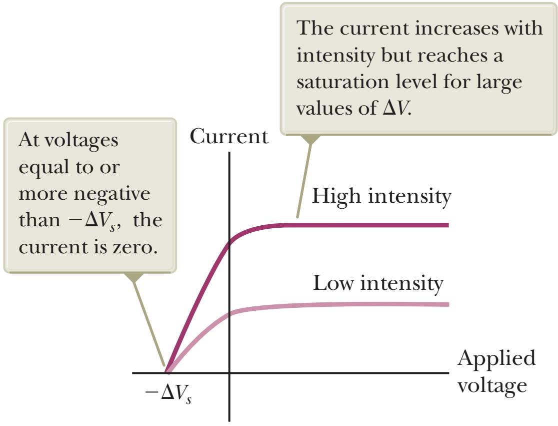


Figure 40.10 Photoelectric current versus applied potential difference for two light intensities.

because more electrons are ejected by the higher-intensity light. Finally, when ΔV is negative—that is, when the battery in the circuit is reversed to make plate E positive and plate C negative—the current drops because many of the photoelectrons emitted from E are repelled by the now negative plate C. In this situation, only those photoelectrons having a kinetic energy greater than $e|\Delta V|$ reach plate C, where e is the magnitude of the charge on the electron. When ΔV is equal to or more negative than $-\Delta V_s$, where ΔV_s is the **stopping potential,** no photoelectrons reach C and the current is zero.

Let's model the combination of the electric field between the plates and an electron ejected from plate E as an isolated system. Suppose this electron stops just as it reaches plate C. Because the system is isolated, the appropriate reduction of Equation 8.2 is

$$\Delta K + \Delta U = 0$$

where the initial configuration is at the instant the electron leaves the metal with kinetic energy K_i and the final configuration is when the electron stops just before touching plate C. If we define the electric potential energy of the system in the initial configuration to be zero, we have

$$(0 - K_i) + [(q)(\Delta V) - 0] = 0 \quad \rightarrow \quad K_i = q\Delta V = -e\Delta V$$

Now suppose the potential difference ΔV is increased in the negative direction just until the current is zero at $\Delta V = -\Delta V_s$. In this case, the electron that stops immediately before reaching plate C has the maximum possible kinetic energy upon leaving the metal surface. The previous equation can then be written as

$$K_{\max} = e\,\Delta V_s \qquad \textbf{(40.8)}$$

This equation allows us to measure $K_{\max}$ experimentally by determining the magnitude of the voltage ΔV_s at which the current drops to zero.

Several features of the photoelectric effect are listed below. For each feature, we compare the predictions made by a classical approach, using the wave model for light, with the experimental results.

1. Dependence of photoelectron kinetic energy on light intensity

Classical prediction: Electrons should absorb energy continuously from the electromagnetic waves. As the light intensity incident on a metal is increased, energy should be transferred into the metal at a higher rate and the electrons should be ejected with more kinetic energy.
Experimental result: The maximum kinetic energy of photoelectrons is *independent* of light intensity as shown in Figure 40.10 with both curves falling to zero at the *same* negative voltage. (According to Equation 40.8, the maximum kinetic energy is proportional to the stopping potential.)

2. Time interval between incidence of light and ejection of photoelectrons
Classical prediction: At low light intensities, a measurable time interval should pass between the instant the light is turned on and the time an electron is ejected from the metal. This time interval is required for the electron to absorb the incident radiation before it acquires enough energy to escape from the metal.
Experimental result: Electrons are emitted from the surface of the metal almost *instantaneously* (less than 10^{-9} s after the surface is illuminated), even at very low light intensities.

3. Dependence of ejection of electrons on light frequency
Classical prediction: Electrons should be ejected from the metal at any incident light frequency, as long as the light intensity is high enough, because energy is transferred to the metal regardless of the incident light frequency.
Experimental result: No electrons are emitted if the incident light frequency falls below some **cutoff frequency** f_c, whose value is characteristic of the material being illuminated. No electrons are ejected below this cutoff frequency *regardless* of the light intensity.

4. Dependence of photoelectron kinetic energy on light frequency
Classical prediction: There should be *no* relationship between the frequency of the light and the electron kinetic energy. The kinetic energy should be related to the intensity of the light.
Experimental result: The maximum kinetic energy of the photoelectrons increases with increasing light frequency.

For these features, experimental results contradict *all four* classical predictions. A successful explanation of the photoelectric effect was given by Einstein in 1905, the same year he published his special theory of relativity. As part of a general paper on electromagnetic radiation, for which he received a Nobel Prize in Physics in 1921, Einstein extended Planck's concept of quantization to electromagnetic waves as mentioned in Section 40.1. Einstein assumed light (or any other electromagnetic wave) of frequency f from *any* source can be considered a stream of quanta. Today we call these quanta **photons.** Each photon has an energy E given by Equation 40.5, $E = hf$, and each moves in a vacuum at the speed of light c, where $c = 3.00 \times 10^8$ m/s.

Quick Quiz 40.2 While standing outdoors one evening, you are exposed to the following four types of electromagnetic radiation: yellow light from a sodium street lamp, radio waves from an AM radio station, radio waves from an FM radio station, and microwaves from an antenna of a communications system. Rank these types of waves in terms of photon energy from highest to lowest.

Let us organize Einstein's model for the photoelectric effect using the properties of structural models:

1. *Physical components:*
We imagine the system to consist of two physical components: (1) an electron that is to be ejected by an incoming photon and (2) the remainder of the metal.
2. *Behavior of the components:*
(a) In Einstein's model, a photon of the incident light gives *all* its energy hf to a *single* electron in the metal. Therefore, the absorption of energy by the electrons is not a continuous process as envisioned in the wave model, but rather a discontinuous process in which energy is delivered to the electrons in bundles. The energy transfer is accomplished via a one-photon/one-electron event.[4]

[4]In principle, two photons could combine to provide an electron with their combined energy. That is highly improbable, however, without the high intensity of radiation available from very strong lasers.

(b) We can describe the time evolution of the system by applying the nonisolated system model for energy over a time interval that includes the absorption of one photon and the ejection of the corresponding electron. Energy is transferred into the system by electromagnetic radiation, the photon. The system has two types of energy: the potential energy of the metal–electron system and the kinetic energy of the ejected electron. Therefore, we can write the conservation of energy equation (Eq. 8.2) as

$$\Delta K + \Delta U = T_{ER} \quad \textbf{(40.9)}$$

The energy transfer into the system is that of the photon, $T_{ER} = hf$. During the process, the kinetic energy of the electron increases from zero to its final value, which we assume to be the maximum possible value K_{max}. The potential energy of the system increases because the electron is pulled away from the metal to which it is attracted. We define the potential energy of the system when the electron is outside the metal as zero. The potential energy of the system when the electron is in the metal is $U = -\phi$, where ϕ is called the **work function** of the metal. The work function represents the minimum energy with which an electron is bound in the metal and is on the order of a few electron volts. Table 40.1 lists selected values. The increase in potential energy of the system when the electron is removed from the metal is the work function ϕ. Substituting these energies into Equation 40.9, we have

$$(K_{max} - 0) + [0 - (-\phi)] = hf$$

$$K_{max} + \phi = hf \quad \textbf{(40.10)}$$

If the electron makes collisions with other electrons or metal ions as it is being ejected, some of the incoming energy is transferred to the metal and the electron is ejected with less kinetic energy than K_{max}.

Table 40.1 Work Functions of Selected Metals

Metal	ϕ (eV)
Na	2.46
Al	4.08
Fe	4.50
Cu	4.70
Zn	4.31
Ag	4.73
Pt	6.35
Pb	4.14

Note: Values are typical for metals listed. Actual values may vary depending on whether the metal is a single crystal or polycrystalline. Values may also depend on the face from which electrons are ejected from crystalline metals. Furthermore, different experimental procedures may produce differing values.

The prediction made by Einstein is an equation for the maximum kinetic energy of an ejected electron as a function of frequency of the illuminating radiation. This equation can be found by rearranging Equation 40.10:

$$K_{max} = hf - \phi \quad \textbf{(40.11)}$$

◀ **Photoelectric effect equation**

With Einstein's structural model, one can explain the observed features of the photoelectric effect that cannot be understood using classical concepts:

1. Dependence of photoelectron kinetic energy on light intensity

   Equation 40.11 shows that K_{max} is independent of the light intensity. The maximum kinetic energy of any one electron, which equals $hf - \phi$, depends only on the light frequency and the work function. If the light intensity is doubled, the number of photons arriving per unit time is doubled, which doubles the rate at which photoelectrons are emitted. The maximum kinetic energy of any one photoelectron, however, is unchanged.

2. Time interval between incidence of light and ejection of photoelectrons

   Near-instantaneous emission of electrons is consistent with the photon model of light. The incident energy appears in small packets, and there is a one-to-one interaction between photons and electrons. If the incident light has very low intensity, there are very few photons arriving per unit time interval; each photon, however, can have sufficient energy to eject an electron immediately.

3. Dependence of ejection of electrons on light frequency

   Because the photon must have energy greater than the work function ϕ to eject an electron, the photoelectric effect cannot be observed below a

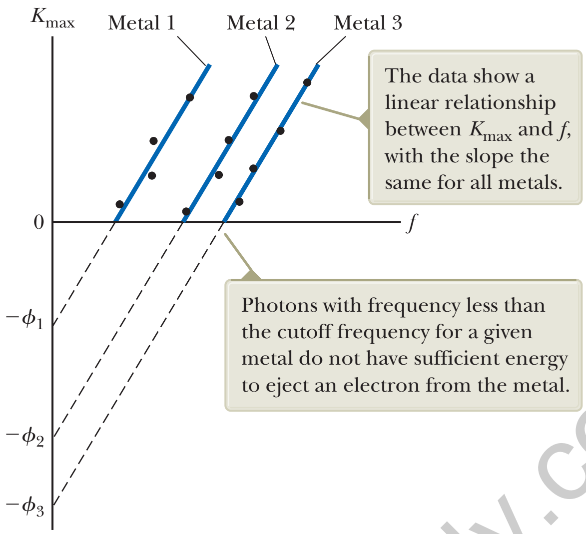


Figure 40.11 A plot of K_{max} for photoelectrons versus frequency of incident light in a typical photoelectric effect experiment.

certain cutoff frequency. If the energy of an incoming photon does not satisfy this requirement, an electron cannot be ejected from the surface, even though many photons per unit time are incident on the metal in a very intense light beam.

4. Dependence of photoelectron kinetic energy on light frequency
A photon of higher frequency carries more energy and therefore ejects a photoelectron with more kinetic energy than does a photon of lower frequency.

Einstein's model predicts a linear relationship (Eq. 40.11) between the maximum electron kinetic energy K_{max} and the light frequency f. Experimental observation of a linear relationship between K_{max} and f would be a final confirmation of Einstein's theory. Indeed, such a linear relationship was observed experimentally within a few years of Einstein's theory and is sketched in Figure 40.11. The slope of the lines in such a plot is Planck's constant h. The intercept on the horizontal axis gives the cutoff frequency below which no photoelectrons are emitted. The cutoff frequency is related to the work function through the relationship $f_c = \phi/h$. The cutoff frequency corresponds to a **cutoff wavelength** λ_c, where

Cutoff wavelength ▶

$$\lambda_c = \frac{c}{f_c} = \frac{c}{\phi/h} = \frac{hc}{\phi} \tag{40.12}$$

and c is the speed of light. Wavelengths greater than λ_c incident on a material having a work function ϕ do not result in the emission of photoelectrons.

The combination hc in Equation 40.12 often occurs when relating a photon's energy to its wavelength. A common shortcut when solving problems is to express this combination in useful units according to the following approximation:

$$hc = 1\,240 \text{ eV} \cdot \text{nm}$$

One of the first practical uses of the photoelectric effect was as the detector in a camera's light meter. Light reflected from the object to be photographed strikes a photoelectric surface in the meter, causing it to emit photoelectrons that then pass through a sensitive ammeter. The magnitude of the current in the ammeter depends on the light intensity.

The phototube, another early application of the photoelectric effect, acts much like a switch in an electric circuit. It produces a current in the circuit when light of sufficiently high frequency falls on a metal plate in the phototube, but produces no current in the dark. Phototubes were used in burglar alarms and in the detection of the soundtrack on motion picture film. Modern semiconductor devices have now replaced older devices based on the photoelectric effect.

Today, the photoelectric effect is used in the operation of photomultiplier tubes. Figure 40.12 shows the structure of such a device. A photon striking the photocathode ejects an electron by means of the photoelectric effect. This electron accelerates across the potential difference between the photocathode and the first *dynode,* shown as being at +200 V relative to the photocathode in Figure 40.12. This high-energy electron strikes the dynode and ejects several more electrons. The same process is repeated through a series of dynodes at ever higher potentials until an electrical pulse is produced as millions of electrons strike the last dynode. The tube is therefore called a *multiplier:* one photon at the input has resulted in millions of electrons at the output.

The photomultiplier tube is used in nuclear detectors to detect photons produced by the interaction of energetic charged particles or gamma rays with certain materials. It is also used in astronomy in a technique called *photoelectric photometry*. In that technique, the light collected by a telescope from a single star is allowed to fall on a photomultiplier tube for a time interval. The tube measures the total energy transferred by light during the time interval, which can then be converted to a luminosity of the star.

The photomultiplier tube is being replaced in many astronomical observations with a *charge-coupled device* (CCD), which is the same device used in a digital camera (Section 36.6). Half of the 2009 Nobel Prize in Physics was awarded to Willard S. Boyle (b. 1924) and George E. Smith (b. 1930) for their 1969 invention of the charge-coupled device. In a CCD, an array of pixels is formed on the silicon surface of an integrated circuit (Section 43.7). When the surface is exposed to light from an astronomical scene through a telescope or a terrestrial scene through a digital camera, electrons generated by the photoelectric effect are caught in "traps" beneath the surface. The number of electrons is related to the intensity of the light striking the surface. A signal processor measures the number of electrons associated with each pixel and converts this information into a digital code that a computer can use to reconstruct and display the scene.

The *electron bombardment CCD camera* allows higher sensitivity than a conventional CCD. In this device, electrons ejected from a photocathode by the photoelectric effect are accelerated through a high voltage before striking a CCD array. The higher energy of the electrons results in a very sensitive detector of low-intensity radiation.

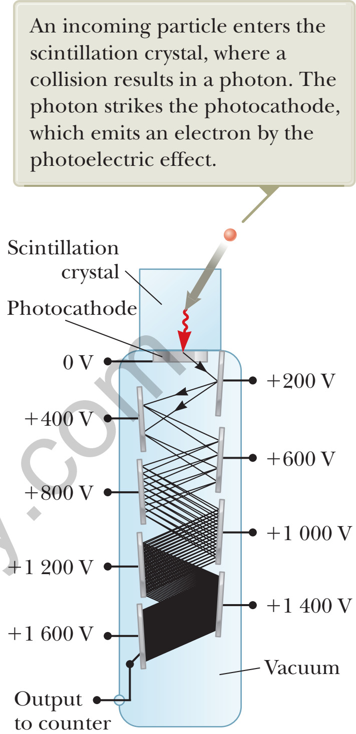


Figure 40.12 The multiplication of electrons in a photomultiplier tube.

Quick Quiz 40.3 Consider one of the curves in Figure 40.10. Suppose the intensity of the incident light is held fixed but its frequency is increased. Does the stopping potential in Figure 40.10 **(a)** remain fixed, **(b)** move to the right, or **(c)** move to the left?

Quick Quiz 40.4 Suppose classical physicists had the idea of plotting K_{max} versus f as in Figure 40.11. Draw a graph of what the expected plot would look like, based on the wave model for light.

Example 40.3 The Photoelectric Effect for Sodium

A sodium surface is illuminated with light having a wavelength of 300 nm. As indicated in Table 40.1, the work function for sodium metal is 2.46 eV.

(A) Find the maximum kinetic energy of the ejected photoelectrons.

SOLUTION

Conceptualize Imagine a photon striking the metal surface and ejecting an electron. The electron with the maximum energy is one near the surface that experiences no interactions with other particles in the metal that would reduce its energy on its way out of the metal.

continued

40.3 continued

Categorize We evaluate the results using equations developed in this section, so we categorize this example as a substitution problem.

Find the energy of each photon in the illuminating light beam from Equation 40.5:

$$E = hf = \frac{hc}{\lambda}$$

From Equation 40.11, find the maximum kinetic energy of an electron:

$$K_{\max} = \frac{hc}{\lambda} - \phi = \frac{1\,240 \text{ eV} \cdot \text{nm}}{300 \text{ nm}} - 2.46 \text{ eV} = \boxed{1.67 \text{ eV}}$$

(B) Find the cutoff wavelength λ_c for sodium.

SOLUTION

Calculate λ_c using Equation 40.12:

$$\lambda_c = \frac{hc}{\phi} = \frac{1\,240 \text{ eV} \cdot \text{nm}}{2.46 \text{ eV}} = \boxed{504 \text{ nm}}$$

40.3 The Compton Effect

Arthur Holly Compton
American Physicist (1892–1962)
Compton was born in Wooster, Ohio, and attended Wooster College and Princeton University. He became the director of the laboratory at the University of Chicago, where experimental work concerned with sustained nuclear chain reactions was conducted. This work was of central importance to the construction of the first nuclear weapon. His discovery of the Compton effect led to his sharing of the 1927 Nobel Prize in Physics with Charles Wilson.

In 1919, Einstein concluded that a photon of energy E travels in a single direction and carries a momentum equal to $E/c = hf/c$. In 1923, Arthur Holly Compton (1892–1962) and Peter Debye (1884–1966) independently carried Einstein's idea of photon momentum further.

Prior to 1922, Compton and his coworkers had accumulated evidence showing that the classical wave theory of light failed to explain the scattering of x-rays from electrons. According to classical theory, electromagnetic waves of frequency f incident on electrons should have two effects: (1) radiation pressure (see Section 34.5) should cause the electrons to accelerate in the direction of propagation of the waves, and (2) the oscillating electric field of the incident radiation should set the electrons into oscillation at the apparent frequency f', where f' is the frequency in the frame of the moving electrons. This apparent frequency is different from the frequency f of the incident radiation because of the Doppler effect (see Section 17.4). Each electron first absorbs radiation as a moving particle and then reradiates as a moving particle, thereby exhibiting two Doppler shifts in the frequency of radiation.

Because different electrons move at different speeds after the interaction, depending on the amount of energy absorbed from the electromagnetic waves, the scattered wave frequency at a given angle to the incoming radiation should show a distribution of Doppler-shifted values. Contrary to this prediction, Compton's experiments showed that at a given angle only *one* frequency of radiation is observed. Compton and his coworkers explained these experiments by treating photons not as waves but rather as point-like particles having energy hf and momentum hf/c and by assuming the energy and momentum of the isolated system of the colliding photon–electron pair are conserved. Compton adopted a particle model for something that was well known as a wave, and today this scattering phenomenon is known as the **Compton effect.** Figure 40.13 shows the quantum picture of the collision between an individual x-ray photon of frequency f_0 and an electron. In the quantum model, the electron is scattered through an angle ϕ with respect to this direction as in a billiard-ball type of collision. (The symbol ϕ used here is an angle and is not to be confused with the work function, which was discussed in the preceding section.) Compare Figure 40.13 with the two-dimensional collision shown in Figure 9.11.

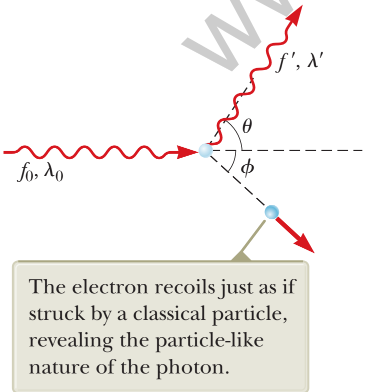


Figure 40.13 The quantum model for x-ray scattering from an electron.

Figure 40.14 is a schematic diagram of the apparatus used by Compton. The x-rays, scattered from a carbon target, were diffracted by a rotating crystal spectrometer, and the intensity was measured with an ionization chamber that generated a current proportional to the intensity. The incident beam consisted of monochromatic x-rays of wavelength $\lambda_0 = 0.071$ nm. The experimental intensity-

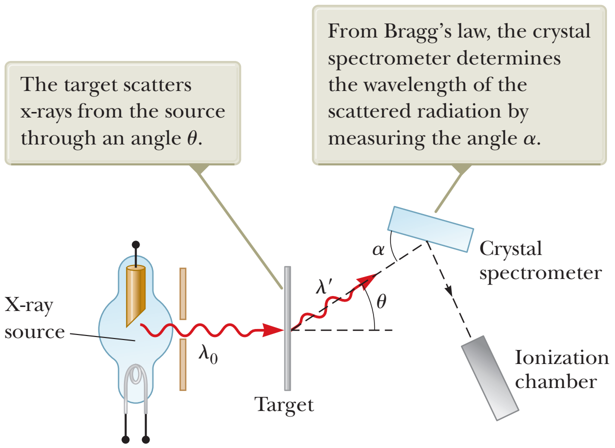


Figure 40.14 Schematic diagram of Compton's apparatus.

versus-wavelength plots observed by Compton for four scattering angles (corresponding to θ in Fig. 40.13) are shown in Figure 40.15. The graphs for the three nonzero angles show two peaks, one at λ_0 and one at $\lambda' > \lambda_0$. The shifted peak at λ' is caused by the scattering of x-rays from free electrons, which was predicted by Compton to depend on scattering angle as

$$\lambda' - \lambda_0 = \frac{h}{m_e c}(1 - \cos\theta) \qquad (40.13)$$

◀ **Compton shift equation**

where m_e is the mass of the electron. This expression is known as the **Compton shift equation** and correctly describes the positions of the peaks in Figure 40.15. The factor $h/m_e c$, called the **Compton wavelength** of the electron, has a currently accepted value of

$$\lambda_C = \frac{h}{m_e c} = 0.002\,43 \text{ nm}$$

◀ **Compton wavelength**

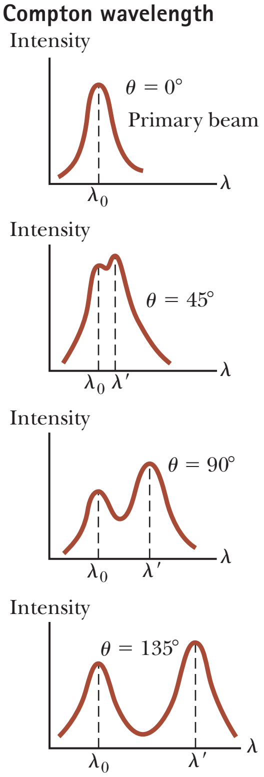


Figure 40.15 Scattered x-ray intensity versus wavelength for Compton scattering at $\theta = 0°$, 45°, 90°, and 135°.

The unshifted peak at λ_0 in Figure 40.15 is caused by x-rays scattered from electrons tightly bound to the target atoms. This unshifted peak also is predicted by Equation 40.13 if the electron mass is replaced with the mass of a carbon atom, which is approximately 23 000 times the mass of the electron. Therefore, there is a wavelength shift for scattering from an electron bound to an atom, but it is so small that it was undetectable in Compton's experiment.

Compton's measurements were in excellent agreement with the predictions of Equation 40.13. These results were the first to convince many physicists of the fundamental validity of quantum theory.

Quick Quiz 40.5 For any given scattering angle θ, Equation 40.13 gives the same value for the Compton shift for any wavelength. Keeping that in mind, for which of the following types of radiation is the fractional shift in wavelength at a given scattering angle the largest? **(a)** radio waves **(b)** microwaves **(c)** visible light **(d)** x-rays

Derivation of the Compton Shift Equation

We can derive the Compton shift equation by assuming the photon behaves like a particle and collides elastically with a free electron initially at rest as shown in Figure 40.13. The photon is treated as a particle having energy $E = hf = hc/\lambda$ and zero rest energy. We apply the isolated system analysis models for energy and momentum to the photon and the electron. In the scattering process, the total energy and total linear momentum of the system are conserved. Applying the isolated system model for energy to this process gives

$$\Delta K_{\text{photon}} + \Delta K_e = 0 \quad \rightarrow \quad \frac{hc}{\lambda_0} = \frac{hc}{\lambda'} + K_e$$

where hc/λ_0 is the energy of the incident photon, hc/λ' is the energy of the scattered photon, and K_e is the kinetic energy of the recoiling electron. Because the electron may recoil at a speed comparable to that of light, we must use the relativistic expression $K_e = (\gamma - 1)m_e c^2$ (Eq. 39.23). Therefore,

$$\frac{hc}{\lambda_0} = \frac{hc}{\lambda'} + (\gamma - 1)m_e c^2 \tag{40.14}$$

where $\gamma = 1/\sqrt{1 - (u^2/c^2)}$ and u is the speed of the electron.

Next, let's apply the isolated system model for momentum to this collision, noting that the x and y components of momentum are each conserved independently. Equation 39.28 shows that the momentum of a photon has a magnitude $p = E/c$, and we know from Equation 40.5 that $E = hf$. Therefore, $p = hf/c$. Substituting λf for c (Eq. 34.20) in this expression gives $p = h/\lambda$. Because the relativistic expression for the momentum of the recoiling electron is $p_e = \gamma m_e u$ (Eq. 39.19), we obtain the following expressions for the x and y components of linear momentum, where the angles are as described in Figure 40.13:

$$x \text{ component:} \quad \frac{h}{\lambda_0} = \frac{h}{\lambda'}\cos\theta + \gamma m_e u \cos\phi \tag{40.15}$$

$$y \text{ component:} \quad 0 = \frac{h}{\lambda'}\sin\theta - \gamma m_e u \sin\phi \tag{40.16}$$

Eliminating u and ϕ from Equations 40.14 through 40.16 gives a single expression that relates the remaining three variables (λ', λ_0, and θ). After some algebra (see Problem 64), we obtain Equation 40.13.

Example 40.4 Compton Scattering at 45°

X-rays of wavelength $\lambda_0 = 0.200\,000$ nm are scattered from a block of material. The scattered x-rays are observed at an angle of 45.0° to the incident beam. Calculate their wavelength.

SOLUTION

Conceptualize Imagine the process in Figure 40.13, with the photon scattered at 45° to its original direction.

Categorize We evaluate the result using an equation developed in this section, so we categorize this example as a substitution problem.

Solve Equation 40.13 for the wavelength of the scattered x-ray:

$$(1) \quad \lambda' = \lambda_0 + \frac{h(1 - \cos\theta)}{m_e c}$$

Substitute numerical values:

$$\lambda' = 0.200\,000 \times 10^{-9}\text{ m} + \frac{(6.626 \times 10^{-34}\text{ J}\cdot\text{s})(1 - \cos 45.0°)}{(9.11 \times 10^{-31}\text{ kg})(3.00 \times 10^{8}\text{ m/s})}$$

$$= 0.200\,000 \times 10^{-9}\text{ m} + 7.10 \times 10^{-13}\text{ m} = \boxed{0.200\,710\text{ nm}}$$

WHAT IF? What if the detector is moved so that scattered x-rays are detected at an angle larger than 45°? Does the wavelength of the scattered x-rays increase or decrease as the angle θ increases?

Answer In Equation (1), if the angle θ increases, $\cos\theta$ decreases. Consequently, the factor $(1 - \cos\theta)$ increases. Therefore, the scattered wavelength increases.

We could also apply an energy argument to achieve this same result. As the scattering angle increases, more energy is transferred from the incident photon to the electron. As a result, the energy of the scattered photon decreases with increasing scattering angle. Because $E = hf$, the frequency of the scattered photon decreases, and because $\lambda = c/f$, the wavelength increases.

40.4 The Nature of Electromagnetic Waves

In Section 35.1, we introduced the notion of competing models of light: particles and waves. Let's expand on that earlier discussion. Phenomena such as the photoelectric effect and the Compton effect offer ironclad evidence that when light (or other forms of electromagnetic radiation) and matter interact, the light behaves as if it were composed of particles having energy hf and momentum h/λ. How can light be considered a photon (in other words, a particle) when we know it is a wave? On the one hand, we describe light in terms of photons having energy and momentum. On the other hand, light and other electromagnetic waves exhibit interference and diffraction effects, which are consistent only with a wave interpretation.

Which model is correct? Is light a wave or a particle? The answer depends on the phenomenon being observed. Some experiments can be explained either better or solely with the photon model, whereas others are explained either better or solely with the wave model. We must accept both models and admit that the true nature of light is not describable in terms of any single classical picture. The same light beam that can eject photoelectrons from a metal (meaning that the beam consists of photons) can also be diffracted by a grating (meaning that the beam is a wave). In other words, the particle model and the wave model of light complement each other.

The success of the particle model of light in explaining the photoelectric effect and the Compton effect raises many other questions. If light is a particle, what is the meaning of the "frequency" and "wavelength" of the particle, and which of these two properties determines its energy and momentum? Is light *simultaneously* a wave and a particle? Although photons have no rest energy (a nonobservable quantity because a photon cannot be at rest), is there a simple expression for the *effective mass* of a moving photon? If photons have effective mass, do they experience gravitational attraction? What is the spatial extent of a photon, and how does an electron absorb or scatter one photon? Some of these questions can be answered, but others demand a view of atomic processes that is too pictorial and literal. Many of them stem from classical analogies such as colliding billiard balls and ocean waves breaking on a seashore. Quantum mechanics gives light a more flexible nature by treating the particle model and the wave model of light as both necessary and complementary. Neither model can be used exclusively to describe all properties of light. A complete understanding of the observed behavior of light can be attained only if the two models are combined in a complementary manner.

40.5 The Wave Properties of Particles

Students introduced to the dual nature of light often find the concept difficult to accept. In the world around us, we are accustomed to regarding such things as baseballs solely as particles and other things such as sound waves solely as forms of wave motion. Every large-scale observation can be interpreted by considering either a wave explanation or a particle explanation, but in the world of photons and electrons, such distinctions are not as sharply drawn.

Even more disconcerting is that, under certain conditions, the things we unambiguously call "particles" exhibit wave characteristics. In his 1923 doctoral dissertation, Louis de Broglie postulated that because photons have both wave and particle characteristics, perhaps all forms of matter have both properties. This highly revolutionary idea had no experimental confirmation at the time. According to de Broglie, electrons, just like light, have a dual particle–wave nature.

In Section 40.3, we found that the momentum of a photon can be expressed as

$$p = \frac{h}{\lambda}$$

SPL/Getty Images

Louis de Broglie
French Physicist (1892–1987)
De Broglie was born in Dieppe, France. At the Sorbonne in Paris, he studied history in preparation for what he hoped would be a career in the diplomatic service. The world of science is lucky he changed his career path to become a theoretical physicist. De Broglie was awarded the Nobel Prize in Physics in 1929 for his prediction of the wave nature of electrons.

This equation shows that the photon wavelength can be specified by its momentum: $\lambda = h/p$. De Broglie suggested that material particles of momentum p have a characteristic wavelength that is given by the *same expression*. Because the magnitude of the momentum of a particle of mass m and speed u is $p = mu$, the **de Broglie wavelength** of that particle is[5]

$$\lambda = \frac{h}{p} = \frac{h}{mu} \tag{40.17}$$

Furthermore, in analogy with photons, de Broglie postulated that particles obey the Einstein relation $E = hf$, where E is the total energy of the particle. The frequency of a particle is then

$$f = \frac{E}{h} \tag{40.18}$$

The dual nature of matter is apparent in Equations 40.17 and 40.18 because each contains both particle quantities (p and E) and wave quantities (λ and f).

The problem of understanding the dual nature of matter and radiation is conceptually difficult because the two models seem to contradict each other. This problem as it applies to light was discussed earlier. The **principle of complementarity** states that

> the wave and particle models of either matter or radiation complement each other.

Neither model can be used exclusively to describe matter or radiation adequately. Because humans tend to generate mental images based on their experiences from the everyday world, we use both descriptions in a complementary manner to explain any given set of data from the quantum world.

The Davisson–Germer Experiment

De Broglie's 1923 proposal that matter exhibits both wave and particle properties was regarded as pure speculation. If particles such as electrons had wave properties, under the correct conditions they should exhibit diffraction effects. Only three years later, C. J. Davisson (1881–1958) and L. H. Germer (1896–1971) succeeded in observing electron diffraction and measuring the wavelength of electrons. Their important discovery provided the first experimental confirmation of the waves proposed by de Broglie.

Interestingly, the intent of the initial Davisson–Germer experiment was not to confirm the de Broglie hypothesis. In fact, their discovery was made by accident (as is often the case). The experiment involved the scattering of low-energy electrons (approximately 54 eV) from a nickel target in a vacuum. During one experiment, the nickel surface was badly oxidized because of an accidental break in the vacuum system. After the target was heated in a flowing stream of hydrogen to remove the oxide coating, electrons scattered by it exhibited intensity maxima and minima at specific angles. The experimenters finally realized that the nickel had formed large crystalline regions upon heating and that the regularly spaced planes of atoms in these regions served as a diffraction grating for electrons. (See the discussion of diffraction of x-rays by crystals in Section 38.5.)

Shortly thereafter, Davisson and Germer performed more extensive diffraction measurements on electrons scattered from single-crystal targets. Their results showed conclusively the wave nature of electrons and confirmed the de Broglie relationship $p = h/\lambda$. In the same year, G. P. Thomson (1892–1975) of Scotland also observed electron diffraction patterns by passing electrons through very thin gold

Pitfall Prevention 40.3

What's Waving? If particles have wave properties, what's waving? You are familiar with waves on strings, which are very concrete. Sound waves are more abstract, but you are likely comfortable with them. Electromagnetic waves are even more abstract, but at least they can be described in terms of physical variables and electric and magnetic fields. In contrast, waves associated with particles are completely abstract and cannot be associated with a physical variable. In Chapter 41, we describe the wave associated with a particle in terms of probability.

[5]The de Broglie wavelength for a particle moving at *any* speed u is $\lambda = h/\gamma mu$, where $\gamma = [1 - (u^2/c^2)]^{-1/2}$.

foils. Diffraction patterns were subsequently observed in the scattering of helium atoms, hydrogen atoms, and neutrons. Hence, the wave nature of particles has been established in various ways.

Quick Quiz 40.6 An electron and a proton both moving at nonrelativistic speeds have the same de Broglie wavelength. Which of the following quantities are also the same for the two particles? **(a)** speed **(b)** kinetic energy **(c)** momentum **(d)** frequency

Example 40.5 Wavelengths for Microscopic and Macroscopic Objects

(A) Calculate the de Broglie wavelength for an electron ($m_e = 9.11 \times 10^{-31}$ kg) moving at 1.00×10^7 m/s.

SOLUTION

Conceptualize Imagine the electron moving through space. From a classical viewpoint, it is a particle under constant velocity. From the quantum viewpoint, the electron has a wavelength associated with it.

Categorize We evaluate the result using an equation developed in this section, so we categorize this example as a substitution problem.

Evaluate the de Broglie wavelength using Equation 40.17:

$$\lambda = \frac{h}{m_e u} = \frac{6.626 \times 10^{-34}\ \text{J}\cdot\text{s}}{(9.11 \times 10^{-31}\ \text{kg})(1.00 \times 10^7\ \text{m/s})} = \boxed{7.27 \times 10^{-11}\ \text{m}}$$

The wave nature of this electron could be detected by diffraction techniques such as those in the Davisson–Germer experiment.

(B) A rock of mass 50 g is thrown with a speed of 40 m/s. What is its de Broglie wavelength?

SOLUTION

Evaluate the de Broglie wavelength using Equation 40.17:

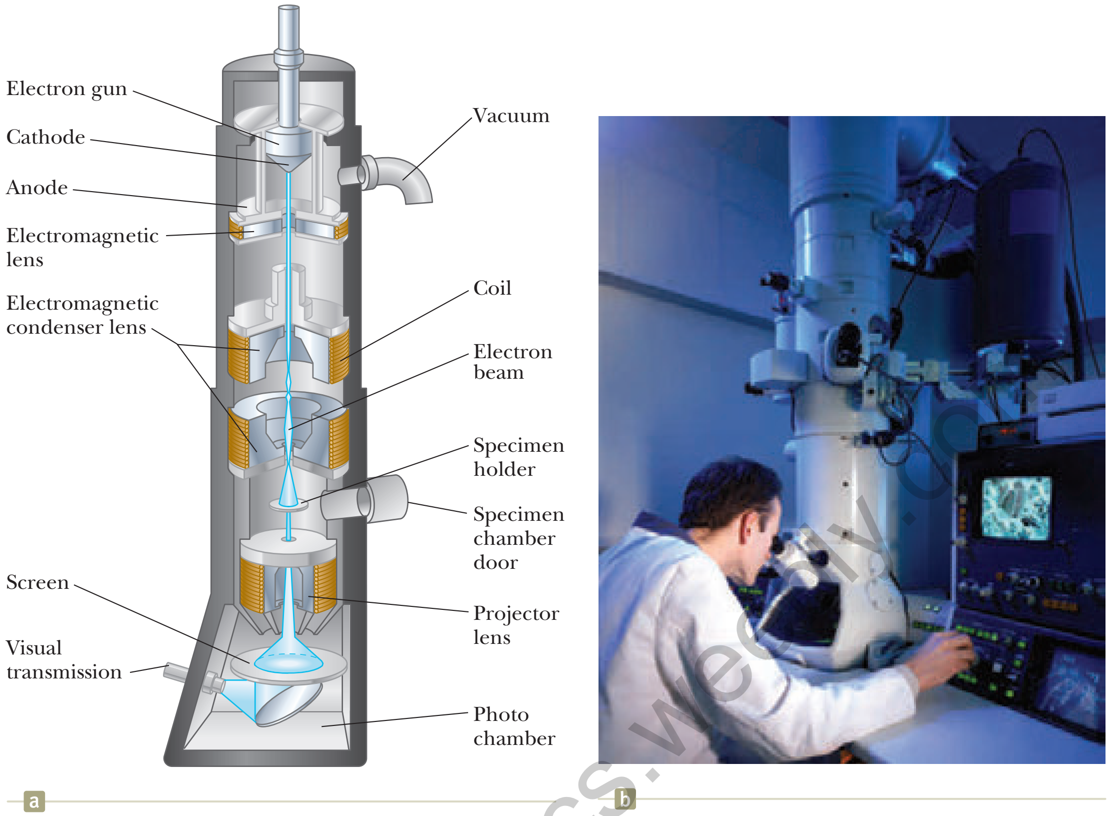
$$\lambda = \frac{h}{mu} = \frac{6.626 \times 10^{-34}\ \text{J}\cdot\text{s}}{(50 \times 10^{-3}\ \text{kg})(40\ \text{m/s})} = \boxed{3.3 \times 10^{-34}\ \text{m}}$$

This wavelength is much smaller than any aperture through which the rock could possibly pass. Hence, we could not observe diffraction effects, and as a result, the wave properties of large-scale objects cannot be observed.

The Electron Microscope

A practical device that relies on the wave characteristics of electrons is the **electron microscope.** A *transmission* electron microscope, used for viewing flat, thin samples, is shown in Figure 40.16 on page 1252. In many respects, it is similar to an optical microscope; the electron microscope, however, has a much greater resolving power because it can accelerate electrons to very high kinetic energies, giving them very short wavelengths. No microscope can resolve details that are significantly smaller than the wavelength of the waves used to illuminate the object. The shorter wavelengths of electrons gives an electron microscope a resolution that can be 1 000 times better than that from the visible light used in optical microscopes. As a result, an electron microscope with ideal lenses would be able to distinguish details approximately 1 000 times smaller than those distinguished by an optical microscope. (Electromagnetic radiation of the same wavelength as the electrons in an electron microscope is in the x-ray region of the spectrum.)

The electron beam in an electron microscope is controlled by electrostatic or magnetic deflection, which acts on the electrons to focus the beam and form an image. Rather than examining the image through an eyepiece as in an optical microscope, the viewer looks at an image formed on a monitor or other type of



Steven Allen/Brand X Pictures/Jupiter Images

Figure 40.16 (a) Diagram of a transmission electron microscope for viewing a thinly sectioned sample. The "lenses" that control the electron beam are magnetic deflection coils. (b) An electron microscope in use.

© Robert Harding Picture Library Ltd./Alamy

Figure 40.17 A scanning electron microscope photograph shows significant detail of a cheese mite, *Tyrolichus casei.* The mite is so small, with a maximum length of 0.70 mm, that ordinary microscopes do not reveal minute anatomical details.

display screen. Figure 40.17 shows the amazing detail available with an electron microscope.

40.6 A New Model: The Quantum Particle

Because in the past we considered the particle and wave models to be distinct, the discussions presented in previous sections may be quite disturbing. The notion that both light and material particles have both particle and wave properties does not fit with this distinction. Experimental evidence shows, however, that this conclusion is exactly what we must accept. The recognition of this dual nature leads to a new model, the **quantum particle,** which is a combination of the particle model introduced in Chapter 2 and the wave model discussed in Chapter 16. In this new model, entities have both particle and wave characteristics, and we must choose one appropriate behavior—particle or wave—to understand a particular phenomenon.

In this section, we shall explore this model in a way that might make you more comfortable with this idea. We shall do so by demonstrating that an entity that exhibits properties of a particle can be constructed from waves.

Let's first recall some characteristics of ideal particles and ideal waves. An ideal particle has zero size. Therefore, an essential feature of a particle is that it is *localized* in space. An ideal wave has a single frequency and is infinitely long as suggested by Figure 40.18a. Therefore, an ideal wave is *unlocalized* in space. A localized entity can be built from infinitely long waves as follows. Imagine drawing one wave along the x axis, with one of its crests located at $x = 0$, as at the top of Figure 40.18b. Now draw a second wave, of the same amplitude but a different frequency, with one of its crests also at $x = 0$. As a result of the superposition of these two waves, *beats* exist as the waves are alternately in phase and out of phase. (Beats were discussed in Section 18.7.) The bottom curve in Figure 40.18b shows the results of superposing these two waves.

Notice that we have already introduced some localization by superposing the two waves. A single wave has the same amplitude everywhere in space; no point in space is any different from any other point. By adding a second wave, however, there is something different about the in-phase points compared with the out-of-phase points.

Now imagine that more and more waves are added to our original two, each new wave having a new frequency. Each new wave is added so that one of its crests is at $x = 0$ with the result that all the waves add constructively at $x = 0$. When we add a large number of waves, the probability of a positive value of a wave function at any point $x \neq 0$ is equal to the probability of a negative value, and there is destructive interference *everywhere* except near $x = 0$, where all the crests are superposed. The result is shown in Figure 40.19. The small region of constructive interference is called a **wave packet.** This localized region of space is different from all other regions. We can identify the wave packet as a particle because it has the localized nature of a particle! The location of the wave packet corresponds to the particle's position.

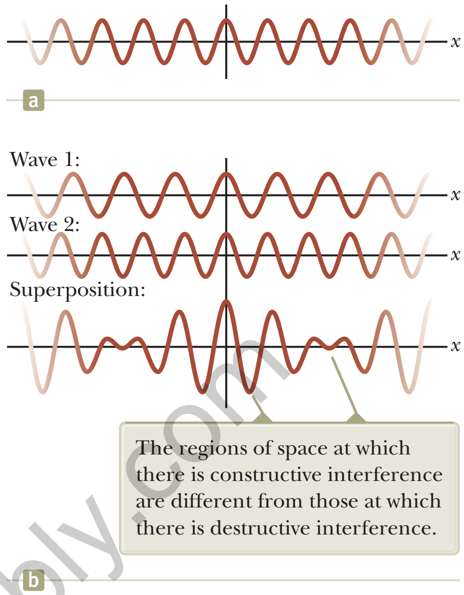


Figure 40.18 (a) An idealized wave of an exact single frequency is the same throughout space and time. (b) If two ideal waves with slightly different frequencies are combined, beats result (Section 18.7).

The localized nature of this entity is the *only* characteristic of a particle that was generated with this process. We have not addressed how the wave packet might achieve such particle characteristics as mass, electric charge, and spin. Therefore, you may not be completely convinced that we have built a particle. As further evidence that the wave packet can represent the particle, let's show that the wave packet has another characteristic of a particle.

To simplify the mathematical representation, we return to our combination of two waves. Consider two waves with equal amplitudes but different angular frequencies ω_1 and ω_2. We can represent the waves mathematically as

$$y_1 = A\cos(k_1x - \omega_1t) \quad \text{and} \quad y_2 = A\cos(k_2x - \omega_2t)$$

where, as in Chapter 16, $k = 2\pi/\lambda$ and $\omega = 2\pi f$. Using the superposition principle, let's add the waves:

$$y = y_1 + y_2 = A\cos(k_1x - \omega_1t) + A\cos(k_2x - \omega_2t)$$

It is convenient to write this expression in a form that uses the trigonometric identity

$$\cos a + \cos b = 2\cos\left(\frac{a-b}{2}\right)\cos\left(\frac{a+b}{2}\right)$$

Letting $a = k_1x - \omega_1t$ and $b = k_2x - \omega_2t$ gives

$$y = 2A\cos\left[\frac{(k_1x - \omega_1t) - (k_2x - \omega_2t)}{2}\right]\cos\left[\frac{(k_1x - \omega_1t) + (k_2x - \omega_2t)}{2}\right]$$

$$y = \left[2A\cos\left(\frac{\Delta k}{2}x - \frac{\Delta\omega}{2}t\right)\right]\cos\left(\frac{k_1 + k_2}{2}x - \frac{\omega_1 + \omega_2}{2}t\right) \qquad \textbf{(40.19)}$$

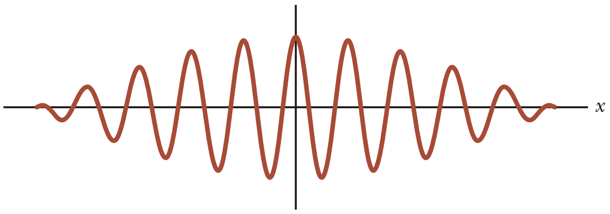


Figure 40.19 If a large number of waves are combined, the result is a wave packet, which represents a particle.

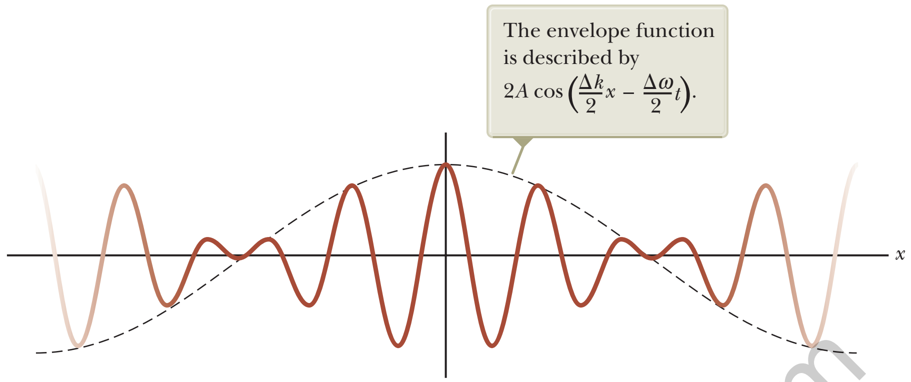


Figure 40.20 The beat pattern of Figure 40.18b, with an envelope function (dashed curve) superimposed.

where $\Delta k = k_1 - k_2$ and $\Delta\omega = \omega_1 - \omega_2$. The second cosine factor represents a wave with a wave number and frequency that are equal to the averages of the values for the individual waves.

In Equation 40.19, the factor in square brackets represents the envelope of the wave as shown by the dashed curve in Figure 40.20. This factor also has the mathematical form of a wave. This envelope of the combination can travel through space with a different speed than the individual waves. As an extreme example of this possibility, imagine combining two identical waves moving in opposite directions. The two waves move with the same speed, but the envelope has a speed of *zero* because we have built a standing wave, which we studied in Section 18.2.

For an individual wave, the speed is given by Equation 16.11,

Phase speed of a wave in a wave packet

$$v_{\text{phase}} = \frac{\omega}{k} \tag{40.20}$$

This speed is called the **phase speed** because it is the rate of advance of a crest on a single wave, which is a point of fixed phase. Equation 40.20 can be interpreted as follows: the phase speed of a wave is the ratio of the coefficient of the time variable t to the coefficient of the space variable x in the equation representing the wave, $y = A\cos(kx - \omega t)$.

The factor in brackets in Equation 40.19 is of the form of a wave, so it moves with a speed given by this same ratio:

$$v_g = \frac{\text{coefficient of time variable } t}{\text{coefficient of space variable } x} = \frac{(\Delta\omega/2)}{(\Delta k/2)} = \frac{\Delta\omega}{\Delta k}$$

The subscript g on the speed indicates that it is commonly called the **group speed,** or the speed of the wave packet (the *group* of waves) we have built. We have generated this expression for a simple addition of two waves. When a large number of waves are superposed to form a wave packet, this ratio becomes a derivative:

Group speed of a wave packet

$$v_g = \frac{d\omega}{dk} \tag{40.21}$$

Multiplying the numerator and the denominator by $\hbar$, where $\hbar = h/2\pi$, gives

$$v_g = \frac{\hbar\, d\omega}{\hbar\, dk} = \frac{d(\hbar\omega)}{d(\hbar k)} \tag{40.22}$$

Let's look at the terms in the parentheses of Equation 40.22 separately. For the numerator,

$$\hbar\omega = \frac{h}{2\pi}(2\pi f) = hf = E$$

For the denominator,

$$\hbar k = \frac{h}{2\pi}\left(\frac{2\pi}{\lambda}\right) = \frac{h}{\lambda} = p$$

Therefore, Equation 40.22 can be written as

$$v_g = \frac{d(\hbar\omega)}{d(\hbar k)} = \frac{dE}{dp} \tag{40.23}$$

Because we are exploring the possibility that the envelope of the combined waves represents the particle, consider a free particle moving with a speed u that is small compared with the speed of light. The energy of the particle is its kinetic energy:

$$E = \tfrac{1}{2}mu^2 = \frac{p^2}{2m}$$

Differentiating this equation with respect to p gives

$$v_g = \frac{dE}{dp} = \frac{d}{dp}\left(\frac{p^2}{2m}\right) = \frac{1}{2m}(2p) = u \tag{40.24}$$

Therefore, the group speed of the wave packet is identical to the speed of the particle that it is modeled to represent, giving us further confidence that the wave packet is a reasonable way to build a particle.

Quick Quiz 40.7 As an analogy to wave packets, consider an "automobile packet" that occurs near the scene of an accident on a freeway. The phase speed is analogous to the speed of individual automobiles as they move through the backup caused by the accident. The group speed can be identified as the speed of the leading edge of the packet of cars. For the automobile packet, is the group speed **(a)** the same as the phase speed, **(b)** less than the phase speed, or **(c)** greater than the phase speed?

40.7 The Double-Slit Experiment Revisited

Wave–particle duality is now a firmly accepted concept reinforced by experimental results, including those of the Davisson–Germer experiment. As with the postulates of special relativity, however, this concept often leads to clashes with familiar thought patterns we hold from everyday experience.

One way to crystallize our ideas about the electron's wave–particle duality is through an experiment in which electrons are fired at a double slit. Consider a parallel beam of mono-energetic electrons incident on a double slit as in Figure 40.21. Let's assume the slit widths are small compared with the electron wavelength so that we need not worry about diffraction maxima and minima as discussed for light in Section 38.2. An electron detector screen is positioned far from the slits at a distance much greater than d, the separation distance of the slits. If the detector screen collects electrons for a long enough time, we find a typical wave interference pattern for the counts per minute, or probability of arrival of electrons. Such an interference

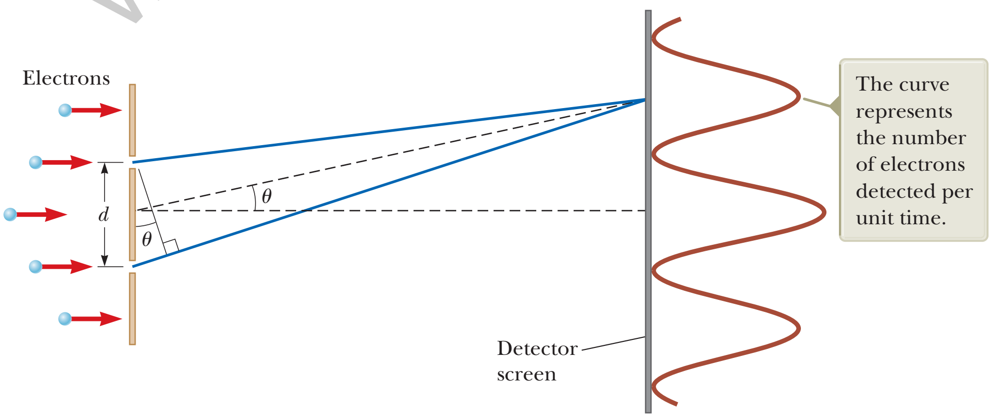


Figure 40.21 Electron interference. The slit separation d is much greater than the individual slit widths and much less than the distance between the slit and the detector screen.

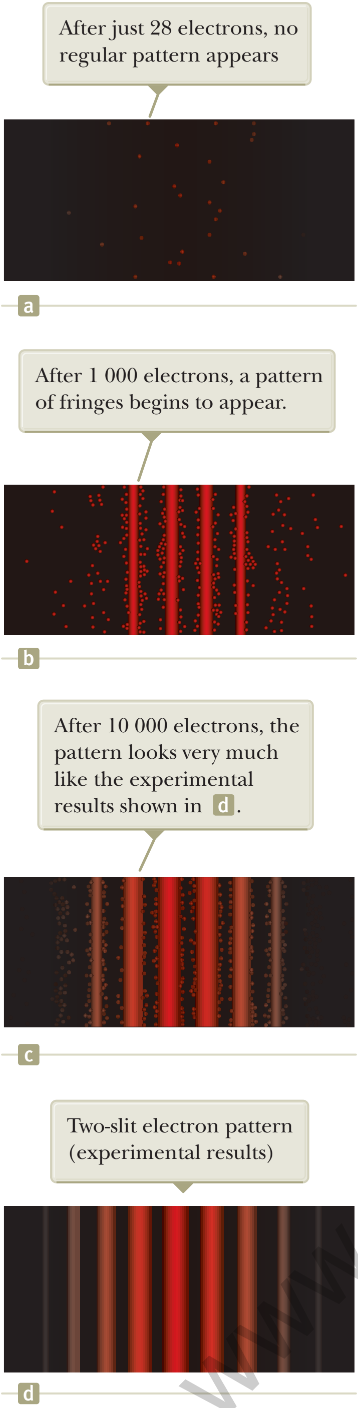


Figure 40.22 (a)–(c) Computer-simulated interference patterns for a small number of electrons incident on a double slit. (d) Computer simulation of a double-slit interference pattern produced by many electrons.

pattern would not be expected if the electrons behaved as classical particles, giving clear evidence that electrons are interfering, a distinct wave-like behavior.

If we measure the angles θ at which the maximum intensity of electrons arrives at the detector screen in Figure 40.21, we find they are described by exactly the same equation as that for light, $d \sin \theta = m\lambda$ (Eq. 37.2), where m is the order number and λ is the electron wavelength. Therefore, the dual nature of the electron is clearly shown in this experiment: the electrons are detected as particles at a localized spot on the detector screen at some instant of time, but the probability of arrival at that spot is determined by finding the intensity of two interfering waves.

Now imagine that we lower the beam intensity so that one electron at a time arrives at the double slit. It is tempting to assume the electron goes through either slit 1 or slit 2. You might argue that there are no interference effects because there is not a second electron going through the other slit to interfere with the first. This assumption places too much emphasis on the particle model of the electron, however. The interference pattern is still observed if the time interval for the measurement is sufficiently long for many electrons to pass one at a time through the slits and arrive at the detector screen! This situation is illustrated by the computer-simulated patterns in Figure 40.22 where the interference pattern becomes clearer as the number of electrons reaching the detector screen increases. Hence, our assumption that the electron is localized and goes through only one slit when both slits are open must be wrong (a painful conclusion!).

To interpret these results, we are forced to conclude that an electron interacts with both slits *simultaneously*. If you try to determine experimentally which slit the electron goes through, the act of measuring destroys the interference pattern. It is impossible to determine which slit the electron goes through. In effect, we can say only that the electron passes through *both* slits! The same arguments apply to photons.

If we restrict ourselves to a pure particle model, it is an uncomfortable notion that the electron can be present at both slits at once. From the quantum particle model, however, the particle can be considered to be built from waves that exist throughout space. Therefore, the wave components of the electron are present at both slits at the same time, and this model leads to a more comfortable interpretation of this experiment.

40.8 The Uncertainty Principle

Whenever one measures the position or velocity of a particle at any instant, experimental uncertainties are built into the measurements. According to classical mechanics, there is no fundamental barrier to an ultimate refinement of the apparatus or experimental procedures. In other words, it is possible, in principle, to make such measurements with arbitrarily small uncertainty. Quantum theory predicts, however, that it is fundamentally impossible to make simultaneous measurements of a particle's position and momentum with infinite accuracy.

In 1927, Werner Heisenberg (1901–1976) introduced this notion, which is now known as the **Heisenberg uncertainty principle:**

If a measurement of the position of a particle is made with uncertainty Δx and a simultaneous measurement of its x component of momentum is made with uncertainty Δp_x, the product of the two uncertainties can never be smaller than $\hbar/2$:

$$\Delta x \, \Delta p_x \geq \frac{\hbar}{2} \quad \textbf{(40.25)}$$

That is, it is physically impossible to measure simultaneously the exact position and exact momentum of a particle. Heisenberg was careful to point out that the inescapable uncertainties Δx and Δp_x do not arise from imperfections in practical measuring instruments. Rather, the uncertainties arise from the quantum structure of matter.

To understand the uncertainty principle, imagine that a particle has a single wavelength that is known *exactly*. According to the de Broglie relation, $\lambda = h/p$, we would therefore know the momentum to be precisely $p = h/\lambda$. In reality, a single-wavelength wave would exist throughout space. Any region along this wave is the same as any other region (Fig. 40.18a). Suppose we ask, Where is the particle this wave represents? No special location in space along the wave could be identified with the particle; all points along the wave are the same. Therefore, we have *infinite* uncertainty in the position of the particle, and we know nothing about its location. Perfect knowledge of the particle's momentum has cost us all information about its location.

In comparison, now consider a particle whose momentum is uncertain so that it has a range of possible values of momentum. According to the de Broglie relation, the result is a range of wavelengths. Therefore, the particle is not represented by a single wavelength, but rather by a combination of wavelengths within this range. This combination forms a wave packet as we discussed in Section 40.6 and illustrated in Figure 40.19. If you were asked to determine the location of the particle, you could only say that it is somewhere in the region defined by the wave packet because there is a distinct difference between this region and the rest of space. Therefore, by losing some information about the momentum of the particle, we have gained information about its position.

If you were to lose *all* information about the momentum, you would be adding together waves of all possible wavelengths, resulting in a wave packet of zero length. Therefore, if you know nothing about the momentum, you know exactly where the particle is.

The mathematical form of the uncertainty principle states that the product of the uncertainties in position and momentum is always larger than some minimum value. This value can be calculated from the types of arguments discussed above, and the result is the value of $\hbar/2$ in Equation 40.25.

Another form of the uncertainty principle can be generated by reconsidering Figure 40.19. Imagine that the horizontal axis is time rather than spatial position x. We can then make the same arguments that were made about knowledge of wavelength and position in the time domain. The corresponding variables would be frequency and time. Because frequency is related to the energy of the particle by $E = hf$, the uncertainty principle in this form is

$$\Delta E\,\Delta t \geq \frac{\hbar}{2} \qquad \textbf{(40.26)}$$

The form of the uncertainty principle given in Equation 40.26 suggests that energy conservation can appear to be violated by an amount ΔE as long as it is only for a short time interval Δt consistent with that equation. We shall use this notion to estimate the rest energies of particles in Chapter 46.

Quick Quiz 40.8 A particle's location is measured and specified as being exactly at $x = 0$, with *zero* uncertainty in the x direction. How does that location affect the uncertainty of its velocity component in the y direction? **(a)** It does not affect it. **(b)** It makes it infinite. **(c)** It makes it zero.

Werner Heisenberg
German Theoretical Physicist (1901–1976)
Heisenberg obtained his Ph.D. in 1923 at the University of Munich. While other physicists tried to develop physical models of quantum phenomena, Heisenberg developed an abstract mathematical model called *matrix mechanics*. The more widely accepted physical models were shown to be equivalent to matrix mechanics. Heisenberg made many other significant contributions to physics, including his famous uncertainty principle for which he received a Nobel Prize in Physics in 1932, the prediction of two forms of molecular hydrogen, and theoretical models of the nucleus.

Pitfall Prevention 40.4
The Uncertainty Principle Some students incorrectly interpret the uncertainty principle as meaning that a measurement interferes with the system. For example, if an electron is observed in a hypothetical experiment using an optical microscope, the photon used to see the electron collides with it and makes it move, giving it an uncertainty in momentum. This scenario does *not* represent the basis of the uncertainty principle. The uncertainty principle is independent of the measurement process and is based on the wave nature of matter.

Example 40.6 Locating an Electron

The speed of an electron is measured to be 5.00×10^3 m/s to an accuracy of 0.003 00%. Find the minimum uncertainty in determining the position of this electron.

SOLUTION

Conceptualize The fractional value given for the accuracy of the electron's speed can be interpreted as the fractional uncertainty in its momentum. This uncertainty corresponds to a minimum uncertainty in the electron's position through the uncertainty principle.

continued

40.6 continued

Categorize We evaluate the result using concepts developed in this section, so we categorize this example as a substitution problem.

Assume the electron is moving along the x axis and find the uncertainty in p_x, letting f represent the accuracy of the measurement of its speed:

$$\Delta p_x = m\,\Delta v_x = mfv_x$$

Solve Equation 40.25 for the uncertainty in the electron's position and substitute numerical values:

$$\Delta x \geq \frac{\hbar}{2\,\Delta p_x} = \frac{\hbar}{2mfv_x} = \frac{1.055 \times 10^{-34}\ \text{J}\cdot\text{s}}{2(9.11 \times 10^{-31}\ \text{kg})(0.000\,030\,0)(5.00 \times 10^{3}\ \text{m/s})}$$

$$= 3.86 \times 10^{-4}\ \text{m} = \boxed{0.386\ \text{mm}}$$

Example 40.7 The Line Width of Atomic Emissions

Atoms have quantized energy levels similar to those of Planck's oscillators, although the energy levels of an atom are usually not evenly spaced. When an atom makes a transition between states separated in energy by ΔE, energy is emitted in the form of a photon of frequency $f = \Delta E/h$. Although an excited atom can radiate at any time from $t = 0$ to $t = \infty$, the average time interval after excitation during which an atom radiates is called the **lifetime** τ. If $\tau = 1.0 \times 10^{-8}$ s, use the uncertainty principle to compute the line width Δf produced by this finite lifetime.

SOLUTION

Conceptualize The lifetime τ given for the excited state can be interpreted as the uncertainty Δt in the time at which the transition occurs. This uncertainty corresponds to a minimum uncertainty in the frequency of the radiated photon through the uncertainty principle.

Categorize We evaluate the result using concepts developed in this section, so we categorize this example as a substitution problem.

Use Equation 40.5 to relate the uncertainty in the photon's frequency to the uncertainty in its energy:

$$E = hf \;\rightarrow\; \Delta E = h\,\Delta f \;\rightarrow\; \Delta f = \frac{\Delta E}{h}$$

Use Equation 40.26 to substitute for the uncertainty in the photon's energy, giving the minimum value of Δf:

$$\Delta f \geq \frac{1}{h}\,\frac{\hbar}{2\,\Delta t} = \frac{1}{h}\,\frac{h/2\pi}{2\,\Delta t} = \frac{1}{4\pi\,\Delta t} = \frac{1}{4\pi\tau}$$

Substitute for the lifetime of the excited state:

$$\Delta f \geq \frac{1}{4\pi(1.0 \times 10^{-8}\ \text{s})} = \boxed{8.0 \times 10^{6}\ \text{Hz}}$$

WHAT IF? What if this same lifetime were associated with a transition that emits a radio wave rather than a visible light wave from an atom? Is the fractional line width $\Delta f/f$ larger or smaller than for the visible light?

Answer Because we are assuming the same lifetime for both transitions, Δf is independent of the frequency of radiation. Radio waves have lower frequencies than light waves, so the ratio $\Delta f/f$ will be larger for the radio waves. Assuming a light-wave frequency f of 6.00×10^{14} Hz, the fractional line width is

$$\frac{\Delta f}{f} = \frac{8.0 \times 10^{6}\ \text{Hz}}{6.00 \times 10^{14}\ \text{Hz}} = 1.3 \times 10^{-8}$$

This narrow fractional line width can be measured with a sensitive interferometer. Usually, however, temperature and pressure effects overshadow the natural line width and broaden the line through mechanisms associated with the Doppler effect and collisions.

Assuming a radio-wave frequency f of 94.7×10^{6} Hz, the fractional line width is

$$\frac{\Delta f}{f} = \frac{8.0 \times 10^{6}\ \text{Hz}}{94.7 \times 10^{6}\ \text{Hz}} = 8.4 \times 10^{-2}$$

Therefore, for the radio wave, this same absolute line width corresponds to a fractional line width of more than 8%.

Summary

Concepts and Principles

The characteristics of **blackbody radiation** cannot be explained using classical concepts. Planck introduced the quantum concept and Planck's constant h when he assumed atomic oscillators existing only in discrete energy states were responsible for this radiation. In Planck's model, radiation is emitted in single quantized packets whenever an oscillator makes a transition between discrete energy states. The energy of a packet is

$$E = hf \quad \textbf{(40.5)}$$

where f is the frequency of the oscillator. Einstein successfully extended Planck's quantum hypothesis to the standing waves of electromagnetic radiation in a cavity used in the blackbody radiation model.

The **photoelectric effect** is a process whereby electrons are ejected from a metal surface when light is incident on that surface. In Einstein's model, light is viewed as a stream of particles, or **photons,** each having energy $E = hf$, where h is Planck's constant and f is the frequency. The maximum kinetic energy of the ejected photoelectron is

$$K_{\max} = hf - \phi \quad \textbf{(40.11)}$$

where ϕ is the **work function** of the metal.

X-rays are scattered at various angles by electrons in a target. In such a scattering event, a shift in wavelength is observed for the scattered x-rays, a phenomenon known as the **Compton effect.** Classical physics does not predict the correct behavior in this effect. If the x-ray is treated as a photon, conservation of energy and linear momentum applied to the photon–electron collisions yields, for the Compton shift,

$$\lambda' - \lambda_0 = \frac{h}{m_e c}(1 - \cos\theta) \quad \textbf{(40.13)}$$

where m_e is the mass of the electron, c is the speed of light, and θ is the scattering angle.

Light has a dual nature in that it has both wave and particle characteristics. Some experiments can be explained either better or solely by the particle model, whereas others can be explained either better or solely by the wave model.

Every object of mass m and momentum $p = mu$ has wave properties, with a **de Broglie wavelength** given by

$$\lambda = \frac{h}{p} = \frac{h}{mu} \quad \textbf{(40.17)}$$

By combining a large number of waves, a single region of constructive interference, called a **wave packet,** can be created. The wave packet carries the characteristic of localization like a particle does, but it has wave properties because it is built from waves. For an individual wave in the wave packet, the **phase speed** is

$$v_{\text{phase}} = \frac{\omega}{k} \quad \textbf{(40.20)}$$

For the wave packet as a whole, the **group speed** is

$$v_g = \frac{d\omega}{dk} \quad \textbf{(40.21)}$$

For a wave packet representing a particle, the group speed can be shown to be the same as the speed of the particle.

The **Heisenberg uncertainty principle** states that if a measurement of the position of a particle is made with uncertainty Δx and a simultaneous measurement of its linear momentum is made with uncertainty Δp_x, the product of the two uncertainties is restricted to

$$\Delta x\,\Delta p_x \geq \frac{\hbar}{2} \quad \textbf{(40.25)}$$

Another form of the uncertainty principle relates measurements of energy and time:

$$\Delta E\,\Delta t \geq \frac{\hbar}{2} \quad \textbf{(40.26)}$$

Objective Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. Rank the wavelengths of the following quantum particles from the largest to the smallest. If any have equal wavelengths, display the equality in your ranking. (a) a photon with energy 3 eV (b) an electron with kinetic energy 3 eV (c) a proton with kinetic energy 3 eV (d) a photon with energy 0.3 eV (e) an electron with momentum 3 eV/c

2. An x-ray photon is scattered by an originally stationary electron. Relative to the frequency of the incident photon, is the frequency of the scattered photon (a) lower, (b) higher, or (c) unchanged?

3. In a Compton scattering experiment, a photon of energy E is scattered from an electron at rest. After the scattering event occurs, which of the following statements is true? (a) The frequency of the photon is greater than E/h. (b) The energy of the photon is less than E. (c) The wavelength of the photon is less than hc/E. (d) The momentum of the photon increases. (e) None of those statements is true.

4. In a certain experiment, a filament in an evacuated lightbulb carries a current I_1 and you measure the spectrum of light emitted by the filament, which behaves as a black body at temperature T_1. The wavelength emitted with highest intensity (symbolized by λ_{max}) has the value λ_1. You then increase the potential difference across the filament by a factor of 8, and the current increases by a factor of 2. **(i)** After this change, what is the new value of the temperature of the filament? (a) $16T_1$ (b) $8T_1$ (c) $4T_1$ (d) $2T_1$ (e) still T_1 **(ii)** What is the new value of the wavelength emitted with highest intensity? (a) $4\lambda_1$ (b) $2\lambda_1$ (c) λ_1 (d) $\frac{1}{2}\lambda_1$ (e) $\frac{1}{4}\lambda_1$

5. Which of the following statements are true according to the uncertainty principle? More than one statement may be correct. (a) It is impossible to simultaneously determine both the position and the momentum of a particle along the same axis with arbitrary accuracy. (b) It is impossible to simultaneously determine both the energy and momentum of a particle with arbitrary accuracy. (c) It is impossible to determine a particle's energy with arbitrary accuracy in a finite amount of time. (d) It is impossible to measure the position of a particle with arbitrary accuracy in a finite amount of time. (e) It is impossible to simultaneously measure both the energy and position of a particle with arbitrary accuracy.

6. A monochromatic light beam is incident on a barium target that has a work function of 2.50 eV. If a potential difference of 1.00 V is required to turn back all the ejected electrons, what is the wavelength of the light beam? (a) 355 nm (b) 497 nm (c) 744 nm (d) 1.42 pm (e) none of those answers

7. Which of the following is most likely to cause sunburn by delivering more energy to individual molecules in skin cells? (a) infrared light (b) visible light (c) ultraviolet light (d) microwaves (e) Choices (a) through (d) are equally likely.

8. Which of the following phenomena most clearly demonstrates the wave nature of electrons? (a) the photoelectric effect (b) blackbody radiation (c) the Compton effect (d) diffraction of electrons by crystals (e) none of those answers

9. What is the de Broglie wavelength of an electron accelerated from rest through a potential difference of 50.0 V? (a) 0.100 nm (b) 0.139 nm (c) 0.174 nm (d) 0.834 nm (e) none of those answers

10. A proton, an electron, and a helium nucleus all move at speed v. Rank their de Broglie wavelengths from largest to smallest.

11. Consider (a) an electron, (b) a photon, and (c) a proton, all moving in vacuum. Choose all correct answers for each question. **(i)** Which of the three possess rest energy? **(ii)** Which have charge? **(iii)** Which carry energy? **(iv)** Which carry momentum? **(v)** Which move at the speed of light? **(vi)** Which have a wavelength characterizing their motion?

12. An electron and a proton, moving in opposite directions, are accelerated from rest through the same potential difference. Which particle has the longer wavelength? (a) The electron does. (b) The proton does. (c) Both are the same. (d) Neither has a wavelength.

13. Which of the following phenomena most clearly demonstrates the particle nature of light? (a) diffraction (b) the photoelectric effect (c) polarization (d) interference (e) refraction

14. Both an electron and a proton are accelerated to the same speed, and the experimental uncertainty in the speed is the same for the two particles. The positions of the two particles are also measured. Is the minimum possible uncertainty in the electron's position (a) less than the minimum possible uncertainty in the proton's position, (b) the same as that for the proton, (c) more than that for the proton, or (d) impossible to tell from the given information?

Conceptual Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. The opening photograph for this chapter shows a filament of a lightbulb in operation. Look carefully at the last turns of wire at the upper and lower ends of the filament. Why are these turns dimmer than the others?

2. How does the Compton effect differ from the photoelectric effect?

3. If matter has a wave nature, why is this wave-like characteristic not observable in our daily experiences?

4. If the photoelectric effect is observed for one metal, can you conclude that the effect will also be observed for another metal under the same conditions? Explain.

5. In the photoelectric effect, explain why the stopping potential depends on the frequency of light but not on the intensity.

6. Why does the existence of a cutoff frequency in the photoelectric effect favor a particle theory for light over a wave theory?

7. Which has more energy, a photon of ultraviolet radiation or a photon of yellow light? Explain.

8. All objects radiate energy. Why, then, are we not able to see all objects in a dark room?

9. Is an electron a wave or a particle? Support your answer by citing some experimental results.

10. Suppose a photograph were made of a person's face using only a few photons. Would the result be simply a very faint image of the face? Explain your answer.

11. Why is an electron microscope more suitable than an optical microscope for "seeing" objects less than 1 μm in size?

12. Is light a wave or a particle? Support your answer by citing specific experimental evidence.

13. (a) What does the slope of the lines in Figure 40.11 represent? (b) What does the y intercept represent? (c) How would such graphs for different metals compare with one another?

14. Why was the demonstration of electron diffraction by Davisson and Germer an important experiment?

15. *Iridescence* is the phenomenon that gives shining colors to the feathers of peacocks, hummingbirds (see page 1134), resplendent quetzals, and even ducks and grackles. Without pigments, it colors Morpho butterflies (Fig. CQ40.15), Urania moths, some beetles and flies, rainbow trout, and mother-of-pearl in abalone shells. Iridescent colors change as you turn an object. They are produced by a wide variety of intricate structures in different species. Problem 64 in Chapter 38 describes the structures that produce iridescence in a peacock feather. These structures were all unknown until the invention of the electron microscope. Explain why light microscopes cannot reveal them.

Figure CQ40.15

16. In describing the passage of electrons through a slit and arriving at a screen, physicist Richard Feynman said that "electrons arrive in lumps, like particles, but the probability of arrival of these lumps is determined as the intensity of the waves would be. It is in this sense that the electron behaves sometimes like a particle and sometimes like a wave." Elaborate on this point in your own words. For further discussion, see R. Feynman, *The Character of Physical Law* (Cambridge, MA: MIT Press, 1980), chap. 6.

17. The classical model of blackbody radiation given by the Rayleigh–Jeans law has two major flaws. (a) Identify the flaws and (b) explain how Planck's law deals with them.

Problems

ENHANCED WebAssign The problems found in this chapter may be assigned online in Enhanced WebAssign

1. straightforward; **2.** intermediate; **3.** challenging

1. full solution available in the *Student Solutions Manual/Study Guide*

AMT Analysis Model tutorial available in Enhanced WebAssign

GP Guided Problem

M Master It tutorial available in Enhanced WebAssign

W Watch It video solution available in Enhanced WebAssign

Section 40.1 Blackbody Radiation and Planck's Hypothesis

1. The temperature of an electric heating element is 150°C. At what wavelength does the radiation emitted from the heating element reach its peak?

2. Model the tungsten filament of a lightbulb as a black body at temperature 2 900 K. (a) Determine the wavelength of light it emits most strongly. (b) Explain why the answer to part (a) suggests that more energy from the lightbulb goes into infrared radiation than into visible light.

3. W Lightning produces a maximum air temperature on the order of 10^4 K, whereas a nuclear explosion produces a temperature on the order of 10^7 K. (a) Use Wien's displacement law to find the order of magnitude of the wavelength of the thermally produced photons radiated with greatest intensity by each of these sources. (b) Name the part of the electromagnetic spectrum where you would expect each to radiate most strongly.

4. Figure P40.4 shows the spectrum of light emitted by a firefly. (a) Determine the temperature of a black body that would emit radiation peaked at the same wavelength. (b) Based on your result, explain whether firefly radiation is blackbody radiation.

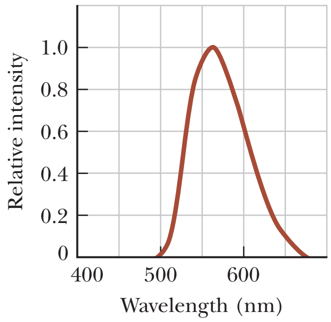


Figure P40.4

5. W The average threshold of dark-adapted (scotopic) vision is 4.00×10^{-11} W/m^2 at a central wavelength of 500 nm. If light with this intensity and wavelength enters the eye and the pupil is open to its maximum diameter of 8.50 mm, how many photons per second enter the eye?

6. M **(i)** Calculate the energy, in electron volts, of a photon whose frequency is (a) 620 THz, (b) 3.10 GHz, and (c) 46.0 MHz. **(ii)** Determine the corresponding wavelengths for the photons listed in part (i) and **(iii)** state the classification of each on the electromagnetic spectrum.

7. (a) What is the surface temperature of Betelgeuse, a red giant star in the constellation Orion (Fig. 40.4), which radiates with a peak wavelength of about 970 nm? (b) Rigel, a bluish-white star in Orion, radiates with a peak wavelength of 145 nm. Find the temperature of Rigel's surface.

8. M An FM radio transmitter has a power output of 150 kW and operates at a frequency of 99.7 MHz. How many photons per second does the transmitter emit?

9. W The human eye is most sensitive to 560-nm (green) light. What is the temperature of a black body that would radiate most intensely at this wavelength?

10. W The radius of our Sun is 6.96×10^8 m, and its total power output is 3.85×10^{26} W. (a) Assuming the Sun's surface emits as a black body, calculate its surface temperature. (b) Using the result of part (a), find λ_{max} for the Sun.

11. W A black body at 7 500 K consists of an opening of diameter 0.050 0 mm, looking into an oven. Find the number of photons per second escaping the opening and having wavelengths between 500 nm and 501 nm.

12. Consider a black body of surface area 20.0 cm^2 and temperature 5 000 K. (a) How much power does it radiate? (b) At what wavelength does it radiate most intensely? Find the spectral power per wavelength interval at (c) this wavelength and at wavelengths of (d) 1.00 nm (an x- or gamma ray), (e) 5.00 nm (ultraviolet light or an x-ray), (f) 400 nm (at the boundary between UV and visible light), (g) 700 nm (at the boundary between visible and infrared light), (h) 1.00 mm (infrared light or a microwave), and (i) 10.0 cm (a microwave or radio wave). (j) Approximately how much power does the object radiate as visible light?

13. **Review.** This problem is about how strongly matter is coupled to radiation, the subject with which quantum mechanics began. For a simple model, consider a solid iron sphere 2.00 cm in radius. Assume its temperature is always uniform throughout its volume. (a) Find the mass of the sphere. (b) Assume the sphere is at 20.0°C and has emissivity 0.860. Find the power with which it radiates electromagnetic waves. (c) If it were alone in the Universe, at what rate would the sphere's temperature be changing? (d) Assume Wien's law describes the sphere. Find the wavelength λ_{max} of electromagnetic radiation it emits most strongly. Although it emits a spectrum of waves having all different wavelengths, assume its power output is carried by photons of wavelength λ_{max}. Find (e) the energy of one photon and (f) the number of photons it emits each second.

14. Show that at long wavelengths, Planck's radiation law (Eq. 40.6) reduces to the Rayleigh–Jeans law (Eq. 40.3).

15. A simple pendulum has a length of 1.00 m and a mass of 1.00 kg. The maximum horizontal displacement of the pendulum bob from equilibrium is 3.00 cm. Calculate the quantum number n for the pendulum.

16. A pulsed ruby laser emits light at 694.3 nm. For a 14.0-ps pulse containing 3.00 J of energy, find (a) the physical length of the pulse as it travels through space and (b) the number of photons in it. (c) Assuming that the beam has a circular cross-section of 0.600 cm diameter, find the number of photons per cubic millimeter.

Section 40.2 The Photoelectric Effect

17. Molybdenum has a work function of 4.20 eV. (a) Find the cutoff wavelength and cutoff frequency for the photoelectric effect. (b) What is the stopping potential if the incident light has a wavelength of 180 nm?

18. The work function for zinc is 4.31 eV. (a) Find the cutoff wavelength for zinc. (b) What is the lowest frequency of light incident on zinc that releases photoelectrons from its surface? (c) If photons of energy 5.50 eV are incident on zinc, what is the maximum kinetic energy of the ejected photoelectrons?

19. Two light sources are used in a photoelectric experiment to determine the work function for a particular metal surface. When green light from a mercury lamp ($\lambda = 546.1$ nm) is used, a stopping potential of 0.376 V reduces the photocurrent to zero. (a) Based on this measurement, what is the work function for this metal? (b) What stopping potential would be observed when using the yellow light from a helium discharge tube ($\lambda = 587.5$ nm)?

20. W Lithium, beryllium, and mercury have work functions of 2.30 eV, 3.90 eV, and 4.50 eV, respectively. Light with a wavelength of 400 nm is incident on each of these metals. (a) Determine which of these metals exhibit the photoelectric effect for this incident light. Explain your reasoning. (b) Find the maximum kinetic energy for the photoelectrons in each case.

21. Electrons are ejected from a metallic surface with speeds of up to 4.60×10^5 m/s when light with a wavelength of 625 nm is used. (a) What is the work function of the surface? (b) What is the cutoff frequency for this surface?

22. From the scattering of sunlight, J. J. Thomson calculated the classical radius of the electron as having the value 2.82×10^{-15} m. Sunlight with an intensity of 500 W/m^2 falls on a disk with this radius. Assume light is a classical wave and the light striking the disk is completely absorbed. (a) Calculate the time interval required to accumulate 1.00 eV of energy. (b) Explain how your result for part (a) compares with the observation that photoelectrons are emitted promptly (within 10^{-9} s).

23. **Review.** An isolated copper sphere of radius 5.00 cm, initially uncharged, is illuminated by ultraviolet light of wavelength 200 nm. The work function for copper is 4.70 eV. What charge does the photoelectric effect induce on the sphere?

24. The work function for platinum is 6.35 eV. Ultraviolet light of wavelength 150 nm is incident on the clean surface of a platinum sample. We wish to predict the stopping voltage we will need for electrons ejected from the surface. (a) What is the photon energy of the ultraviolet light? (b) How do you know that these photons will eject electrons from platinum? (c) What is the maximum kinetic energy of the ejected photoelectrons? (d) What stopping voltage would be required to arrest the current of photoelectrons?

Section 40.3 The Compton Effect

25. X-rays are scattered from a target at an angle of 55.0° with the direction of the incident beam. Find the wavelength shift of the scattered x-rays.

26. A photon having wavelength λ scatters off a free electron at A (Fig. P40.26), producing a second photon having wavelength λ'. This photon then scatters off another free electron at B, producing a third photon having wavelength λ'' and moving in a direction directly opposite the original photon as shown in the figure. Determine the value of $\Delta\lambda = \lambda'' - \lambda$.

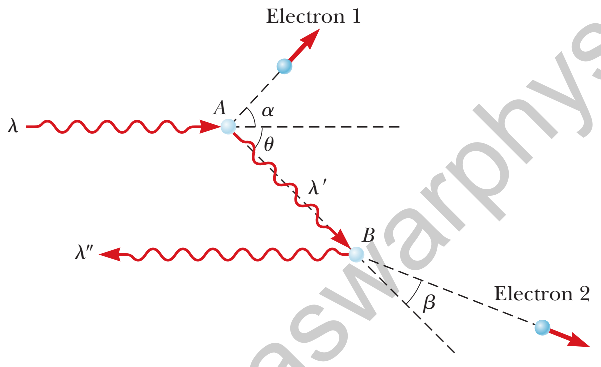


Figure P40.26

27. A 0.110-nm photon collides with a stationary electron. After the collision, the electron moves forward and the photon recoils backward. Find the momentum and the kinetic energy of the electron.

28. X-rays with a wavelength of 120.0 pm undergo Compton scattering. (a) Find the wavelengths of the photons scattered at angles of 30.0°, 60.0°, 90.0°, 120°, 150°, and 180°. (b) Find the energy of the scattered electron in each case. (c) Which of the scattering angles provides the electron with the greatest energy? Explain whether you could answer this question without doing any calculations.

29. A 0.001 60-nm photon scatters from a free electron. For what (photon) scattering angle does the recoiling electron have kinetic energy equal to the energy of the scattered photon?

30. After a 0.800-nm x-ray photon scatters from a free electron, the electron recoils at 1.40×10^6 m/s. (a) What is the Compton shift in the photon's wavelength? (b) Through what angle is the photon scattered?

31. A photon having energy $E_0 = 0.880$ MeV is scattered by a free electron initially at rest such that the scattering angle of the scattered electron is equal to that of the scattered photon as shown in Figure P40.31. (a) Determine the scattering angle of the photon and the electron. (b) Determine the energy and momentum of the scattered photon. (c) Determine the kinetic energy and momentum of the scattered electron.

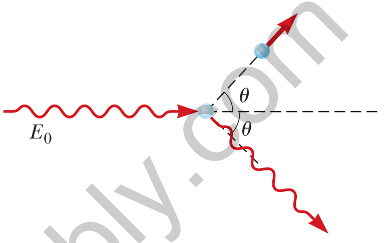


Figure P40.31 Problems 31 and 32.

32. A photon having energy E_0 is scattered by a free electron initially at rest such that the scattering angle of the scattered electron is equal to that of the scattered photon as shown in Figure P40.31. (a) Determine the angle θ. (b) Determine the energy and momentum of the scattered photon. (c) Determine the kinetic energy and momentum of the scattered electron.

33. X-rays having an energy of 300 keV undergo Compton scattering from a target. The scattered rays are detected at 37.0° relative to the incident rays. Find (a) the Compton shift at this angle, (b) the energy of the scattered x-ray, and (c) the energy of the recoiling electron.

34. In a Compton scattering experiment, a photon is scattered through an angle of 90.0° and the electron is set into motion in a direction at an angle of 20.0° to the original direction of the photon. (a) Explain how this information is sufficient to determine uniquely the wavelength of the scattered photon and (b) find this wavelength.

35. In a Compton scattering experiment, an x-ray photon scatters through an angle of 17.4° from a free electron that is initially at rest. The electron recoils with a speed of 2 180 km/s. Calculate (a) the wavelength of the incident photon and (b) the angle through which the electron scatters.

36. Find the maximum fractional energy loss for a 0.511-MeV gamma ray that is Compton scattered from (a) a free electron and (b) a free proton.

Section 40.4 The Nature of Electromagnetic Waves

37. An electromagnetic wave is called *ionizing radiation* if its photon energy is larger than, say, 10.0 eV so that a single photon has enough energy to break apart an atom. With reference to Figure P40.37 (page 1264), explain what region or regions of the electromagnetic spectrum fit this definition of ionizing radiation and what do not. (If you wish to consult a larger version of Fig. P40.37, see Fig. 34.13.)

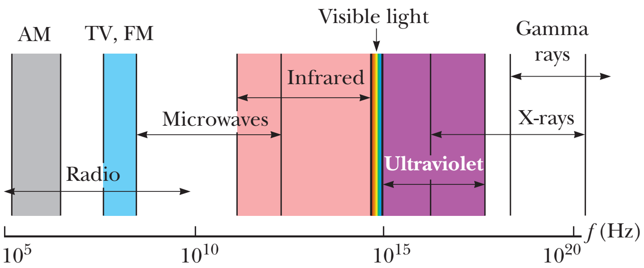


Figure P40.37

38. W **Review.** A helium–neon laser produces a beam of diameter 1.75 mm, delivering 2.00×10^{18} photons/s. Each photon has a wavelength of 633 nm. Calculate the amplitudes of (a) the electric fields and (b) the magnetic fields inside the beam. (c) If the beam shines perpendicularly onto a perfectly reflecting surface, what force does it exert on the surface? (d) If the beam is absorbed by a block of ice at 0°C for 1.50 h, what mass of ice is melted?

Section 40.5 The Wave Properties of Particles

39. (a) Calculate the momentum of a photon whose wavelength is 4.00×10^{-7} m. (b) Find the speed of an electron having the same momentum as the photon in part (a).

40. (a) An electron has a kinetic energy of 3.00 eV. Find its wavelength. (b) **What If?** A photon has energy 3.00 eV. Find its wavelength.

41. The resolving power of a microscope depends on the wavelength used. If you wanted to "see" an atom, a wavelength of approximately 1.00×10^{-11} m would be required. (a) If electrons are used (in an electron microscope), what minimum kinetic energy is required for the electrons? (b) **What If?** If photons are used, what minimum photon energy is needed to obtain the required resolution?

42. Calculate the de Broglie wavelength for a proton moving with a speed of 1.00×10^6 m/s.

43. In the Davisson–Germer experiment, 54.0-eV electrons were diffracted from a nickel lattice. If the first maximum in the diffraction pattern was observed at $\phi = 50.0°$ (Fig. P40.43), what was the lattice spacing a between the vertical columns of atoms in the figure?

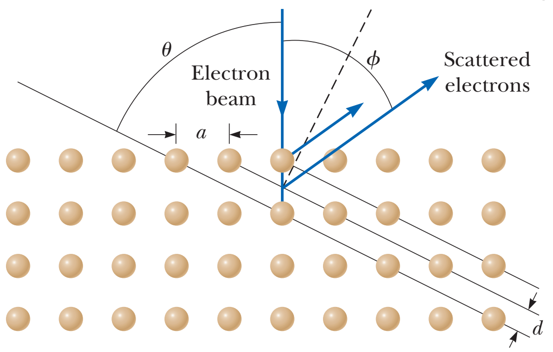


Figure P40.43

44. The nucleus of an atom is on the order of 10^{-14} m in diameter. For an electron to be confined to a nucleus, its de Broglie wavelength would have to be on this order of magnitude or smaller. (a) What would be the kinetic energy of an electron confined to this region? (b) Make an order-of-magnitude estimate of the electric potential energy of a system of an electron inside an atomic nucleus. (c) Would you expect to find an electron in a nucleus? Explain.

45. Robert Hofstadter won the 1961 Nobel Prize in Physics for his pioneering work in studying the scattering of 20-GeV electrons from nuclei. (a) What is the γ factor for an electron with total energy 20.0 GeV, defined by $\gamma = 1/\sqrt{1 - u^2/c^2}$? (b) Find the momentum of the electron. (c) Find the wavelength of the electron. (d) State how the wavelength compares with the diameter of an atomic nucleus, typically on the order of 10^{-14} m.

46. *Why is the following situation impossible?* After learning about de Broglie's hypothesis that material particles of momentum p move as waves with wavelength $\lambda = h/p$, an 80-kg student has grown concerned about being diffracted when passing through a doorway of width w = 75 cm. Assume significant diffraction occurs when the width of the diffraction aperture is less than ten times the wavelength of the wave being diffracted. Together with his classmates, the student performs precision experiments and finds that he does indeed experience measurable diffraction.

47. A photon has an energy equal to the kinetic energy of an electron with speed u, which may be close to the speed of light c. (a) Calculate the ratio of the wavelength of the photon to the wavelength of the electron. (b) Evaluate the ratio for the particle speed $u = 0.900c$. (c) **What If?** What would happen to the answer to part (b) if the material particle were a proton instead of an electron? (d) Evaluate the ratio for the particle speed $u = 0.001\,00c$. (e) What value does the ratio of the wavelengths approach at high particle speeds? (f) At low particle speeds?

48. (a) Show that the frequency f and wavelength λ of a freely moving quantum particle with mass are related by the expression

$$\left(\frac{f}{c}\right)^2 = \frac{1}{\lambda^2} + \frac{1}{\lambda_C^{\,2}}$$

where $\lambda_C = h/mc$ is the Compton wavelength of the particle. (b) Is it ever possible for a particle having nonzero mass to have the same wavelength *and* frequency as a photon? Explain.

Section 40.6 A New Model: The Quantum Particle

49. Consider a freely moving quantum particle with mass m and speed u. Its energy is $E = K = \frac{1}{2}mu^2$. (a) Determine the phase speed of the quantum wave representing the particle and (b) show that it is different from the speed at which the particle transports mass and energy.

50. For a free relativistic quantum particle moving with speed u, the total energy of the particle is $E = hf = \hbar\omega = \sqrt{p^2c^2 + m^2c^4}$ and the momentum is $p = h/\lambda = \hbar k =$

γmu. For the quantum wave representing the particle, the group speed is $v_g = d\omega/dk$. Prove that the group speed of the wave is the same as the speed of the particle.

Section 40.7 The Double-Slit Experiment Revisited

51. M Neutrons traveling at 0.400 m/s are directed through a pair of slits separated by 1.00 mm. An array of detectors is placed 10.0 m from the slits. (a) What is the de Broglie wavelength of the neutrons? (b) How far off axis is the first zero-intensity point on the detector array? (c) When a neutron reaches a detector, can we say which slit the neutron passed through? Explain.

52. In a certain vacuum tube, electrons evaporate from a hot cathode at a slow, steady rate and accelerate from rest through a potential difference of 45.0 V. Then they travel 28.0 cm as they pass through an array of slits and fall on a screen to produce an interference pattern. If the beam current is below a certain value, only one electron at a time will be in flight in the tube. In this situation, the interference pattern still appears, showing that each individual electron can interfere with itself. What is the maximum value for the beam current that will result in only one electron at a time in flight in the tube?

53. W A modified oscilloscope is used to perform an electron interference experiment. Electrons are incident on a pair of narrow slits 0.060 0 μm apart. The bright bands in the interference pattern are separated by 0.400 mm on a screen 20.0 cm from the slits. Determine the potential difference through which the electrons were accelerated to give this pattern.

Section 40.8 The Uncertainty Principle

54. Suppose a duck lives in a universe in which $h = 2\pi$ J · s. The duck has a mass of 2.00 kg and is initially known to be within a pond 1.00 m wide. (a) What is the minimum uncertainty in the component of the duck's velocity parallel to the pond's width? (b) Assuming this uncertainty in speed prevails for 5.00 s, determine the uncertainty in the duck's position after this time interval.

55. M An electron and a 0.020 0-kg bullet each have a velocity of magnitude 500 m/s, accurate to within 0.010 0%. Within what lower limit could we determine the position of each object along the direction of the velocity?

56. A 0.500-kg block rests on the frictionless, icy surface of a frozen pond. If the location of the block is measured to a precision of 0.150 cm and its mass is known exactly, what is the minimum uncertainty in the block's speed?

57. The average lifetime of a muon is about 2 μs. Estimate the minimum uncertainty in the rest energy of a muon.

58. *Why is the following situation impossible?* An air rifle is used to shoot 1.00-g particles at a speed of v_x = 100 m/s. The rifle's barrel has a diameter of 2.00 mm. The rifle is mounted on a perfectly rigid support so that it is fired in exactly the same way each time. Because of the uncertainty principle, however, after many firings, the diameter of the spray of pellets on a paper target is 1.00 cm.

59. Use the uncertainty principle to show that if an electron were confined inside an atomic nucleus of diameter on the order of 10^{-14} m, it would have to be moving relativistically, whereas a proton confined to the same nucleus can be moving nonrelativistically.

Additional Problems

60. The accompanying table shows data obtained in a photoelectric experiment. (a) Using these data, make a graph similar to Figure 40.11 that plots as a straight line. From the graph, determine (b) an experimental value for Planck's constant (in joule-seconds) and (c) the work function (in electron volts) for the surface. (Two significant figures for each answer are sufficient.)

Wavelength (nm)	Maximum Kinetic Energy of Photoelectrons (eV)
588	0.67
505	0.98
445	1.35
399	1.63

61. AMT M Photons of wavelength 450 nm are incident on a metal. The most energetic electrons ejected from the metal are bent into a circular arc of radius 20.0 cm by a magnetic field with a magnitude of 2.00×10^{-5} T. What is the work function of the metal?

62. **Review.** Photons of wavelength λ are incident on a metal. The most energetic electrons ejected from the metal are bent into a circular arc of radius R by a magnetic field having a magnitude B. What is the work function of the metal?

63. **Review.** Design an incandescent lamp filament. A tungsten wire radiates electromagnetic waves with power 75.0 W when its ends are connected across a 120-V power supply. Assume its constant operating temperature is 2 900 K and its emissivity is 0.450. Also assume it takes in energy only by electric transmission and emits energy only by electromagnetic radiation. You may take the resistivity of tungsten at 2 900 K as 7.13×10^{-7} Ω · m. Specify (a) the radius and (b) the length of the filament.

64. Derive the equation for the Compton shift (Eq. 40.13) from Equations 40.14 through 40.16.

65. Figure P40.65 shows the stopping potential versus the incident photon frequency for the photoelectric effect

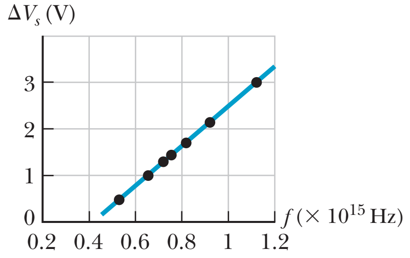


Figure P40.65

for sodium. Use the graph to find (a) the work function of sodium, (b) the ratio h/e, and (c) the cutoff wavelength. The data are taken from R. A. Millikan, *Physical Review* **7**:362 (1916).

66. A photon of initial energy E_0 undergoes Compton scattering at an angle θ from a free electron (mass m_e) initially at rest. Derive the following relationship for the final energy E' of the scattered photon:

$$E' = \frac{E_0}{1 + \left(\dfrac{E_0}{m_e c^2}\right)(1 - \cos\theta)}$$

67. A daredevil's favorite trick is to step out of a 16th-story window and fall 50.0 m into a pool. A news reporter takes a picture of the 75.0-kg daredevil just before he makes a splash, using an exposure time of 5.00 ms. Find (a) the daredevil's de Broglie wavelength at this moment, (b) the uncertainty of his kinetic energy measurement during the 5.00-ms time interval, and (c) the percent error caused by such an uncertainty.

68. Show that the ratio of the Compton wavelength λ_C to the de Broglie wavelength $\lambda = h/p$ for a relativistic electron is

$$\frac{\lambda_C}{\lambda} = \left[\left(\frac{E}{m_e c^2}\right)^2 - 1\right]^{1/2}$$

where E is the total energy of the electron and m_e is its mass.

69. Monochromatic ultraviolet light with intensity 550 W/m^2 is incident normally on the surface of a metal that has a work function of 3.44 eV. Photoelectrons are emitted with a maximum speed of 420 km/s. (a) Find the maximum possible rate of photoelectron emission from 1.00 cm^2 of the surface by imagining that every photon produces one photoelectron. (b) Find the electric current these electrons constitute. (c) How do you suppose the actual current compares with this maximum possible current?

70. A π^0 meson is an unstable particle produced in high-energy particle collisions. Its rest energy is approximately 135 MeV, and it exists for a lifetime of only 8.70×10^{-17} s before decaying into two gamma rays. Using the uncertainty principle, estimate the fractional uncertainty $\Delta m/m$ in its mass determination.

71. The neutron has a mass of 1.67×10^{-27} kg. Neutrons emitted in nuclear reactions can be slowed down by collisions with matter. They are referred to as thermal neutrons after they come into thermal equilibrium with the environment. The average kinetic energy $(\frac{3}{2}k_B T)$ of a thermal neutron is approximately 0.04 eV. (a) Calculate the de Broglie wavelength of a neutron with a kinetic energy of 0.040 0 eV. (b) How does your answer compare with the characteristic atomic spacing in a crystal? (c) Explain whether you expect thermal neutrons to exhibit diffraction effects when scattered by a crystal.

Challenge Problems

72. A woman on a ladder drops small pellets toward a point target on the floor. (a) Show that, according to the uncertainty principle, the average miss distance must be at least

$$\Delta x_f = \left(\frac{2\hbar}{m}\right)^{1/2}\left(\frac{2H}{g}\right)^{1/4}$$

where H is the initial height of each pellet above the floor and m is the mass of each pellet. Assume that the spread in impact points is given by $\Delta x_f = \Delta x_i + (\Delta v_x)t$. (b) If $H = 2.00$ m and $m = 0.500$ g, what is Δx_f?

73. Review. A light source emitting radiation at frequency 7.00×10^{14} Hz is incapable of ejecting photoelectrons from a certain metal. In an attempt to use this source to eject photoelectrons from the metal, the source is given a velocity toward the metal. (a) Explain how this procedure can produce photoelectrons. (b) When the speed of the light source is equal to $0.280c$, photoelectrons just begin to be ejected from the metal. What is the work function of the metal? (c) When the speed of the light source is increased to $0.900c$, determine the maximum kinetic energy of the photoelectrons.

74. Using conservation principles, prove that a photon cannot transfer all its energy to a free electron.

75. The total power per unit area radiated by a black body at a temperature T is the area under the $I(\lambda, T)$-versus-λ curve as shown in Figure 40.3. (a) Show that this power per unit area is

$$\int_0^\infty I(\lambda, T)\, d\lambda = \sigma T^4$$

where $I(\lambda, T)$ is given by Planck's radiation law and σ is a constant independent of T. This result is known as Stefan's law. (See Section 20.7.) To carry out the integration, you should make the change of variable $x = hc/\lambda k_B T$ and use

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

(b) Show that the Stefan–Boltzmann constant σ has the value

$$\sigma = \frac{2\pi^5 k_B^4}{15 c^2 h^3} = 5.67 \times 10^{-8}\ \text{W/m}^2 \cdot \text{K}^4$$

76. (a) Derive Wien's displacement law from Planck's law. Proceed as follows. In Figure 40.3, notice that the wavelength at which a black body radiates with greatest intensity is the wavelength for which the graph of $I(\lambda, T)$ versus λ has a horizontal tangent. From Equation 40.6, evaluate the derivative $dI/d\lambda$. Set it equal to zero. Solve the resulting transcendental equation numerically to prove that $hc/\lambda_{max} k_B T = 4.965\ldots$ or $\lambda_{max} T = hc/4.965 k_B$. (b) Evaluate the constant as precisely as possible and compare it with Wien's experimental value.

Quantum Mechanics

CHAPTER

41

An opened flash drive of the type used as an external data storage device for a computer. Flash drives are employed extensively in computers, digital cameras, cell phones, and other devices. Writing data to and erasing data from flash drives incorporate the phenomenon of quantum tunneling, which we explore in this chapter. *(Image copyright Vasilius, 2009. Used under license from Shutterstock.com)*

In this chapter, we introduce quantum mechanics, an extremely successful theory for explaining the behavior of microscopic particles. This theory, developed in the 1920s by Erwin Schrödinger, Werner Heisenberg, and others, enables us to understand a host of phenomena involving atoms, molecules, nuclei, and solids. The discussion in this chapter follows from the quantum particle model that was developed in Chapter 40 and incorporates some of the features of the waves under boundary conditions model that was explored in Chapter 18. We also discuss practical applications of quantum mechanics, including the scanning tunneling microscope and nanoscale devices that may be used in future quantum computers. Finally, we shall return to the simple harmonic oscillator that was introduced in Chapter 15 and examine it from a quantum mechanical point of view.

41.1 The Wave Function

In Chapter 40, we introduced some new and strange ideas. In particular, we concluded on the basis of experimental evidence that both matter and electromagnetic radiation are sometimes best modeled as particles and sometimes as waves, depending on the phenomenon being observed. We can improve our understanding of quantum physics by making another connection between particles and waves using the notion of probability, a concept that was introduced in Chapter 40.

We begin by discussing electromagnetic radiation using the particle model. The probability per unit volume of finding a photon in a given region of space at an instant of time is proportional to the number of photons per unit volume at that time:

$$\frac{\text{Probability}}{V} \propto \frac{N}{V}$$

The number of photons per unit volume is proportional to the intensity of the radiation:

$$\frac{N}{V} \propto I$$

Now, let's form a connection between the particle model and the wave model by recalling that the intensity of electromagnetic radiation is proportional to the square of the electric field amplitude E for the electromagnetic wave (Eq. 34.24):

$$I \propto E^2$$

Equating the beginning and the end of this series of proportionalities gives

$$\frac{\text{Probability}}{V} \propto E^2 \qquad \textbf{(41.1)}$$

Therefore, for electromagnetic radiation, the probability per unit volume of finding a particle associated with this radiation (the photon) is proportional to the square of the amplitude of the associated electromagnetic wave.

Recognizing the wave–particle duality of both electromagnetic radiation and matter, we should suspect a parallel proportionality for a material particle: the probability per unit volume of finding the particle is proportional to the square of the amplitude of a wave representing the particle. In Chapter 40, we learned that there is a de Broglie wave associated with every particle. The amplitude of the de Broglie wave associated with a particle is not a measurable quantity because the wave function representing a particle is generally a complex function as we discuss below. In contrast, the electric field for an electromagnetic wave is a real function. The matter analog to Equation 41.1 relates the square of the amplitude of the wave to the probability per unit volume of finding the particle. Hence, the amplitude of the wave associated with the particle is called the **probability amplitude,** or the **wave function,** and it has the symbol Ψ.

In general, the complete wave function Ψ for a system depends on the positions of all the particles in the system and on time; therefore, it can be written $\Psi(\vec{\mathbf{r}}_1, \vec{\mathbf{r}}_2, \vec{\mathbf{r}}_3, \ldots, \vec{\mathbf{r}}_j, \ldots, t)$, where $\vec{\mathbf{r}}_j$ is the position vector of the jth particle in the system. For many systems of interest, including all those we study in this text, the wave function Ψ is mathematically separable in space and time and can be written as a product of a space function ψ for one particle of the system and a complex time function:[1]

Space- and time-dependent wave function Ψ

$$\Psi(\vec{\mathbf{r}}_1, \vec{\mathbf{r}}_2, \vec{\mathbf{r}}_3, \ldots, \vec{\mathbf{r}}_j, \ldots, t) = \psi(\vec{\mathbf{r}}_j)e^{-i\omega t} \qquad \textbf{(41.2)}$$

where ω ($= 2\pi f$) is the angular frequency of the wave function and $i = \sqrt{-1}$.

For any system in which the potential energy is time-independent and depends only on the positions of particles within the system, the important information about the system is contained within the space part of the wave function. The time part is simply the factor $e^{-i\omega t}$. Therefore, an understanding of ψ is the critical aspect of a given problem.

The wave function ψ is often complex-valued. The absolute square $|\psi|^2 = \psi^*\psi$, where ψ^* is the complex conjugate[2] of ψ, is always real and positive and is proportional to the probability per unit volume of finding a particle at a given point at some instant. The wave function contains within it all the information that can be known about the particle.

---

[1]The standard form of a complex number is $a + ib$. The notation $e^{i\theta}$ is equivalent to the standard form as follows:

$$e^{i\theta} = \cos\theta + i\sin\theta$$

Therefore, the notation $e^{-i\omega t}$ in Equation 41.2 is equivalent to $\cos(-\omega t) + i\sin(-\omega t) = \cos\omega t - i\sin\omega t$.

[2]For a complex number $z = a + ib$, the complex conjugate is found by changing i to $-i$: $z^* = a - ib$. The product of a complex number and its complex conjugate is always real and positive. That is, $z^*z = (a - ib)(a + ib) = a^2 - (ib)^2 = a^2 - (i)^2b^2 = a^2 + b^2$.

Although ψ cannot be measured, we can measure the real quantity $|\psi|^2$, which can be interpreted as follows. If ψ represents a single particle, then $|\psi|^2$—called the **probability density**—is the relative probability per unit volume that the particle will be found at any given point in the volume. This interpretation can also be stated in the following manner. If dV is a small volume element surrounding some point, the probability of finding the particle in that volume element is

Probability density $|\psi|^2$

$$P(x, y, z)\ dV = |\psi|^2\ dV \tag{41.3}$$

This probabilistic interpretation of the wave function was first suggested by Max Born (1882–1970) in 1928. In 1926, Erwin Schrödinger proposed a wave equation that describes the manner in which the wave function changes in space and time. The *Schrödinger wave equation,* which we shall examine in Section 41.3, represents a key element in the theory of quantum mechanics.

The concepts of quantum mechanics, strange as they sometimes may seem, developed from classical ideas. In fact, when the techniques of quantum mechanics are applied to macroscopic systems, the results are essentially identical to those of classical physics. This blending of the two approaches occurs when the de Broglie wavelength is small compared with the dimensions of the system. The situation is similar to the agreement between relativistic mechanics and classical mechanics when $v << c$.

Pitfall Prevention 41.1

The Wave Function Belongs to a System The common language in quantum mechanics is to associate a wave function with a particle. The wave function, however, is determined by the particle *and* its interaction with its environment, so it more rightfully belongs to a system. In many cases, the particle is the only part of the system that experiences a change, which is why the common language has developed. You will see examples in the future in which it is more proper to think of the system wave function rather than the particle wave function.

In Section 40.5, we found that the de Broglie equation relates the momentum of a particle to its wavelength through the relation $p = h/\lambda$. If an ideal free particle has a precisely known momentum p_x, its wave function is an infinitely long sinusoidal wave of wavelength $\lambda = h/p_x$ and the particle has equal probability of being at any point along the x axis (Fig. 40.18a). The wave function ψ for such a free particle moving along the x axis can be written as

Wave function for a free particle

$$\psi(x) = Ae^{ikx} \tag{41.4}$$

where A is a constant amplitude and $k = 2\pi/\lambda$ is the angular wave number (Eq. 16.8) of the wave representing the particle.[3]

One-Dimensional Wave Functions and Expectation Values

This section discusses only one-dimensional systems, where the particle must be located along the x axis, so the probability $|\psi|^2\ dV$ in Equation 41.3 is modified to become $|\psi|^2\ dx$. The probability that the particle will be found in the infinitesimal interval dx around the point x is

$$P(x)\ dx = |\psi|^2\ dx \tag{41.5}$$

Although it is not possible to specify the position of a particle with complete certainty, it is possible through $|\psi|^2$ to specify the probability of observing it in a region surrounding a given point x. The probability of finding the particle in the arbitrary interval $a \leq x \leq b$ is

$$P_{ab} = \int_a^b |\psi|^2\, dx \tag{41.6}$$

The probability P_{ab} is the area under the curve of $|\psi|^2$ versus x between the points $x = a$ and $x = b$ as in Figure 41.1.

Experimentally, there is a finite probability of finding a particle in an interval near some point at some instant. The value of that probability must lie between the

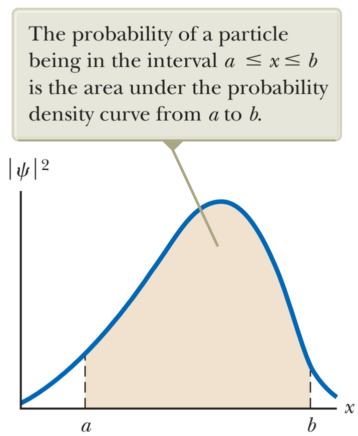


Figure 41.1 An arbitrary probability density curve for a particle.

[3]For the free particle, the full wave function, based on Equation 41.2, is

$$\Psi(x, t) = Ae^{ikx}e^{-i\omega t} = Ae^{i(kx-\omega t)} = A[\cos(kx - \omega t) + i \sin(kx - \omega t)]$$

The real part of this wave function has the same form as the waves we added together to form wave packets in Section 40.6.

limits 0 and 1. For example, if the probability is 0.30, there is a 30% chance of finding the particle in the interval.

Because the particle must be somewhere along the x axis, the sum of the probabilities over all values of x must be 1:

Normalization condition on ψ

$$\int_{-\infty}^{\infty} |\psi|^2 \, dx = 1 \tag{41.7}$$

Any wave function satisfying Equation 41.7 is said to be **normalized.** Normalization is simply a statement that the particle exists at some point in space.

Once the wave function for a particle is known, it is possible to calculate the average position at which you would expect to find the particle after many measurements. This average position is called the **expectation value** of x and is defined by the equation

Expectation value for position x

$$\langle x \rangle \equiv \int_{-\infty}^{\infty} \psi^* x \psi \, dx \tag{41.8}$$

(Brackets, $\langle \ldots \rangle$, are used to denote expectation values.) Furthermore, one can find the expectation value of any function $f(x)$ associated with the particle by using the following equation:[4]

Expectation value for a function $f(x)$

$$\langle f(x) \rangle \equiv \int_{-\infty}^{\infty} \psi^* f(x) \psi \, dx \tag{41.9}$$

Quick Quiz 41.1 Consider the wave function for the free particle, Equation 41.4. At what value of x is the particle most likely to be found at a given time? **(a)** at $x = 0$ **(b)** at small nonzero values of x **(c)** at large values of x **(d)** anywhere along the x axis

Example 41.1 A Wave Function for a Particle

Consider a particle whose wave function is graphed in Figure 41.2 and is given by

$$\psi(x) = Ae^{-ax^2}$$

(A) What is the value of A if this wave function is normalized?

SOLUTION

Conceptualize The particle is not a free particle because the wave function is not a sinusoidal function. Figure 41.2 indicates that the particle is constrained to remain close to $x = 0$ at all times. Think of a physical system in which the particle always stays close to a given point. Examples of such systems are a block on a spring, a marble at the bottom of a bowl, and the bob of a simple pendulum.

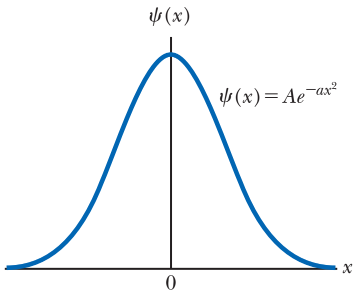


Figure 41.2 (Example 41.1) A symmetric wave function for a particle, given by $\psi(x) = Ae^{-ax^2}$.

Categorize Because the statement of the problem describes the wave nature of a particle, this example requires a quantum approach rather than a classical approach.

Analyze Apply the normalization condition, Equation 41.7, to the wave function:

$$\int_{-\infty}^{\infty} |\psi|^2 \, dx = \int_{-\infty}^{\infty} (Ae^{-ax^2})^2 \, dx = A^2 \int_{-\infty}^{\infty} e^{-2ax^2} \, dx = 1$$

[4] Expectation values are analogous to "weighted averages," in which each possible value of a function is multiplied by the probability of the occurrence of that value before summing over all possible values. We write the expectation value as $\int_{-\infty}^{\infty} \psi^* f(x) \psi \, dx$ rather than $\int_{-\infty}^{\infty} f(x) \psi^2 \, dx$ because $f(x)$ may be represented by an operator (such as a derivative) rather than a simple multiplicative function in more advanced treatments of quantum mechanics. In these situations, the operator is applied only to ψ and not to ψ^*.

41.1 continued

Express the integral as the sum of two integrals:

$$(1)\quad A^2\int_{-\infty}^{\infty} e^{-2ax^2}\,dx = A^2\left(\int_0^{\infty} e^{-2ax^2}\,dx + \int_{-\infty}^{0} e^{-2ax^2}\,dx\right) = 1$$

Change the integration variable from x to $-x$ in the second integral:

$$\int_{-\infty}^{0} e^{-2ax^2}\,dx = \int_{\infty}^{0} e^{-2a(-x)^2}\,(-dx) = -\int_{\infty}^{0} e^{-2ax^2}\,dx$$

Reverse the order of the limits, which introduces a negative sign:

$$-\int_{\infty}^{0} e^{-2ax^2}\,dx = \int_0^{\infty} e^{-2ax^2}\,dx$$

Substitute this expression for the second integral in Equation (1):

$$A^2\left(\int_0^{\infty} e^{-2ax^2}\,dx + \int_0^{\infty} e^{-2ax^2}\,dx\right) = 1$$

$$(2)\quad 2A^2\int_0^{\infty} e^{-2ax^2}\,dx = 1$$

Evaluate the integral with the help of Table B.6 in Appendix B:

$$\int_0^{\infty} e^{-2ax^2}\,dx = \tfrac{1}{2}\sqrt{\frac{\pi}{2a}}$$

Substitute this result into Equation (2) and solve for A:

$$2A^2\left(\tfrac{1}{2}\sqrt{\frac{\pi}{2a}}\right) = 1 \;\rightarrow\; A = \boxed{\left(\frac{2a}{\pi}\right)^{1/4}}$$

(B) What is the expectation value of x for this particle?

SOLUTION

Evaluate the expectation value using Equation 41.8:

$$\langle x\rangle \equiv \int_{-\infty}^{\infty} \psi^* x\psi\,dx = \int_{-\infty}^{\infty} (Ae^{-ax^2})x(Ae^{-ax^2})\,dx$$

$$= A^2\int_{-\infty}^{\infty} xe^{-2ax^2}\,dx$$

As in part (A), express the integral as a sum of two integrals:

$$(3)\quad \langle x\rangle = A^2\left(\int_0^{\infty} xe^{-2ax^2}\,dx + \int_{-\infty}^{0} xe^{-2ax^2}\,dx\right)$$

Change the integration variable from x to $-x$ in the second integral:

$$\int_{-\infty}^{0} xe^{-2ax^2}\,dx = \int_{\infty}^{0} -xe^{-2a(-x)^2}\,(-dx) = \int_{\infty}^{0} xe^{-2ax^2}\,dx$$

Reverse the order of the limits, which introduces a negative sign:

$$\int_{\infty}^{0} xe^{-2ax^2}\,dx = -\int_0^{\infty} xe^{-2ax^2}\,dx$$

Substitute this expression for the second integral in Equation (3):

$$\langle x\rangle = A^2\left(\int_0^{\infty} xe^{-2ax^2}\,dx - \int_0^{\infty} xe^{-2ax^2}\,dx\right) = \boxed{0}$$

Finalize Given the symmetry of the wave function around $x = 0$ in Figure 41.2, it is not surprising that the average position of the particle is at $x = 0$. In Section 41.7, we show that the wave function studied in this example represents the lowest-energy state of the quantum harmonic oscillator.

41.2 Analysis Model: Quantum Particle Under Boundary Conditions

The free particle discussed in Section 41.1 has no boundary conditions; it can be anywhere in space. The particle in Example 41.1 is not a free particle. Figure 41.2 shows that the particle is always restricted to positions near $x = 0$. In this section, we shall investigate the effects of restrictions on the motion of a quantum particle.

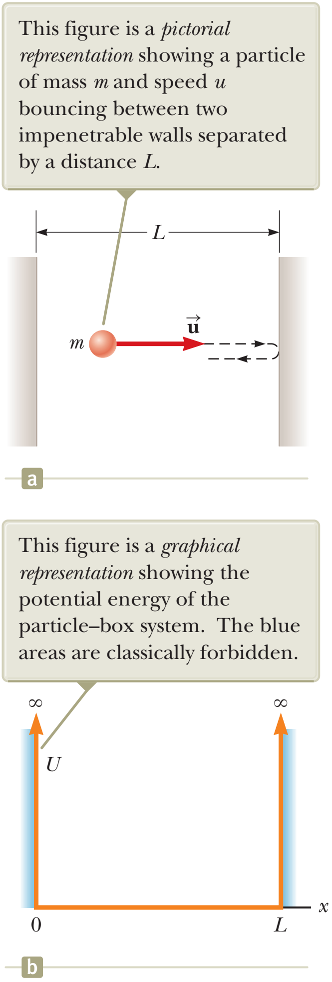


Figure 41.3 (a) The particle in a box. (b) The potential energy function for the system.

A Particle in a Box

We begin by applying some of the ideas we have developed to a simple physical problem, a particle confined to a one-dimensional region of space, called the *particle-in-a-box* problem (even though the "box" is one-dimensional!). From a classical viewpoint, if a particle is bouncing elastically back and forth along the x axis between two impenetrable walls separated by a distance L as in Figure 41.3a, it can be modeled as a particle under constant speed. If the speed of the particle is u, the magnitude of its momentum mu remains constant as does its kinetic energy. (Recall that in Chapter 39 we used u for particle speed to distinguish it from v, the speed of a reference frame.) Classical physics places no restrictions on the values of a particle's momentum and energy. The quantum-mechanical approach to this problem is quite different and requires that we find the appropriate wave function consistent with the conditions of the situation.

Because the walls are impenetrable, there is zero probability of finding the particle outside the box, so the wave function $\psi(x)$ must be zero for $x < 0$ and $x > L$. To be a mathematically well-behaved function, $\psi(x)$ must be continuous in space. There must be no discontinuous jumps in the value of the wave function at any point.[5] Therefore, if ψ is zero outside the walls, it must also be zero *at* the walls; that is, $\psi(0) = 0$ and $\psi(L) = 0$. Only those wave functions that satisfy these boundary conditions are allowed.

Figure 41.3b, a graphical representation of the particle-in-a-box problem, shows the potential energy of the particle–environment system as a function of the position of the particle. As long as the particle is inside the box, the potential energy of the system does not depend on the location of the particle and we can choose its constant value to be zero. Outside the box, we must ensure that the wave function is zero. We can do so by defining the system's potential energy as infinitely large if the particle were outside the box. Therefore, the only way a particle could be outside the box is if the system has an infinite amount of energy, which is impossible.

The wave function for a particle in the box can be expressed as a real sinusoidal function:[6]

$$\psi(x) = A \sin\left(\frac{2\pi x}{\lambda}\right) \tag{41.10}$$

where λ is the de Broglie wavelength associated with the particle. This wave function must satisfy the boundary conditions at the walls. The boundary condition $\psi(0) = 0$ is satisfied already because the sine function is zero when $x = 0$. The boundary condition $\psi(L) = 0$ gives

$$\psi(L) = 0 = A \sin\left(\frac{2\pi L}{\lambda}\right)$$

which can only be true if

$$\frac{2\pi L}{\lambda} = n\pi \quad \to \quad \lambda = \frac{2L}{n} \tag{41.11}$$

where $n = 1, 2, 3, \ldots$. Therefore, only certain wavelengths for the particle are allowed! Each of the allowed wavelengths corresponds to a quantum state for the system, and n is the quantum number. Incorporating Equation 41.11 in Equation 41.10 gives

Wave functions for a particle in a box ▶

$$\psi_n(x) = A \sin\left(\frac{2\pi x}{2L/n}\right) = A \sin\left(\frac{n\pi x}{L}\right) \tag{41.12}$$

[5]If the wave function were not continuous at a point, the derivative of the wave function at that point would be infinite. This result leads to difficulties in the Schrödinger equation, for which the wave function is a solution as discussed in Section 41.3.

[6]We shall show this result explicitly in Section 41.3.

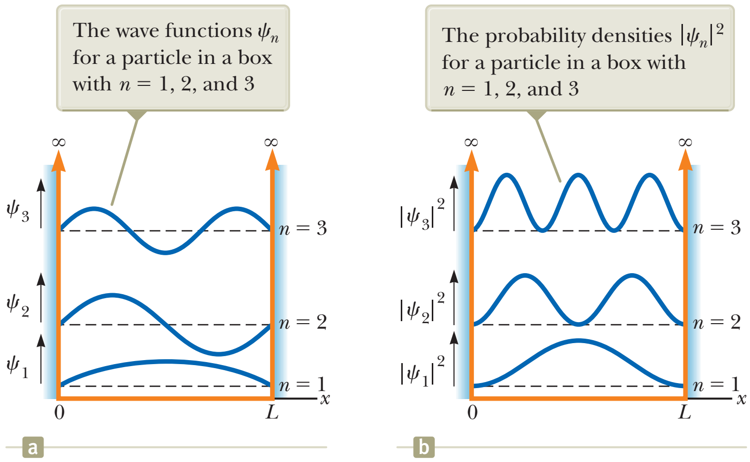


Figure 41.4 The first three allowed states for a particle confined to a one-dimensional box. The states are shown superimposed on the potential energy function of Figure 41.3b. The wave functions and probability densities are plotted vertically from separate axes that are offset vertically for clarity. The positions of these axes on the potential energy function suggest the relative energies of the states.

Normalizing this wave function shows that $A = \sqrt{2/L}$. (See Problem 18.) Therefore, the normalized wave function for the particle in a box is

$$\psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right) \quad (41.13)$$

◀ **Normalized wave function for a particle in a box**

Figures 41.4a and b are graphical representations of ψ_n versus x and $|\psi_n|^2$ versus x for n = 1, 2, and 3 for the particle in a box.[7] Although a general wave function ψ can have positive and negative values, $|\psi|^2$ is always positive. Because $|\psi|^2$ represents a probability density, a negative value for $|\psi|^2$ would be meaningless.

Further inspection of Figure 41.4b shows that $|\psi|^2$ is zero at the boundaries, satisfying our boundary conditions. In addition, $|\psi|^2$ is zero at other points, depending on the value of n. For $n = 2$, $|\psi_2|^2 = 0$ at $x = L/2$; for $n = 3$, $|\psi_3|^2 = 0$ at $x = L/3$ and at $x = 2L/3$. The number of zero points increases by one each time the quantum number increases by one.

Because the wavelengths of the particle are restricted by the condition $\lambda = 2L/n$, the magnitude of the momentum of the particle is also restricted to specific values, which can be found from the expression for the de Broglie wavelength, Equation 40.17:

$$p = \frac{h}{\lambda} = \frac{h}{2L/n} = \frac{nh}{2L}$$

We have chosen the potential energy of the system to be zero when the particle is inside the box. Therefore, the energy of the system is simply the kinetic energy of the particle and the allowed values are given by

$$E_n = \tfrac{1}{2}mu^2 = \frac{p^2}{2m} = \frac{(nh/2L)^2}{2m}$$

$$\boxed{E_n = \left(\frac{h^2}{8mL^2}\right)n^2 \quad n = 1, 2, 3, \ldots} \quad (41.14)$$

◀ **Quantized energies for a particle in a box**

Pitfall Prevention 41.2

Reminder: Energy Belongs to a System We often refer to the energy of a particle in commonly used language. As in Pitfall Prevention 41.1, we are actually describing the energy of the *system* of the particle and whatever environment is establishing the impenetrable walls. For the particle in a box, the only type of energy is kinetic energy belonging to the particle, which is the origin of the common description.

This expression shows that the energy of the particle is quantized. The lowest allowed energy corresponds to the **ground state,** which is the lowest energy state for any system. For the particle in a box, the ground state corresponds to $n = 1$, for which $E_1 = h^2/8mL^2$. Because $E_n = n^2E_1$, the **excited states** corresponding to n = 2, 3, 4, . . . have energies given by $4E_1$, $9E_1$, $16E_1$, ...

[7]Note that $n = 0$ is not allowed because, according to Equation 41.12, the wave function would be $\psi = 0$, which is not a physically reasonable wave function. For example, it cannot be normalized because $\int_{-\infty}^{\infty}|\psi|^2\,dx = \int_{-\infty}^{\infty}(0)\,dx = 0$, but Equation 41.7 tells us that this integral must equal 1.

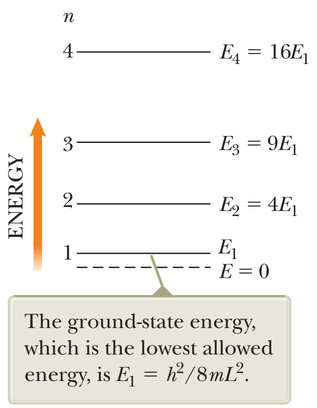


Figure 41.5 Energy-level diagram for a particle confined to a one-dimensional box of length L.

Figure 41.5 is an energy-level diagram describing the energy values of the allowed states. Because the lowest energy of the particle in a box is not zero, then, according to quantum mechanics, the particle can never be at rest! The smallest energy it can have, corresponding to $n = 1$, is called the **ground-state energy.** This result contradicts the classical viewpoint, in which $E = 0$ is an acceptable state, as are *all* positive values of E.

Quick Quiz 41.2 Consider an electron, a proton, and an alpha particle (a helium nucleus), each trapped separately in identical boxes. **(i)** Which particle corresponds to the highest ground-state energy? (a) the electron (b) the proton (c) the alpha particle (d) The ground-state energy is the same in all three cases. **(ii)** Which particle has the longest wavelength when the system is in the ground state? (a) the electron (b) the proton (c) the alpha particle (d) All three particles have the same wavelength.

Quick Quiz 41.3 A particle is in a box of length L. Suddenly, the length of the box is increased to $2L$. What happens to the energy levels shown in Figure 41.5? **(a)** nothing; they are unaffected. **(b)** They move farther apart. **(c)** They move closer together.

Example 41.2 Microscopic and Macroscopic Particles in Boxes

(A) An electron is confined between two impenetrable walls 0.200 nm apart. Determine the energy levels for the states $n = 1$, 2, and 3.

SOLUTION

Conceptualize In Figure 41.3a, imagine that the particle is an electron and the walls are very close together.

Categorize We evaluate the energy levels using an equation developed in this section, so we categorize this example as a substitution problem.

Use Equation 41.14 for the $n = 1$ state:

$$E_1 = \frac{h^2}{8m_eL^2}(1)^2 = \frac{(6.63 \times 10^{-34}\ \mathrm{J\cdot s})^2}{8(9.11 \times 10^{-31}\ \mathrm{kg})(2.00 \times 10^{-10}\ \mathrm{m})^2}$$

$$= 1.51 \times 10^{-18}\ \mathrm{J} = \boxed{9.42\ \mathrm{eV}}$$

Using $E_n = n^2E_1$, find the energies of the $n = 2$ and $n = 3$ states:

$$E_2 = (2)^2E_1 = 4(9.42\ \mathrm{eV}) = \boxed{37.7\ \mathrm{eV}}$$

$$E_3 = (3)^2E_1 = 9(9.42\ \mathrm{eV}) = \boxed{84.8\ \mathrm{eV}}$$

(B) Find the speed of the electron in the $n = 1$ state.

SOLUTION

Solve the classical expression for kinetic energy for the particle speed:

$$K = \tfrac{1}{2}m_eu^2 \rightarrow u = \sqrt{\frac{2K}{m_e}}$$

Recognize that the kinetic energy of the particle is equal to the system energy and substitute E_n for K:

$$(1)\quad u = \sqrt{\frac{2E_n}{m_e}}$$

Substitute numerical values from part (A):

$$u = \sqrt{\frac{2(1.51 \times 10^{-18}\ \mathrm{J})}{9.11 \times 10^{-31}\ \mathrm{kg}}} = \boxed{1.82 \times 10^6\ \mathrm{m/s}}$$

Simply placing the electron in the box results in a *minimum* speed of the electron equal to 0.6% of the speed of light!

(C) A 0.500-kg baseball is confined between two rigid walls of a stadium that can be modeled as a box of length 100 m. Calculate the minimum speed of the baseball.

41.2 continued

SOLUTION

Conceptualize In Figure 41.3a, imagine that the particle is a baseball and the walls are those of the stadium.

Categorize This part of the example is a substitution problem in which we apply a quantum approach to a macroscopic object.

Use Equation 41.14 for the $n = 1$ state:

$$E_1 = \frac{h^2}{8mL^2}(1)^2 = \frac{(6.63 \times 10^{-34}\ \mathrm{J \cdot s})^2}{8(0.500\ \mathrm{kg})(100\ \mathrm{m})^2} = 1.10 \times 10^{-71}\ \mathrm{J}$$

Use Equation (1) to find the speed:

$$u = \sqrt{\frac{2(1.10 \times 10^{-71}\ \mathrm{J})}{0.500\ \mathrm{kg}}} = \boxed{6.63 \times 10^{-36}\ \mathrm{m/s}}$$

This speed is so small that the object can be considered to be at rest, which is what one would expect for the minimum speed of a macroscopic object.

WHAT IF? What if a sharp line drive is hit so that the baseball is moving with a speed of 150 m/s? What is the quantum number of the state in which the baseball now resides?

Answer We expect the quantum number to be very large because the baseball is a macroscopic object.

Evaluate the kinetic energy of the baseball:

$$\tfrac{1}{2}mu^2 = \tfrac{1}{2}(0.500\ \mathrm{kg})(150\ \mathrm{m/s})^2 = 5.62 \times 10^3\ \mathrm{J}$$

From Equation 41.14, calculate the quantum number n:

$$n = \sqrt{\frac{8mL^2E_n}{h^2}} = \sqrt{\frac{8(0.500\ \mathrm{kg})(100\ \mathrm{m})^2(5.62 \times 10^3\ \mathrm{J})}{(6.63 \times 10^{-34}\ \mathrm{J \cdot s})^2}} = 2.26 \times 10^{37}$$

This result is a tremendously large quantum number. As the baseball pushes air out of the way, hits the ground, and rolls to a stop, it moves through more than 10^{37} quantum states. These states are so close together in energy that we cannot observe the transitions from one state to the next. Rather, we see what appears to be a smooth variation in the speed of the ball. The quantum nature of the universe is simply not evident in the motion of macroscopic objects.

Example 41.3 The Expectation Values for the Particle in a Box

A particle of mass m is confined to a one-dimensional box between $x = 0$ and $x = L$. Find the expectation value of the position x of the particle in the state characterized by quantum number n.

SOLUTION

Conceptualize Figure 41.4b shows that the probability for the particle to be at a given location varies with position within the box. Can you predict what the expectation value of x will be from the symmetry of the wave functions?

Categorize The statement of the example categorizes the problem for us: we focus on a quantum particle in a box and on the calculation of its expectation value of x.

Analyze In Equation 41.8, the integration from $-\infty$ to ∞ reduces to the limits 0 to L because $\psi = 0$ everywhere except in the box.

Substitute Equation 41.13 into Equation 41.8 to find the expectation value for x:

$$\langle x \rangle = \int_{-\infty}^{\infty} \psi_n{}^* x \psi_n\, dx = \int_0^L x\left[\sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right)\right]^2 dx$$

$$= \frac{2}{L}\int_0^L x \sin^2\left(\frac{n\pi x}{L}\right) dx$$

continued

41.3 continued

Evaluate the integral by consulting an integral table or by mathematical integration:[8]

$$\langle x \rangle = \frac{2}{L}\left[\frac{x^2}{4} - \frac{x\sin\left(2\dfrac{n\pi x}{L}\right)}{4\dfrac{n\pi}{L}} - \frac{\cos\left(2\dfrac{n\pi x}{L}\right)}{8\left(\dfrac{n\pi}{L}\right)^2}\right]_0^L$$

$$= \frac{2}{L}\left[\frac{L^2}{4}\right] = \boxed{\frac{L}{2}}$$

Finalize This result shows that the expectation value of x is at the center of the box for all values of n, which you would expect from the symmetry of the square of the wave functions (the probability density) about the center (Fig. 41.4b).

The $n = 2$ wave function in Figure 41.4b has a value of zero at the midpoint of the box. Can the expectation value of the particle be at a position at which the particle has zero probability of existing? Remember that the expectation value is the *average* position. Therefore, the particle is as likely to be found to the right of the midpoint as to the left, so its average position is at the midpoint even though its probability of being there is zero. As an analogy, consider a group of students for whom the average final examination score is 50%. There is no requirement that some student achieve a score of exactly 50% for the average of all students to be 50%.

Boundary Conditions on Particles in General

The discussion of the particle in a box is very similar to the discussion in Chapter 18 of standing waves on strings:

- Because the ends of the string must be nodes, the wave functions for allowed waves must be zero at the boundaries of the string. Because the particle in a box cannot exist outside the box, the allowed wave functions for the particle must be zero at the boundaries.
- The boundary conditions on the string waves lead to quantized wavelengths and frequencies of the waves. The boundary conditions on the wave function for the particle in a box lead to quantized wavelengths and frequencies of the particle.

In quantum mechanics, it is very common for particles to be subject to boundary conditions. We therefore introduce a new analysis model, the **quantum particle under boundary conditions.** In many ways, this model is similar to the waves under boundary conditions model studied in Section 18.3. In fact, the allowed wavelengths for the wave function of a particle in a box (Eq. 41.11) are identical in form to the allowed wavelengths for mechanical waves on a string fixed at both ends (Eq. 18.4).

The quantum particle under boundary conditions model *differs* in some ways from the waves under boundary conditions model:

- In most cases of quantum particles, the wave function is *not* a simple sinusoidal function like the wave function for waves on strings. Furthermore, the wave function for a quantum particle may be a complex function.
- For a quantum particle, frequency is related to energy through $E = hf$, so the quantized frequencies lead to quantized energies.
- There may be no stationary "nodes" associated with the wave function of a quantum particle under boundary conditions. Systems more complicated than the particle in a box have more complicated wave functions, and some boundary conditions may not lead to zeroes of the wave function at fixed points.

[8]To integrate this function, first replace $\sin^2(n\pi x/L)$ with $\frac{1}{2}(1 - \cos 2n\pi x/L)$ (refer to Table B.3 in Appendix B), which allows $\langle x \rangle$ to be expressed as two integrals. The second integral can then be evaluated by partial integration (Section B.7 in Appendix B).

In general,

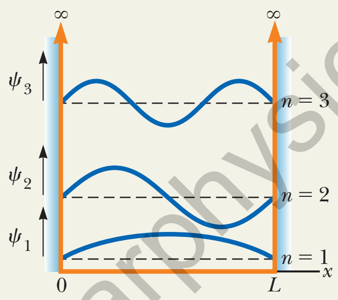
> an interaction of a quantum particle with its environment represents one or more boundary conditions, and, if the interaction restricts the particle to a finite region of space, results in quantization of the energy of the system.

Boundary conditions on quantum wave functions are related to the coordinates describing the problem. For the particle in a box, the wave function must be zero at two values of x. In the case of a three-dimensional system such as the hydrogen atom we shall discuss in Chapter 42, the problem is best presented in *spherical coordinates*. These coordinates, an extension of the plane polar coordinates introduced in Section 3.1, consist of a radial coordinate r and two angular coordinates. The generation of the wave function and application of the boundary conditions for the hydrogen atom are beyond the scope of this book. We shall, however, examine the behavior of some of the hydrogen-atom wave functions in Chapter 42.

Boundary conditions on wave functions that exist for all values of x require that the wave function approach zero as $x \to \infty$ (so that the wave function can be normalized) and remain finite as $x \to 0$. One boundary condition on any angular parts of wave functions is that adding 2π radians to the angle must return the wave function to the same value because an addition of 2π results in the same angular position.

Analysis Model **Quantum Particle Under Boundary Conditions**

Imagine a particle described by quantum physics that is subject to one or more boundary conditions. If the particle is restricted to a finite region of space by the boundary conditions, the energy of the system is quantized. Associated with each quantized energy is a quantum state characterized by a wave function and a quantum number.



Examples:

- an electron in a quantum dot cannot escape, quantizing the energies of the electron (Section 41.4)
- an electron in a hydrogen atom is restricted to stay near the nucleus of the atom, quantizing the energies of the atom (Chapter 42)
- two atoms are bound to form a diatomic molecule, quantizing the energies of vibration and rotation of the molecule (Chapter 43)
- a proton is trapped in a nucleus, quantizing its energy levels (Chapter 44)

41.3 The Schrödinger Equation

In Section 34.3, we discussed a linear wave equation for electromagnetic radiation that follows from Maxwell's equations. The waves associated with particles also satisfy a wave equation. The wave equation for material particles is different from that associated with photons because material particles have a nonzero rest energy. The appropriate wave equation was developed by Schrödinger in 1926. In analyzing the behavior of a quantum system, the approach is to determine a solution to this equation and then apply the appropriate boundary conditions to the solution. This process yields the allowed wave functions and energy levels of the system under consideration. Proper manipulation of the wave function then enables one to calculate all measurable features of the system.

The Schrödinger equation as it applies to a particle of mass m confined to moving along the x axis and interacting with its environment through a potential energy function $U(x)$ is

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + U\psi = E\psi \tag{41.15}$$

◀ **Time-independent Schrödinger equation**

© INTERFOTO/Alamy

Erwin Schrödinger
Austrian Theoretical Physicist (1887–1961)
Schrödinger is best known as one of the creators of quantum mechanics. His approach to quantum mechanics was demonstrated to be mathematically equivalent to the more abstract matrix mechanics developed by Heisenberg. Schrödinger also produced important papers in the fields of statistical mechanics, color vision, and general relativity.

where E is a constant equal to the total energy of the system (the particle and its environment). Because this equation is independent of time, it is commonly referred to as the **time-independent Schrödinger equation.** (We shall not discuss the time-dependent Schrödinger equation in this book.)

The Schrödinger equation is consistent with the principle of conservation of mechanical energy for an isolated system with no nonconservative forces acting. Problem 44 shows, both for a free particle and a particle in a box, that the first term in the Schrödinger equation reduces to the kinetic energy of the particle multiplied by the wave function. Therefore, Equation 41.15 indicates that the total energy of the system is the sum of the kinetic energy and the potential energy and that the total energy is a constant: $K + U = E =$ constant.

In principle, if the potential energy function U for a system is known, one can solve Equation 41.15 and obtain the wave functions and energies for the allowed states of the system. In addition, in many cases, the wave function ψ must satisfy boundary conditions. Therefore, once we have a preliminary solution to the Schrödinger equation, we impose the following conditions to find the exact solution and the allowed energies:

- ψ must be normalizable. That is, Equation 41.7 must be satisfied.
- ψ must go to 0 as $x \rightarrow \pm\infty$ and remain finite as $x \rightarrow 0$.
- ψ must be continuous in x and be single-valued everywhere; solutions to Equation 41.15 in different regions must join smoothly at the boundaries between the regions.
- $d\psi/dx$ must be finite, continuous, and single-valued everywhere for finite values of U. If $d\psi/dx$ were not continuous, we would not be able to evaluate the second derivative $d^2\psi/dx^2$ in Equation 41.15 at the point of discontinuity.

The task of solving the Schrödinger equation may be very difficult, depending on the form of the potential energy function. As it turns out, the Schrödinger equation is extremely successful in explaining the behavior of atomic and nuclear systems, whereas classical physics fails to explain this behavior. Furthermore, when quantum mechanics is applied to macroscopic objects, the results agree with classical physics.

The Particle in a Box Revisited

Pitfall Prevention 41.3

Potential Wells A potential well such as that in Figure 41.3b is a graphical representation of energy, not a pictorial representation, so you would not see this shape if you were able to observe the situation. A particle moves *only horizontally* at a fixed vertical position in a potential-energy diagram, representing the conserved energy of the system of the particle and its environment.

To see how the quantum particle under boundary conditions model is applied to a problem, let's return to our particle in a one-dimensional box of length L (see Fig. 41.3) and analyze it with the Schrödinger equation. Figure 41.3b is the potential-energy diagram that describes this problem. Potential-energy diagrams are a useful representation for understanding and solving problems with the Schrödinger equation.

Because of the shape of the curve in Figure 41.3b, the particle in a box is sometimes said to be in a **square well,**[9] where a **well** is an upward-facing region of the curve in a potential-energy diagram. (A downward-facing region is called a *barrier,* which we investigate in Section 41.5.) Figure 41.3b shows an infinite square well.

In the region $0 < x < L$, where $U = 0$, we can express the Schrödinger equation in the form

$$\frac{d^2\psi}{dx^2} = -\frac{2mE}{\hbar^2}\psi = -k^2\psi \tag{41.16}$$

where

$$k = \frac{\sqrt{2mE}}{\hbar}$$

[9]It is called a square well even if it has a rectangular shape in a potential-energy diagram.

The solution to Equation 41.16 is a function ψ whose second derivative is the negative of the same function multiplied by a constant k^2. Both the sine and cosine functions satisfy this requirement. Therefore, the most general solution to the equation is a linear combination of both solutions:

$$\psi(x) = A \sin kx + B \cos kx$$

where A and B are constants that are determined by the boundary and normalization conditions.

The first boundary condition on the wave function is that $\psi(0) = 0$:

$$\psi(0) = A \sin 0 + B \cos 0 = 0 + B = 0$$

which means that $B = 0$. Therefore, our solution reduces to

$$\psi(x) = A \sin kx$$

The second boundary condition, $\psi(L) = 0$, when applied to the reduced solution gives

$$\psi(L) = A \sin kL = 0$$

This equation could be satisfied by setting $A = 0$, but that would mean that $\psi = 0$ everywhere, which is not a valid wave function. The boundary condition is also satisfied if kL is an integral multiple of π, that is, if $kL = n\pi$, where n is an integer. Substituting $k = \sqrt{2mE}/\hbar$ into this expression gives

$$kL = \frac{\sqrt{2mE}}{\hbar} L = n\pi$$

Each value of the integer n corresponds to a quantized energy that we call E_n. Solving for the allowed energies E_n gives

$$E_n = \left(\frac{h^2}{8mL^2}\right) n^2 \qquad \textbf{(41.17)}$$

which are identical to the allowed energies in Equation 41.14.

Substituting the values of k in the wave function, the allowed wave functions $\psi_n(x)$ are given by

$$\psi_n(x) = A \sin\left(\frac{n\pi x}{L}\right) \qquad \textbf{(41.18)}$$

which is the wave function (Eq. 41.12) used in our initial discussion of the particle in a box.

41.4 A Particle in a Well of Finite Height

Now consider a particle in a *finite* potential well, that is, a system having a potential energy that is zero when the particle is in the region $0 < x < L$ and a finite value U when the particle is outside this region as in Figure 41.6. Classically, if the total energy E of the system is less than U, the particle is permanently bound in the potential well. If the particle were outside the well, its kinetic energy would have to be negative, which is an impossibility. According to quantum mechanics, however, a finite probability exists that the particle can be found outside the well even if $E < U$. That is, the wave function ψ is generally nonzero outside the well—regions I and III in Figure 41.6—so the probability density $|\psi|^2$ is also nonzero in these regions. Although this notion may be uncomfortable to accept, the uncertainty principle indicates that the energy of the system is uncertain. This uncertainty allows the particle to be outside the well as long as the apparent violation of conservation of energy does not exist in any measurable way.

In region II, where $U = 0$, the allowed wave functions are again sinusoidal because they represent solutions of Equation 41.16. The boundary conditions, however,

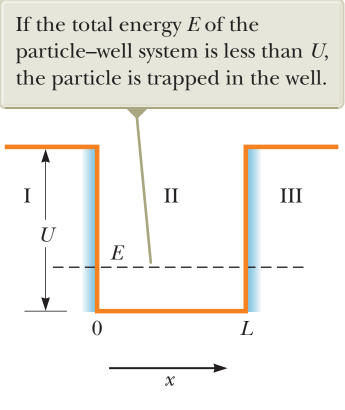


Figure 41.6 Potential-energy diagram of a well of finite height U and length L.

no longer require that ψ be zero at the ends of the well, as was the case with the infinite square well.

The Schrödinger equation for regions I and III may be written

$$\frac{d^2\psi}{dx^2} = \frac{2m(U - E)}{\hbar^2}\psi \tag{41.19}$$

Because $U > E$, the coefficient of ψ on the right-hand side is necessarily positive. Therefore, we can express Equation 41.19 as

$$\frac{d^2\psi}{dx^2} = C^2\psi \tag{41.20}$$

where $C^2 = 2m(U - E)/\hbar^2$ is a positive constant in regions I and III. As you can verify by substitution, the general solution of Equation 41.20 is

$$\psi = Ae^{Cx} + Be^{-Cx} \tag{41.21}$$

where A and B are constants.

We can use this general solution as a starting point for determining the appropriate solution for regions I and III. The solution must remain finite as $x \to \pm\infty$. Therefore, in region I, where $x < 0$, the function ψ cannot contain the term Be^{-Cx}. This requirement is handled by taking $B = 0$ in this region to avoid an infinite value for ψ for large negative values of x. Likewise, in region III, where $x > L$, the function ψ cannot contain the term Ae^{Cx}. This requirement is handled by taking $A = 0$ in this region to avoid an infinite value for ψ for large positive x values. Hence, the solutions in regions I and III are

$$\psi_{\mathrm{I}} = Ae^{Cx} \quad \text{for } x < 0$$

$$\psi_{\mathrm{III}} = Be^{-Cx} \quad \text{for } x > L$$

In region II, the wave function is sinusoidal and has the general form

$$\psi_{\mathrm{II}}(x) = F\sin kx + G\cos kx$$

where F and G are constants.

These results show that the wave functions outside the potential well (where classical physics forbids the presence of the particle) decay exponentially with distance. At large negative x values, ψ_{I} approaches zero; at large positive x values, ψ_{III} approaches zero. These functions, together with the sinusoidal solution in region II, are shown in Figure 41.7a for the first three energy states. In evaluating the complete wave function, we impose the following boundary conditions:

$$\psi_{\mathrm{I}} = \psi_{\mathrm{II}} \quad \text{and} \quad \frac{d\psi_{\mathrm{I}}}{dx} = \frac{d\psi_{\mathrm{II}}}{dx} \quad \text{at } x = 0$$

$$\psi_{\mathrm{II}} = \psi_{\mathrm{III}} \quad \text{and} \quad \frac{d\psi_{\mathrm{II}}}{dx} = \frac{d\psi_{\mathrm{III}}}{dx} \quad \text{at } x = L$$

These four boundary conditions and the normalization condition (Eq. 41.7) are sufficient to determine the four constants A, B, F, and G and the allowed values of the energy E. Figure 41.7b plots the probability densities for these states. In each case, the wave functions inside and outside the potential well join smoothly at the boundaries.

The notion of trapping particles in potential wells is used in the burgeoning field of **nanotechnology,** which refers to the design and application of devices having dimensions ranging from 1 to 100 nm. The fabrication of these devices often involves manipulating single atoms or small groups of atoms to form very tiny structures or mechanisms.

One area of nanotechnology of interest to researchers is the **quantum dot,** a small region that is grown in a silicon crystal and acts as a potential well. This region can trap electrons into states with quantized energies. The wave functions

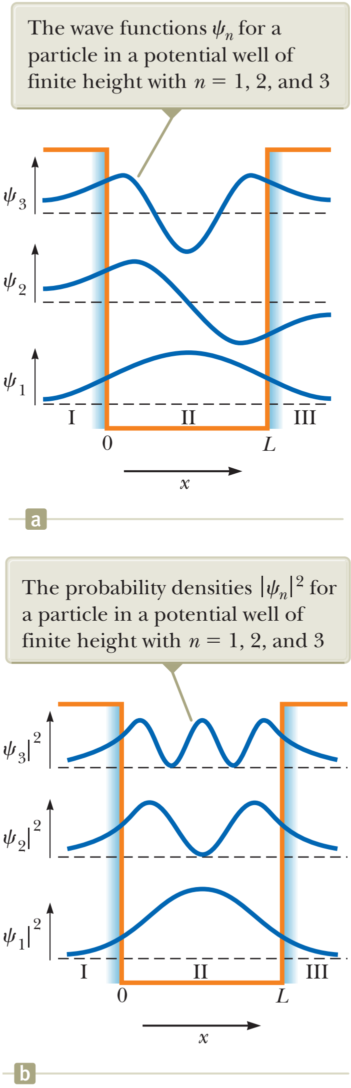


Figure 41.7 The first three allowed states for a particle in a potential well of finite height. The states are shown superimposed on the potential energy function of Figure 41.6. The wave functions and probability densities are plotted vertically from separate axes that are offset vertically for clarity. The positions of these axes on the potential energy function suggest the relative energies of the states.

for a particle in a quantum dot look similar to those in Figure 41.7a if L is on the order of nanometers. The storage of binary information using quantum dots is an active field of research. A simple binary scheme would involve associating a one with a quantum dot containing an electron and a zero with an empty dot. Other schemes involve cells of multiple dots such that arrangements of electrons among the dots correspond to ones and zeroes. Several research laboratories are studying the properties and potential applications of quantum dots. Information should be forthcoming from these laboratories at a steady rate in the next few years.

41.5 Tunneling Through a Potential Energy Barrier

Consider the potential energy function shown in Figure 41.8. In this situation, the potential energy has a constant value of U in the region of width L and is zero in all other regions.[10] A potential energy function of this shape is called a **square barrier,** and U is called the **barrier height.** A very interesting and peculiar phenomenon occurs when a moving particle encounters such a barrier of finite height and width. Suppose a particle of energy $E < U$ is incident on the barrier from the left (Fig. 41.8). Classically, the particle is reflected by the barrier. If the particle were located in region II, its kinetic energy would be negative, which is not classically allowed. Consequently, region II and therefore region III are both classically *forbidden* to the particle incident from the left. According to quantum mechanics, however, all regions are accessible to the particle, regardless of its energy. (Although all regions are accessible, the probability of the particle being in a classically forbidden region is very low.) According to the uncertainty principle, the particle could be within the barrier as long as the time interval during which it is in the barrier is short and consistent with Equation 40.26. If the barrier is relatively narrow, this short time interval can allow the particle to pass through the barrier.

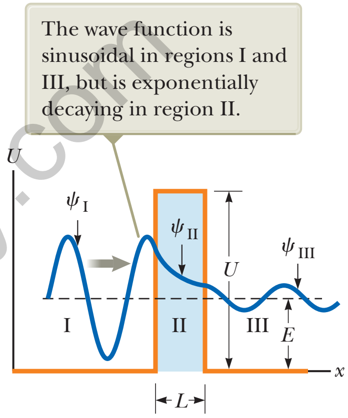


Figure 41.8 Wave function ψ for a particle incident from the left on a barrier of height U and width L. The wave function is plotted vertically from an axis positioned at the energy of the particle.

Pitfall Prevention 41.4

"Height" on an Energy Diagram The word *height* (as in *barrier height*) refers to an energy in discussions of barriers in potential-energy diagrams. For example, we might say the height of the barrier is 10 eV. On the other hand, the barrier *width* refers to the traditional usage of such a word and is an actual physical length measurement between the locations of the two vertical sides of the barrier.

Let's approach this situation using a mathematical representation. The Schrödinger equation has valid solutions in all three regions. The solutions in regions I and III are sinusoidal like Equation 41.18. In region II, the solution is exponential like Equation 41.21. Applying the boundary conditions that the wave functions in the three regions and their derivatives must join smoothly at the boundaries, a full solution, such as the one represented by the curve in Figure 41.8, can be found. Because the probability of locating the particle is proportional to $|\psi|^2$, the probability of finding the particle beyond the barrier in region III is nonzero. This result is in complete disagreement with classical physics. The movement of the particle to the far side of the barrier is called **tunneling** or **barrier penetration.**

The probability of tunneling can be described with a **transmission coefficient** T and a **reflection coefficient** R. The transmission coefficient represents the probability that the particle penetrates to the other side of the barrier, and the reflection coefficient is the probability that the particle is reflected by the barrier. Because the incident particle is either reflected or transmitted, we require that $T + R = 1$. An approximate expression for the transmission coefficient that is obtained in the case of $T << 1$ (a very wide barrier or a very high barrier, that is, $U >> E$) is

$$T \approx e^{-2CL} \tag{41.22}$$

where

$$C = \frac{\sqrt{2m(U - E)}}{\hbar} \tag{41.23}$$

This quantum model of barrier penetration and specifically Equation 41.22 show that T can be nonzero. That the phenomenon of tunneling is observed experimentally provides further confidence in the principles of quantum physics.

[10] It is common in physics to refer to L as the *length* of a well but the *width* of a barrier.

Quick Quiz 41.4 Which of the following changes would increase the probability of transmission of a particle through a potential barrier? (You may choose more than one answer.) **(a)** decreasing the width of the barrier **(b)** increasing the width of the barrier **(c)** decreasing the height of the barrier **(d)** increasing the height of the barrier **(e)** decreasing the kinetic energy of the incident particle **(f)** increasing the kinetic energy of the incident particle

Example 41.4 Transmission Coefficient for an Electron

A 30-eV electron is incident on a square barrier of height 40 eV.

(A) What is the probability that the electron tunnels through the barrier if its width is 1.0 nm?

SOLUTION

Conceptualize Because the particle energy is smaller than the height of the potential barrier, we expect the electron to reflect from the barrier with a probability of 100% according to classical physics. Because of the tunneling phenomenon, however, there is a finite probability that the particle can appear on the other side of the barrier.

Categorize We evaluate the probability using an equation developed in this section, so we categorize this example as a substitution problem.

Evaluate the quantity $U - E$ that appears in Equation 41.23:

$$U - E = 40\text{ eV} - 30\text{ eV} = 10\text{ eV}\left(\frac{1.6 \times 10^{-19}\text{ J}}{1\text{ eV}}\right) = 1.6 \times 10^{-18}\text{ J}$$

Evaluate the quantity $2CL$ using Equation 41.23:

$$(1) \quad 2CL = 2\frac{\sqrt{2(9.11 \times 10^{-31}\text{ kg})(1.6 \times 10^{-18}\text{ J})}}{1.055 \times 10^{-34}\text{ J}\cdot\text{s}}(1.0 \times 10^{-9}\text{ m}) = 32.4$$

From Equation 41.22, find the probability of tunneling through the barrier:

$$T \approx e^{-2CL} = e^{-32.4} = \boxed{8.5 \times 10^{-15}}$$

(B) What is the probability that the electron tunnels through the barrier if its width is 0.10 nm?

SOLUTION

In this case, the width L in Equation (1) is one-tenth as large, so evaluate the new value of $2CL$:

$$2CL = (0.1)(32.4) = 3.24$$

From Equation 41.22, find the new probability of tunneling through the barrier:

$$T \approx e^{-2CL} = e^{-3.24} = \boxed{0.039}$$

In part (A), the electron has approximately 1 chance in 10^{14} of tunneling through the barrier. In part (B), however, the electron has a much higher probability (3.9%) of penetrating the barrier. Therefore, reducing the width of the barrier by only one order of magnitude increases the probability of tunneling by about 12 orders of magnitude!

41.6 Applications of Tunneling

As we have seen, tunneling is a quantum phenomenon, a manifestation of the wave nature of matter. Many examples exist (on the atomic and nuclear scales) for which tunneling is very important.

Alpha Decay

One form of radioactive decay is the emission of alpha particles (the nuclei of helium atoms) by unstable, heavy nuclei (Chapter 44). To escape from the nucleus, an alpha particle must penetrate a barrier whose height is several times larger than the energy of the nucleus–alpha particle system as shown in Figure 41.9. The barrier results from a combination of the attractive nuclear force (discussed in Chapter 44) and the Coulomb repulsion (discussed in Chapter 23) between the alpha particle and the rest of the nucleus. Occasionally, an alpha particle tunnels through the barrier, which explains the basic mechanism for this type of decay and the large variations in the mean lifetimes of various radioactive nuclei.

Figure 41.8 shows the wave function of a particle tunneling through a barrier in one dimension. A similar wave function having spherical symmetry describes the barrier penetration of an alpha particle leaving a radioactive nucleus. The wave function exists both inside and outside the nucleus, and its amplitude is constant in time. In this way, the wave function correctly describes the small but constant probability that the nucleus will decay. The moment of decay cannot be predicted. In general, quantum mechanics implies that the future is indeterminate. This feature is in contrast to classical mechanics, from which the trajectory of an object can be calculated to arbitrarily high precision from precise knowledge of its initial position and velocity and of the forces exerted on it. Do not think that the future is undetermined simply because we have incomplete information about the present. The wave function contains all the information about the state of a system. Sometimes precise predictions can be made, such as the energy of a bound system, but sometimes only probabilities can be calculated about the future. The fundamental laws of nature are probabilistic. Therefore, it appears that Einstein's famous statement about quantum mechanics, "God does not roll dice," was wrong.

A radiation detector can be used to show that a nucleus decays by emitting a particle at a particular moment and in a particular direction. To point out the contrast between this experimental result and the wave function describing it, Schrödinger imagined a box containing a cat, a radioactive sample, a radiation counter, and a vial of poison. When a nucleus in the sample decays, the counter triggers the administration of lethal poison to the cat. Quantum mechanics correctly predicts the probability of finding the cat dead when the box is opened. Before the box is opened, does the cat have a wave function describing it as fractionally dead, with some chance of being alive?

This question is under continuing investigation, never with actual cats but sometimes with interference experiments building upon the experiment described in Section 40.7. Does the act of measurement change the system from a probabilistic to a definite state? When a particle emitted by a radioactive nucleus is detected at one particular location, does the wave function describing the particle drop instantaneously to zero everywhere else in the Universe? (Einstein called such a state change a "spooky action at a distance.") Is there a fundamental difference between a quantum system and a macroscopic system? The answers to these questions are unknown.

Nuclear Fusion

The basic reaction that powers the Sun and, indirectly, almost everything else in the solar system is fusion, which we shall study in Chapter 45. In one step of the process that occurs at the core of the Sun, protons must approach one another to within such a small distance that they fuse and form a deuterium nucleus. (See Section 45.4.) According to classical physics, these protons cannot overcome and penetrate the barrier caused by their mutual electrical repulsion. Quantum mechanically, however, the protons are able to tunnel through the barrier and fuse together.

Scanning Tunneling Microscopes

The scanning tunneling microscope (STM) enables scientists to obtain highly detailed images of surfaces at resolutions comparable to the size of a *single atom*. Figure 41.10 (page 1284), showing the surface of a piece of graphite, demonstrates what STMs can do. What makes this image so remarkable is that its resolution is

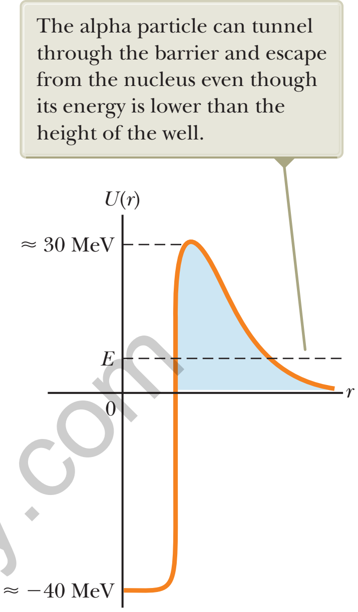


Figure 41.9 The potential well for an alpha particle in a nucleus.

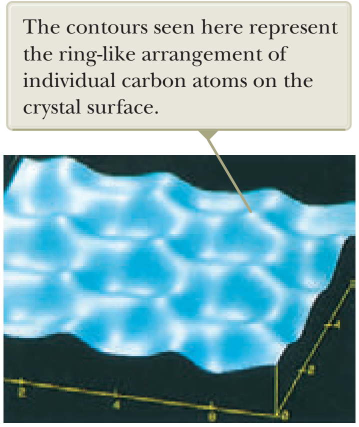


Photo courtesy of Paul K. Hansma, University of California, Santa Barbara

Figure 41.10 The surface of graphite as "viewed" with a scanning tunneling microscope. This type of microscope enables scientists to see details with a lateral resolution of about 0.2 nm and a vertical resolution of 0.001 nm.

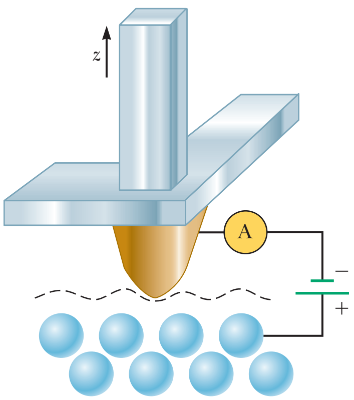


Figure 41.11 Schematic view of a scanning tunneling microscope. A scan of the tip over the sample can reveal surface contours down to the atomic level. An STM image is composed of a series of scans displaced laterally from one another. (Based on a drawing from P. K. Hansma, V. B. Elings, O. Marti, and C. Bracker, *Science* **242:**209, 1988. © 1988 by the AAAS.)

approximately 0.2 nm. For an optical microscope, the resolution is limited by the wavelength of the light used to make the image. Therefore, an optical microscope has a resolution no better than 200 nm, about half the wavelength of visible light, and so could never show the detail displayed in Figure 41.10.

Scanning tunneling microscopes achieve such high resolution by using the basic idea shown in Figure 41.11. An electrically conducting probe with a very sharp tip is brought near the surface to be studied. The empty space between tip and surface represents the "barrier" we have been discussing, and the tip and surface are the two walls of the "potential well." Because electrons obey quantum rules rather than Newtonian rules, they can "tunnel" across the barrier of empty space. If a voltage is applied between surface and tip, electrons in the atoms of the surface material can tunnel preferentially from surface to tip to produce a tunneling current. In this way, the tip samples the distribution of electrons immediately above the surface.

In the empty space between tip and surface, the electron wave function falls off exponentially (see region II in Fig. 41.8 and Example 41.4). For tip-to-surface distances $z > 1$ nm (that is, beyond a few atomic diameters), essentially no tunneling takes place. This exponential behavior causes the current of electrons tunneling from surface to tip to depend very strongly on z. By monitoring the tunneling current as the tip is scanned over the surface, scientists obtain a sensitive measure of the topography of the electron distribution on the surface. The result of this scan is used to make images like that in Figure 41.10. In this way, the STM can measure the height of surface features to within 0.001 nm, approximately 1/100 of an atomic diameter!

You can appreciate the sensitivity of STMs by examining Figure 41.10. Of the six carbon atoms in each ring, three appear lower than the other three. In fact, all six atoms are at the same height, but all have slightly different electron distributions. The three atoms that appear lower are bonded to other carbon atoms directly beneath them in the underlying atomic layer; as a result, their electron distributions, which are responsible for the bonding, extend downward beneath the surface. The atoms in the surface layer that appear higher do not lie directly over subsurface atoms and hence are not bonded to any underlying atoms. For these higher-appearing atoms, the electron distribution extends upward into the space above the surface. Because STMs map the topography of the electron distribution, this extra electron density makes these atoms appear higher in Figure 41.10.

The STM has one serious limitation: Its operation depends on the electrical conductivity of the sample and the tip. Unfortunately, most materials are not electrically conductive at their surfaces. Even metals, which are usually excellent electrical conductors, are covered with nonconductive oxides. A newer microscope, the atomic force microscope, or AFM, overcomes this limitation.

Resonant Tunneling Devices

Let's expand on the quantum-dot discussion in Section 41.4 by exploring the **resonant tunneling device.** Figure 41.12a shows the physical construction of such a device. The island of gallium arsenide in the center is a quantum dot located between two barriers formed from the thin extensions of aluminum arsenide. Figure 41.12b shows both the potential barriers encountered by electrons incident from the left and the quantized energy levels in the quantum dot. This situation differs from the one shown in Figure 41.8 in that there are quantized energy levels on the right of the first barrier. In Figure 41.8, an electron that tunnels through the barrier is considered a free particle and can have any energy. In contrast, the second barrier in Figure 41.12b imposes boundary conditions on the particle and quantizes its energy in the quantum dot. In Figure 41.12b, as the electron with the energy shown encounters the first barrier, it has no matching energy levels available on the right side of the barrier, which greatly reduces the probability of tunneling.

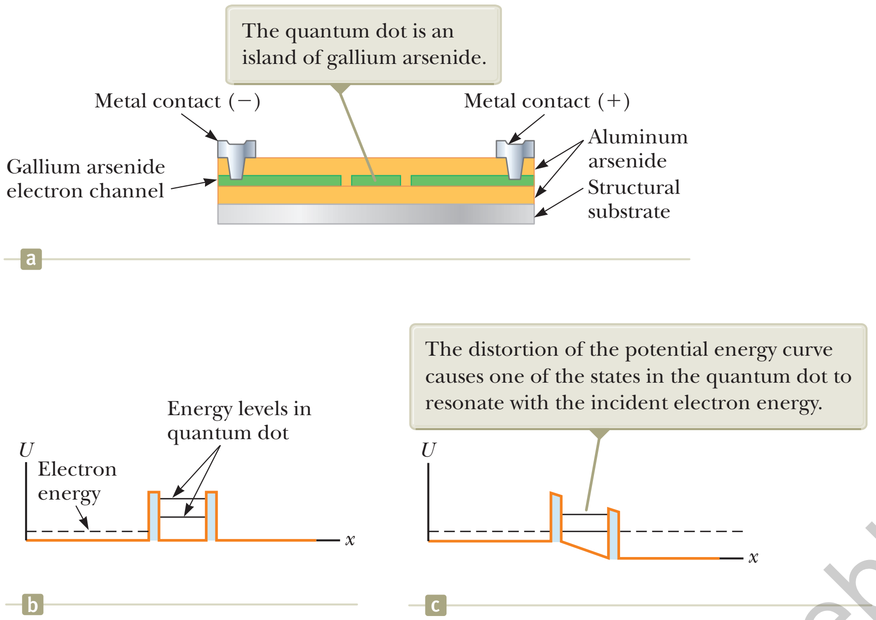


Figure 41.12 (a) The physical structure of a resonant tunneling device. (b) A potential-energy diagram showing the double barrier representing the walls of the quantum dot. (c) A voltage is applied across the device.

Figure 41.12c shows the effect of applying a voltage: the potential decreases with position as we move to the right across the device. The deformation of the potential barrier results in an energy level in the quantum dot coinciding with the energy of the incident electrons. This "resonance" of energies gives the device its name. When the voltage is applied, the probability of tunneling increases tremendously and the device carries current. In this manner, the device can be used as a very fast switch on a nanotechnological scale.

Resonant Tunneling Transistors

Figure 41.13a shows the addition of a gate electrode at the top of the resonant tunneling device over the quantum dot. This electrode turns the device into a **resonant**

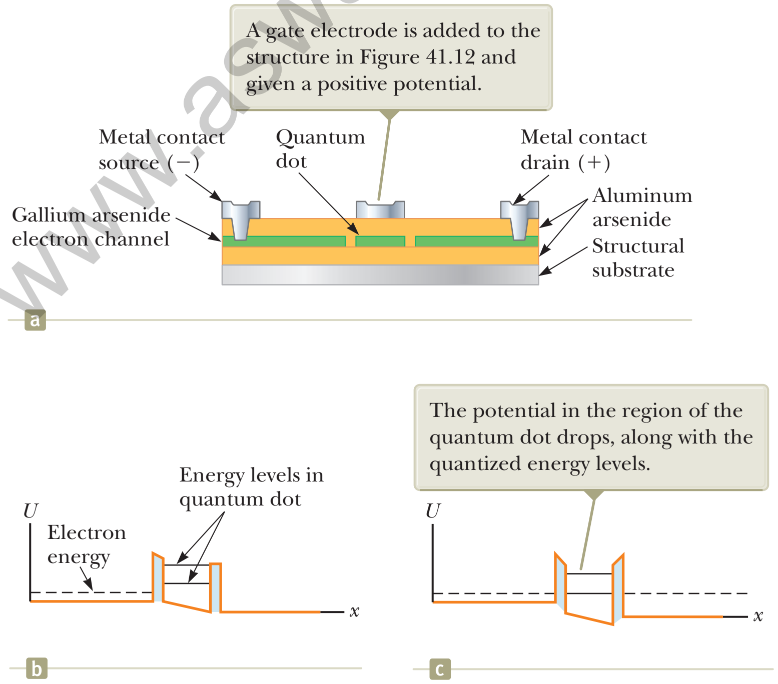


Figure 41.13 (a) A resonant tunneling transistor. (b) A potential-energy diagram showing the double barrier representing the walls of the quantum dot. (c) A voltage is applied to the gate electrode.

tunneling transistor. The basic function of a transistor is amplification, converting a small varying voltage into a large varying voltage. Figure 41.13b, representing the potential-energy diagram for the tunneling transistor, has a slope at the bottom of the quantum dot due to the differing voltages at the source and drain electrodes. In this configuration, there is no resonance between the electron energies outside the quantum dot and the quantized energies within the dot. By applying a small voltage to the gate electrode as in Figure 41.13c, the quantized energies can be brought into resonance with the electron energy outside the well and resonant tunneling occurs. The resulting current causes a voltage across an external resistor that is much larger than that of the gate voltage; hence, the device amplifies the input signal to the gate electrode.

41.7 The Simple Harmonic Oscillator

Consider a particle that is subject to a linear restoring force $F = -kx$, where k is a constant and x is the position of the particle relative to equilibrium ($x = 0$). The classical description of such a situation is provided by the particle in simple harmonic motion analysis model, which was discussed in Chapter 15. The potential energy of the system is, from Equation 15.20,

$$U = \tfrac{1}{2}kx^2 = \tfrac{1}{2}m\omega^2 x^2$$

where the angular frequency of vibration is $\omega = \sqrt{k/m}$. Classically, if the particle is displaced from its equilibrium position and released, it oscillates between the points $x = -A$ and $x = A$, where A is the amplitude of the motion. Furthermore, its total energy E is, from Equation 15.21,

$$E = K + U = \tfrac{1}{2}kA^2 = \tfrac{1}{2}m\omega^2 A^2$$

In the classical model, any value of E is allowed, including $E = 0$, which is the total energy when the particle is at rest at $x = 0$.

Let's investigate how the simple harmonic oscillator is treated from a quantum point of view. The Schrödinger equation for this problem is obtained by substituting $U = \frac{1}{2}m\omega^2x^2$ into Equation 41.15:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \tfrac{1}{2}m\omega^2x^2\psi = E\psi \qquad \textbf{(41.24)}$$

The mathematical technique for solving this equation is beyond the level of this book; nonetheless, it is instructive to guess at a solution. We take as our guess the following wave function:

$$\psi = Be^{-Cx^2} \qquad \textbf{(41.25)}$$

Substituting this function into Equation 41.24 shows that it is a satisfactory solution to the Schrödinger equation, provided that

$$C = \frac{m\omega}{2\hbar} \quad \text{and} \quad E = \tfrac{1}{2}\hbar\omega$$

It turns out that the solution we have guessed corresponds to the ground state of the system, which has an energy $\frac{1}{2}\hbar\omega$. Because $C = m\omega/2\hbar$, it follows from Equation 41.25 that the wave function for this state is

Wave function for the ground state of a simple harmonic oscillator ▶

$$\psi = Be^{-(m\omega/2\hbar)x^2} \qquad \textbf{(41.26)}$$

where B is a constant to be determined from the normalization condition. This result is but one solution to Equation 41.24. The remaining solutions that describe the excited states are more complicated, but all solutions include the exponential factor e^{-Cx^2}.

The energy levels of a harmonic oscillator are quantized as we would expect because the oscillating particle is bound to stay near $x = 0$. The energy of a state having an arbitrary quantum number n is

$$E_n = \left(n + \tfrac{1}{2}\right)\hbar\omega \qquad n = 0, 1, 2, \ldots \tag{41.27}$$

The state $n = 0$ corresponds to the ground state, whose energy is $E_0 = \frac{1}{2}\hbar\omega$; the state $n = 1$ corresponds to the first excited state, whose energy is $E_1 = \frac{3}{2}\hbar\omega$; and so on. The energy-level diagram for this system is shown in Figure 41.14. The separations between adjacent levels are equal and given by

$$\Delta E = \hbar\omega \tag{41.28}$$

Notice that the energy levels for the harmonic oscillator in Figure 41.14 are equally spaced, just as Planck proposed for the oscillators in the walls of the cavity that was used in the model for blackbody radiation in Section 40.1. Planck's Equation 40.4 for the energy levels of the oscillators differs from Equation 41.27 only in the term $\frac{1}{2}$ added to n. This additional term does not affect the energy emitted in a transition, given by Equation 40.5, which is equivalent to Equation 41.28. That Planck generated these concepts without the benefit of the Schrödinger equation is testimony to his genius.

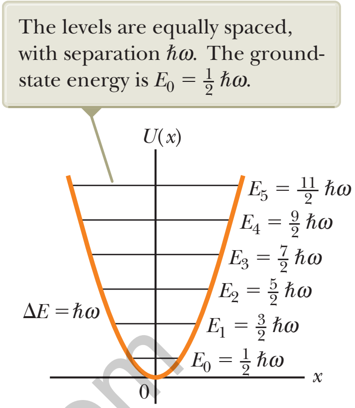


Figure 41.14 Energy-level diagram for a simple harmonic oscillator, superimposed on the potential energy function.

Example 41.5 Molar Specific Heat of Hydrogen Gas

In Figure 21.6 (Section 21.3), which shows the molar specific heat of hydrogen as a function of temperature, vibration does not contribute to the molar specific heat at room temperature. Explain why, modeling the hydrogen molecule as a simple harmonic oscillator. The effective spring constant for the bond in the hydrogen molecule is 573 N/m.

SOLUTION

Conceptualize Imagine the only mode of vibration available to a diatomic molecule. This mode (shown in Fig. 21.5c) consists of the two atoms always moving in opposite directions with equal speeds.

Categorize We categorize this example as a quantum harmonic oscillator problem, with the molecule modeled as a two-particle system.

Analyze The motion of the particles relative to the center of mass can be analyzed by considering the oscillation of a single particle with reduced mass μ. (See Problem 40.)

Use the result of Problem 40 to evaluate the reduced mass of the hydrogen molecule, in which the masses of the two particles are the same:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} = \frac{m^2}{2m} = \tfrac{1}{2}m$$

Using Equation 41.28, calculate the energy necessary to excite the molecule from its ground vibrational state to its first excited vibrational state:

$$\Delta E = \hbar\omega = \hbar\sqrt{\frac{k}{\mu}} = \hbar\sqrt{\frac{k}{\frac{1}{2}m}} = \hbar\sqrt{\frac{2k}{m}}$$

Substitute numerical values, noting that m is the mass of a hydrogen atom:

$$\Delta E = (1.055 \times 10^{-34}\ \text{J}\cdot\text{s})\sqrt{\frac{2(573\ \text{N/m})}{1.67 \times 10^{-27}\ \text{kg}}} = 8.74 \times 10^{-20}\ \text{J}$$

Set this energy equal to $\frac{3}{2}k_BT$ from Equation 21.19 and find the temperature at which the average molecular translational kinetic energy is equal to that required to excite the first vibrational state of the molecule:

$$\tfrac{3}{2}k_BT = \Delta E$$

$$T = \tfrac{2}{3}\left(\frac{\Delta E}{k_B}\right) = \tfrac{2}{3}\left(\frac{8.74 \times 10^{-20}\ \text{J}}{1.38 \times 10^{-23}\ \text{J/K}}\right) = 4.22 \times 10^3\ \text{K}$$

Finalize The temperature of the gas must be more than 4 000 K for the translational kinetic energy to be comparable to the energy required to excite the first vibrational state. This excitation energy must come from collisions between

continued

41.5 continued

molecules, so if the molecules do not have sufficient translational kinetic energy, they cannot be excited to the first vibrational state and vibration does not contribute to the molar specific heat. Hence, the curve in Figure 21.6 does not rise to a value corresponding to the contribution of vibration until the hydrogen gas has been raised to thousands of kelvins.

Figure 21.6 shows that rotational energy levels must be more closely spaced in energy than vibrational levels because they are excited at a lower temperature than the vibrational levels. The translational energy levels are those of a particle in a three-dimensional box, where the box is the container holding the gas. These levels are given by an expression similar to Equation 41.14. Because the box is macroscopic in size, L is very large and the energy levels are very close together. In fact, they are so close together that translational energy levels are excited at the temperature at which liquid hydrogen becomes a gas shown in Figure 21.6.

Summary

Definitions

The **wave function** Ψ for a system is a mathematical function that can be written as a product of a space function ψ for one particle of the system and a complex time function:

$$\Psi(\vec{\mathbf{r}}_1, \vec{\mathbf{r}}_2, \vec{\mathbf{r}}_3, \ldots, \vec{\mathbf{r}}_j, \ldots, t) = \psi(\vec{\mathbf{r}}_j)e^{-i\omega t} \quad \textbf{(41.2)}$$

where $\omega\ (= 2\pi f)$ is the angular frequency of the wave function and $i = \sqrt{-1}$. The wave function contains within it all the information that can be known about the particle.

The measured position x of a particle, averaged over many trials, is called the **expectation value** of x and is defined by

$$\langle x \rangle \equiv \int_{-\infty}^{\infty} \psi^* x \psi \, dx \quad \textbf{(41.8)}$$

Concepts and Principles

In quantum mechanics, a particle in a system can be represented by a wave function $\psi(x, y, z)$. The probability per unit volume (or probability density) that a particle will be found at a point is $|\psi|^2 = \psi^*\psi$, where ψ^* is the complex conjugate of ψ. If the particle is confined to moving along the x axis, the probability that it is located in an interval dx is $|\psi|^2\, dx$. Furthermore, the sum of all these probabilities over all values of x must be 1:

$$\int_{-\infty}^{\infty} |\psi|^2 \, dx = 1 \quad \textbf{(41.7)}$$

This expression is called the **normalization condition.**

If a particle of mass m is confined to moving in a one-dimensional box of length L whose walls are impenetrable, then ψ must be zero at the walls and outside the box. The wave functions for this system are given by

$$\psi(x) = A \sin\left(\frac{n\pi x}{L}\right) \quad n = 1, 2, 3, \ldots \quad \textbf{(41.12)}$$

where A is the maximum value of ψ. The allowed states of a particle in a box have quantized energies given by

$$E_n = \left(\frac{h^2}{8mL^2}\right)n^2 \quad n = 1, 2, 3, \ldots \quad \textbf{(41.14)}$$

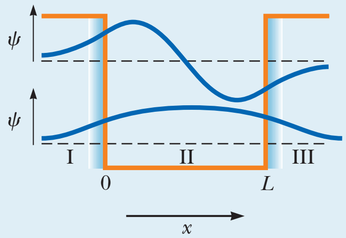
The wave function for a system must satisfy the **Schrödinger equation.** The time-independent Schrödinger equation for a particle confined to moving along the x axis is

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + U\psi = E\psi \quad \textbf{(41.15)}$$

where U is the potential energy of the system and E is the total energy.

Analysis Models for Problem Solving

Quantum Particle Under Boundary Conditions. An interaction of a quantum particle with its environment represents one or more boundary conditions. If the interaction restricts the particle to a finite region of space, the energy of the system is quantized. All wave functions must satisfy the following four boundary conditions: (1) $\psi(x)$ must remain finite as x approaches 0, (2) $\psi(x)$ must approach zero as x approaches $\pm\infty$, (3) $\psi(x)$ must be continuous for all values of x, and (4) $d\psi/dx$ must be continuous for all finite values of $U(x)$. If the solution to Equation 41.15 is piecewise, conditions (3) and (4) must be applied at the boundaries between regions of x in which Equation 41.15 has been solved.



Objective Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. A beam of quantum particles with kinetic energy 2.00 eV is reflected from a potential barrier of small width and original height 3.00 eV. How does the fraction of the particles that are reflected change as the barrier height is reduced to 2.01 eV? (a) It increases. (b) It decreases. (c) It stays constant at zero. (d) It stays constant at 1. (e) It stays constant with some other value.

2. A quantum particle of mass m_1 is in a square well with infinitely high walls and length 3 nm. Rank the situations (a) through (e) according to the particle's energy from highest to lowest, noting any cases of equality. (a) The particle of mass m_1 is in the ground state of the well. (b) The same particle is in the $n = 2$ excited state of the same well. (c) A particle with mass $2m_1$ is in the ground state of the same well. (d) A particle of mass m_1 in the ground state of the same well, and the uncertainty principle has become inoperative; that is, Planck's constant has been reduced to zero. (e) A particle of mass m_1 is in the ground state of a well of length 6 nm.

3. Is each one of the following statements (a) through (e) true or false for an electron? (a) It is a quantum particle, behaving in some experiments like a classical particle and in some experiments like a classical wave. (b) Its rest energy is zero. (c) It carries energy in its motion. (d) It carries momentum in its motion. (e) Its motion is described by a wave function that has a wavelength and satisfies a wave equation.

4. Is each one of the following statements (a) through (e) true or false for a photon? (a) It is a quantum particle, behaving in some experiments like a classical particle and in some experiments like a classical wave. (b) Its rest energy is zero. (c) It carries energy in its motion. (d) It carries momentum in its motion. (e) Its motion is described by a wave function that has a wavelength and satisfies a wave equation.

5. A particle in a rigid box of length L is in the first excited state for which $n = 2$ (Fig. OQ41.5). Where is the particle most likely to be found? (a) At the center of the box. (b) At either end of the box. (c) All points in the box are equally likely. (d) One-fourth of the way from either end of the box. (e) None of those answers is correct.

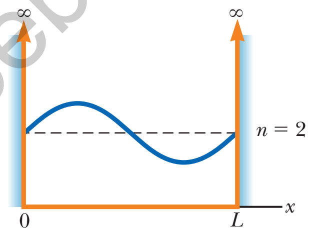


Figure OQ41.5

6. Two square wells have the same length. Well 1 has walls of finite height, and well 2 has walls of infinite height. Both wells contain identical quantum particles, one in each well. **(i)** Is the wavelength of the ground-state wave function (a) greater for well 1, (b) greater for well 2, or (c) equal for both wells? **(ii)** Is the magnitude of the ground-state momentum (a) greater for well 1, (b) greater for well 2, or (c) equal for both wells? **(iii)** Is the ground-state energy of the particle (a) greater for well 1, (b) greater for well 2, or (c) equal for both wells?

7. The probability of finding a certain quantum particle in the section of the x axis between $x = 4$ nm and $x = 7$ nm is 48%. The particle's wave function $\psi(x)$ is constant over this range. What numerical value can be attributed to $\psi(x)$, in units of $\text{nm}^{-1/2}$? (a) 0.48 (b) 0.16 (c) 0.12 (d) 0.69 (e) 0.40

8. Suppose a tunneling current in an electronic device goes through a potential-energy barrier. The tunneling current is small because the width of the barrier is large and the barrier is high. To increase the current most effectively, what should you do? (a) Reduce the width of the barrier. (b) Reduce the height of the barrier. (c) Either choice (a) or choice (b) is equally effective. (d) Neither choice (a) nor choice (b) increases the current.

9. Unlike the idealized diagram of Figure 41.11, a typical tip used for a scanning tunneling microscope is rather jagged on the atomic scale, with several irregularly spaced points. For such a tip, does most of the tunneling current occur between the sample and (a) all the points of the tip equally, (b) the most centrally located point, (c) the point closest to the sample, or (d) the point farthest from the sample?

10. Figure OQ41.10 represents the wave function for a hypothetical quantum particle in a given region. From the choices *a* through *e*, at what value of *x* is the particle most likely to be found?

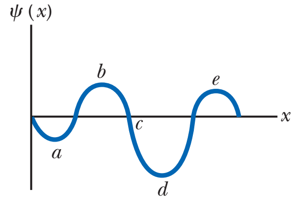


Figure OQ41.10

Conceptual Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. Richard Feynman said, "A philosopher once said that 'it is necessary for the very existence of science that the same conditions always produce the same results.' Well, they don't!" In view of what has been discussed in this chapter, present an argument showing that the philosopher's statement is false. How might the statement be reworded to make it true?

2. Discuss the relationship between ground-state energy and the uncertainty principle.

3. For a quantum particle in a box, the probability density at certain points is zero as seen in Figure CQ41.3. Does this value imply that the particle cannot move across these points? Explain.

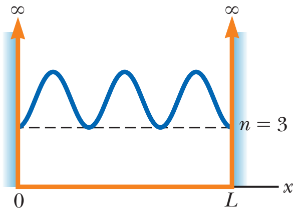


Figure CQ41.3

4. Why are the following wave functions not physically possible for all values of *x*? (a) $\psi(x) = Ae^x$ (b) $\psi(x) = A\tan x$

5. What is the significance of the wave function ψ?

6. In quantum mechanics, it is possible for the energy E of a particle to be less than the potential energy, but classically this condition is not possible. Explain.

7. Consider the wave functions in Figure CQ41.7. Which of them are not physically significant in the interval shown? For those that are not, state why they fail to qualify.

8. How is the Schrödinger equation useful in describing quantum phenomena?

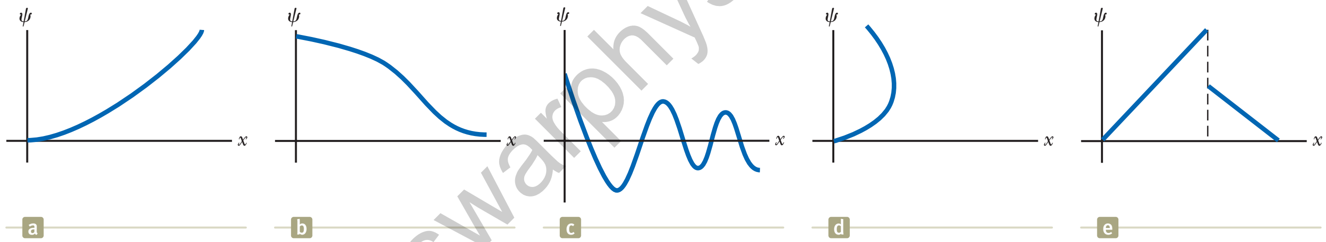


Figure CQ41.7

Problems

ENHANCED WebAssign The problems found in this chapter may be assigned online in Enhanced WebAssign

1. straightforward; **2.** intermediate; **3.** challenging

1. full solution available in the *Student Solutions Manual/Study Guide*

AMT Analysis Model tutorial available in Enhanced WebAssign

GP Guided Problem

M Master It tutorial available in Enhanced WebAssign

W Watch It video solution available in Enhanced WebAssign

Section 41.1 The Wave Function

1. M A free electron has a wave function

$$\psi(x) = Ae^{i(5.00 \times 10^{10}\, x)}$$

where x is in meters. Find its (a) de Broglie wavelength, (b) momentum, and (c) kinetic energy in electron volts.

2. The wave function for a particle is given by $\psi(x) = Ae^{-|x|/a}$, where A and a are constants. (a) Sketch this function for values of x in the interval $-3a < x < 3a$. (b) Determine the value of A. (c) Find the probability that the particle will be found in the interval $-a < x < a$.

3. The wave function for a quantum particle is given by $\psi(x) = Ax$ between $x = 0$ and $x = 1.00$, and $\psi(x) = 0$ elsewhere. Find (a) the value of the normalization constant A, (b) the probability that the particle will be found between $x = 0.300$ and $x = 0.400$, and (c) the expectation value of the particle's position.

4. The wave function for a quantum particle is

$$\psi(x) = \sqrt{\frac{a}{\pi(x^2 + a^2)}}$$

for $a > 0$ and $-\infty < x < +\infty$. Determine the probability that the particle is located somewhere between $x = -a$ and $x = +a$.

Section 41.2 Analysis Model: Quantum Particle Under Boundary Conditions

5. (a) Use the quantum-particle-in-a-box model to calculate the first three energy levels of a neutron trapped in an atomic nucleus of diameter 20.0 fm. (b) Explain whether the energy-level differences have a realistic order of magnitude.

6. An electron that has an energy of approximately 6 eV moves between infinitely high walls 1.00 nm apart. Find (a) the quantum number n for the energy state the electron occupies and (b) the precise energy of the electron.

7. An electron is contained in a one-dimensional box of length 0.100 nm. (a) Draw an energy-level diagram for the electron for levels up to $n = 4$. (b) Photons are emitted by the electron making downward transitions that could eventually carry it from the $n = 4$ state to the $n = 1$ state. Find the wavelengths of all such photons.

8. *Why is the following situation impossible?* A proton is in an infinitely deep potential well of length 1.00 nm. It absorbs a microwave photon of wavelength 6.06 mm and is excited into the next available quantum state.

9. AMT A ruby laser emits 694.3-nm light. Assume light of this wavelength is due to a transition of an electron in a box from its $n = 2$ state to its $n = 1$ state. Find the length of the box.

10. A laser emits light of wavelength λ. Assume this light is due to a transition of an electron in a box from its $n = 2$ state to its $n = 1$ state. Find the length of the box.

11. The nuclear potential energy that binds protons and neutrons in a nucleus is often approximated by a square well. Imagine a proton confined in an infinitely high square well of length 10.0 fm, a typical nuclear diameter. Assuming the proton makes a transition from the $n = 2$ state to the ground state, calculate (a) the energy and (b) the wavelength of the emitted photon. (c) Identify the region of the electromagnetic spectrum to which this wavelength belongs.

12. W A proton is confined to move in a one-dimensional box of length 0.200 nm. (a) Find the lowest possible energy of the proton. (b) **What If?** What is the lowest possible energy of an electron confined to the same box? (c) How do you account for the great difference in your results for parts (a) and (b)?

13. W An electron is confined to a one-dimensional region in which its ground-state ($n = 1$) energy is 2.00 eV. (a) What is the length L of the region? (b) What energy input is required to promote the electron to its first excited state?

14. A 4.00-g particle confined to a box of length L has a speed of 1.00 mm/s. (a) What is the classical kinetic energy of the particle? (b) If the energy of the first excited state ($n = 2$) is equal to the kinetic energy found in part (a), what is the value of L? (c) Is the result found in part (b) realistic? Explain.

15. A photon with wavelength λ is absorbed by an electron confined to a box. As a result, the electron moves from state $n = 1$ to $n = 4$. (a) Find the length of the box. (b) What is the wavelength λ' of the photon emitted in the transition of that electron from the state $n = 4$ to the state $n = 2$?

16. For a quantum particle of mass m in the ground state of a square well with length L and infinitely high walls, the uncertainty in position is $\Delta x \approx L$. (a) Use the uncertainty principle to estimate the uncertainty in its momentum. (b) Because the particle stays inside the box, its average momentum must be zero. Its average squared momentum is then $\langle p^2 \rangle \approx (\Delta p)^2$. Estimate the energy of the particle. (c) State how the result of part (b) compares with the actual ground-state energy.

17. A quantum particle is described by the wave function

$$\psi(x) = \begin{cases} A \cos\left(\dfrac{2\pi x}{L}\right) & \text{for } -\dfrac{L}{4} \le x \le \dfrac{L}{4} \\ 0 & \text{elsewhere} \end{cases}$$

(a) Determine the normalization constant A. (b) What is the probability that the particle will be found between $x = 0$ and $x = L/8$ if its position is measured?

18. The wave function for a quantum particle confined to moving in a one-dimensional box located between $x = 0$ and $x = L$ is

$$\psi(x) = A \sin\left(\frac{n\pi x}{L}\right)$$

Use the normalization condition on ψ to show that

$$A = \sqrt{\frac{2}{L}}$$

19. A quantum particle in an infinitely deep square well has a wave function given by

$$\psi_2(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{2\pi x}{L}\right)$$

for $0 \le x \le L$ and zero otherwise. (a) Determine the expectation value of x. (b) Determine the probability of finding the particle near $\frac{1}{2}L$ by calculating the probability that the particle lies in the range $0.490L \le x \le 0.510L$. (c) **What If?** Determine the probability of finding the particle near $\frac{1}{4}L$ by calculating the probability that the particle lies in the range $0.240L \le x \le 0.260L$.

(d) Argue that the result of part (a) does not contradict the results of parts (b) and (c).

20. An electron in an infinitely deep square well has a wave function that is given by

$$\psi_3(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{3\pi x}{L}\right)$$

for $0 \le x \le L$ and is zero otherwise. (a) What are the most probable positions of the electron? (b) Explain how you identify them.

21. An electron is trapped in an infinitely deep potential well 0.300 nm in length. (a) If the electron is in its ground state, what is the probability of finding it within 0.100 nm of the left-hand wall? (b) Identify the classical probability of finding the electron in this interval and state how it compares with the answer to part (a). (c) Repeat parts (a) and (b) assuming the particle is in the 99th energy state.

22. A quantum particle is in the $n = 1$ state of an infinitely deep square well with walls at $x = 0$ and $x = L$. Let ℓ be an arbitrary value of x between $x = 0$ and $x = L$. (a) Find an expression for the probability, as a function of ℓ, that the particle will be found between $x = 0$ and $x = \ell$. (b) Sketch the probability as a function of the variable ℓ/L. Choose values of ℓ/L ranging from 0 to 1.00 in steps of 0.100. (c) Explain why the probability function must have particular values at $\ell/L = 0$ and at $\ell/L = 1$. (d) Find the value of ℓ for which the probability of finding the particle between $x = 0$ and $x = \ell$ is twice the probability of finding the particle between $x = \ell$ and $x = L$. *Suggestion:* Solve the transcendental equation for ℓ/L numerically.

23. M A quantum particle in an infinitely deep square well has a wave function that is given by

$$\psi_1(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi x}{L}\right)$$

for $0 \le x \le L$ and is zero otherwise. (a) Determine the probability of finding the particle between $x = 0$ and $x = \frac{1}{3}L$. (b) Use the result of this calculation and a symmetry argument to find the probability of finding the particle between $x = \frac{1}{3}L$ and $x = \frac{2}{3}L$. Do not re-evaluate the integral.

Section 41.3 The Schrödinger Equation

24. Show that the wave function $\psi = Ae^{i(kx-\omega t)}$ is a solution to the Schrödinger equation (Eq. 41.15), where $k = 2\pi/\lambda$ and $U = 0$.

25. The wave function of a quantum particle of mass m is

$$\psi(x) = A\cos(kx) + B\sin(kx)$$

where A, B, and k are constants. (a) Assuming the particle is free ($U = 0$), show that $\psi(x)$ is a solution of the Schrödinger equation (Eq. 41.15). (b) Find the corresponding energy E of the particle.

26. Consider a quantum particle moving in a one-dimensional box for which the walls are at $x = -L/2$ and $x = L/2$. (a) Write the wave functions and probability densities for $n = 1$, $n = 2$, and $n = 3$. (b) Sketch the wave functions and probability densities.

27. In a region of space, a quantum particle with zero total energy has a wave function

$$\psi(x) = Axe^{-x^2/L^2}$$

(a) Find the potential energy U as a function of x. (b) Make a sketch of $U(x)$ versus x.

28. A quantum particle of mass m moves in a potential well of length $2L$. Its potential energy is infinite for $x < -L$ and for $x > +L$. In the region $-L < x < L$, its potential energy is given by

$$U(x) = \frac{-\hbar^2 x^2}{mL^2(L^2 - x^2)}$$

In addition, the particle is in a stationary state that is described by the wave function $\psi(x) = A(1 - x^2/L^2)$ for $-L < x < +L$ and by $\psi(x) = 0$ elsewhere. (a) Determine the energy of the particle in terms of $\hbar$, m, and L. (b) Determine the normalization constant A. (c) Determine the probability that the particle is located between $x = -L/3$ and $x = +L/3$.

Section 41.4 A Particle in a Well of Finite Height

29. Sketch (a) the wave function $\psi(x)$ and (b) the probability density $|\psi(x)|^2$ for the $n = 4$ state of a quantum particle in a finite potential well. (See Fig. 41.7.)

30. Suppose a quantum particle is in its ground state in a box that has infinitely high walls (see Fig. 41.4a). Now suppose the left-hand wall is suddenly lowered to a finite height and width. (a) Qualitatively sketch the wave function for the particle a short time later. (b) If the box has a length L, what is the wavelength of the wave that penetrates the left-hand wall?

Section 41.5 Tunneling Through a Potential Energy Barrier

31. M An electron with kinetic energy $E = 5.00$ eV is incident on a barrier of width $L = 0.200$ nm and height $U = 10.0$ eV (Fig. P41.31). What is the probability that the electron (a) tunnels through the barrier? (b) Is reflected?

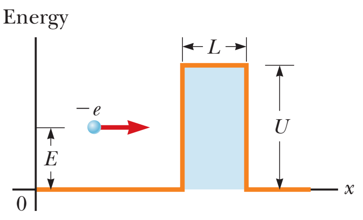


Figure P41.31 Problems 31 and 32.

32. W An electron having total energy $E = 4.50$ eV approaches a rectangular energy barrier with $U = 5.00$ eV and $L = 950$ pm as shown in Figure P41.31. Classically, the electron cannot pass through the barrier because $E < U$. Quantum-mechanically, however, the probability of tunneling is not zero. (a) Calculate this probability, which is the transmission coefficient. (b) To what value would the width L of the potential barrier have to be increased for the chance of an incident 4.50-eV electron tunneling through the barrier to be one in one million?

33. W An electron has a kinetic energy of 12.0 eV. The electron is incident upon a rectangular barrier of height 20.0 eV and width 1.00 nm. If the electron absorbed all the energy of a photon of green light (with wavelength 546 nm) at the instant it reached the barrier, by what factor would the electron's probability of tunneling through the barrier increase?

Section 41.6 Applications of Tunneling

34. W A scanning tunneling microscope (STM) can precisely determine the depths of surface features because the current through its tip is very sensitive to differences in the width of the gap between the tip and the sample surface. Assume the electron wave function falls off exponentially in this direction with a decay length of 0.100 nm, that is, with $C = 10.0\ \text{nm}^{-1}$. Determine the ratio of the current when the STM tip is 0.500 nm above a surface feature to the current when the tip is 0.515 nm above the surface.

35. M The design criterion for a typical scanning tunneling microscope (STM) specifies that it must be able to detect, on the sample below its tip, surface features that differ in height by only 0.002 00 nm. Assuming the electron transmission coefficient is e^{-2CL} with $C = 10.0\ \text{nm}^{-1}$, what percentage change in electron transmission must the electronics of the STM be able to detect to achieve this resolution?

Section 41.7 The Simple Harmonic Oscillator

36. A one-dimensional harmonic oscillator wave function is

$$\psi = Axe^{-bx^2}$$

(a) Show that ψ satisfies Equation 41.24. (b) Find b and the total energy E. (c) Is this wave function for the ground state or for the first excited state?

37. W A quantum simple harmonic oscillator consists of an electron bound by a restoring force proportional to its position relative to a certain equilibrium point. The proportionality constant is 8.99 N/m. What is the longest wavelength of light that can excite the oscillator?

38. A quantum simple harmonic oscillator consists of a particle of mass m bound by a restoring force proportional to its position relative to a certain equilibrium point. The proportionality constant is k. What is the longest wavelength of light that can excite the oscillator?

39. (a) Normalize the wave function for the ground state of a simple harmonic oscillator. That is, apply Equation 41.7 to Equation 41.26 and find the required value for the constant B in terms of m, ω, and fundamental constants. (b) Determine the probability of finding the oscillator in a narrow interval $-\delta/2 < x < \delta/2$ around its equilibrium position.

40. Two particles with masses m_1 and m_2 are joined by a light spring of force constant k. They vibrate along a straight line with their center of mass fixed. (a) Show that the total energy

$$\tfrac{1}{2}m_1u_1^{\,2} + \tfrac{1}{2}m_2u_2^{\,2} + \tfrac{1}{2}kx^2$$

can be written as $\frac{1}{2}\mu u^2 + \frac{1}{2}kx^2$, where $u = |u_1| + |u_2|$ is the *relative* speed of the particles and $\mu = m_1m_2/(m_1 + m_2)$ is the reduced mass of the system. This result demonstrates that the pair of freely vibrating particles can be precisely modeled as a single particle vibrating on one end of a spring that has its other end fixed. (b) Differentiate the equation

$$\tfrac{1}{2}\mu u^2 + \tfrac{1}{2}kx^2 = \text{constant}$$

with respect to x. Proceed to show that the system executes simple harmonic motion. (c) Find its frequency.

41. The total energy of a particle–spring system in which the particle moves with simple harmonic motion along the x axis is

$$E = \frac{p_x^{\,2}}{2m} + \frac{kx^2}{2}$$

where p_x is the momentum of the quantum particle and k is the spring constant. (a) Using the uncertainty principle, show that this expression can also be written as

$$E \geq \frac{p_x^{\,2}}{2m} + \frac{k\hbar^2}{8p_x^{\,2}}$$

(b) Show that the minimum energy of the harmonic oscillator is

$$E_{\min} = K + U = \tfrac{1}{4}\hbar\sqrt{\frac{k}{m}} + \frac{\hbar\omega}{4} = \frac{\hbar\omega}{2}$$

42. Show that Equation 41.26 is a solution of Equation 41.24 with energy $E = \frac{1}{2}\hbar\omega$.

Additional Problems

43. W A particle of mass 2.00×10^{-28} kg is confined to a one-dimensional box of length 1.00×10^{-10} m. For $n = 1$, what are (a) the particle's wavelength, (b) its momentum, and (c) its ground-state energy?

44. Prove that the first term in the Schrödinger equation, $-(\hbar^2/2m)(d^2\psi/dx^2)$, reduces to the kinetic energy of the quantum particle multiplied by the wave function (a) for a freely moving particle, with the wave function given by Equation 41.4, and (b) for a particle in a box, with the wave function given by Equation 41.13.

45. A particle in a one-dimensional box of length L is in its first excited state, corresponding to $n = 2$. Determine the probability of finding the particle between $x = 0$ and $x = L/4$.

46. Prove that assuming $n = 0$ for a quantum particle in an infinitely deep potential well leads to a violation of the uncertainty principle $\Delta p_x \Delta x \geq \hbar/2$.

47. Calculate the transmission probability for quantum-mechanical tunneling in each of the following cases. (a) An electron with an energy deficit of $U - E = 0.010\ 0$ eV is incident on a square barrier of width $L = 0.100$ nm. (b) An electron with an energy deficit of 1.00 eV is incident on the same barrier. (c) An alpha particle (mass 6.64×10^{-27} kg) with an energy deficit of 1.00 MeV is incident on a square barrier of width 1.00 fm. (d) An 8.00-kg bowling ball with an energy deficit of 1.00 J is incident on a square barrier of width 2.00 cm.

48. An electron in an infinitely deep potential well has a ground-state energy of 0.300 eV. (a) Show that the photon emitted in a transition from the $n = 3$ state to the $n = 1$ state has a wavelength of 517 nm, which makes it green visible light. (b) Find the wavelength and the spectral region for each of the other five transitions that take place among the four lowest energy levels.

49. An atom in an excited state 1.80 eV above the ground state remains in that excited state 2.00 μs before moving to the ground state. Find (a) the frequency and (b) the wavelength of the emitted photon. (c) Find the approximate uncertainty in energy of the photon.

50. A marble rolls back and forth across a shoebox at a constant speed of 0.8 m/s. Make an order-of-magnitude estimate of the probability of it escaping through the wall of the box by quantum tunneling. State the quantities you take as data and the values you measure or estimate for them.

51. An electron confined to a box absorbs a photon with wavelength λ. As a result, the electron makes a transition from the $n = 1$ state to the $n = 3$ state. (a) Find the length of the box. (b) What is the wavelength λ' of the photon emitted when the electron makes a transition from the $n = 3$ state to the $n = 2$ state?

52. For a quantum particle described by a wave function $\psi(x)$, the expectation value of a physical quantity $f(x)$ associated with the particle is defined by

$$\langle f(x) \rangle \equiv \int_{-\infty}^{\infty} \psi^* f(x) \psi \, dx$$

For a particle in an infinitely deep one-dimensional box extending from $x = 0$ to $x = L$, show that

$$\langle x^2 \rangle = \frac{L^2}{3} - \frac{L^2}{2n^2\pi^2}$$

53. A quantum particle of mass m is placed in a one-dimensional box of length L. Assume the box is so small that the particle's motion is relativistic and $K = p^2/2m$ is not valid. (a) Derive an expression for the kinetic energy levels of the particle. (b) Assume the particle is an electron in a box of length $L = 1.00 \times 10^{-12}$ m. Find its lowest possible kinetic energy. (c) By what percent is the nonrelativistic equation in error? *Suggestion:* See Equation 39.23.

54. *Why is the following situation impossible?* A particle is in the ground state of an infinite square well of length L. A light source is adjusted so that the photons of wavelength λ are absorbed by the particle as it makes a transition to the first excited state. An identical particle is in the ground state of a finite square well of length L. The light source sends photons of the same wavelength λ toward this particle. The photons are not absorbed because the allowed energies of the finite square well are different from those of the infinite square well. To cause the photons to be absorbed, you move the light source at a high speed toward the particle in the finite square well. You are able to find a speed at which the Doppler-shifted photons are absorbed as the particle makes a transition to the first excited state.

55. A quantum particle has a wave function

$$\psi(x) = \begin{cases} \sqrt{\dfrac{2}{a}}\, e^{-x/a} & \text{for } x > 0 \\ 0 & \text{for } x < 0 \end{cases}$$

(a) Find and sketch the probability density. (b) Find the probability that the particle will be at any point where $x < 0$. (c) Show that ψ is normalized and then (d) find the probability of finding the particle between $x = 0$ and $x = a$.

56. An electron is confined to move in the xy plane in a rectangle whose dimensions are L_x and L_y. That is, the electron is trapped in a two-dimensional potential well having lengths of L_x and L_y. In this situation, the allowed energies of the electron depend on two quantum numbers n_x and n_y and are given by

$$E = \frac{h^2}{8m_e}\left(\frac{n_x^{\,2}}{L_x^{\,2}} + \frac{n_y^{\,2}}{L_y^{\,2}}\right)$$

Using this information, we wish to find the wavelength of a photon needed to excite the electron from the ground state to the second excited state, assuming $L_x = L_y = L$. (a) Using the assumption on the lengths, write an expression for the allowed energies of the electron in terms of the quantum numbers n_x and n_y. (b) What values of n_x and n_y correspond to the ground state? (c) Find the energy of the ground state. (d) What are the possible values of n_x and n_y for the first excited state, that is, the next-highest state in terms of energy? (e) What are the possible values of n_x and n_y for the second excited state? (f) Using the values in part (e), what is the energy of the second excited state? (g) What is the energy difference between the ground state and the second excited state? (h) What is the wavelength of a photon that will cause the transition between the ground state and the second excited state?

57. The normalized wave functions for the ground state, $\psi_0(x)$, and the first excited state, $\psi_1(x)$, of a quantum harmonic oscillator are

$$\psi_0(x) = \left(\frac{a}{\pi}\right)^{1/4} e^{-ax^2/2} \qquad \psi_1(x) = \left(\frac{4a^3}{\pi}\right)^{1/4} x e^{-ax^2/2}$$

where $a = m\omega/\hbar$. A mixed state, $\psi_{01}(x)$, is constructed from these states:

$$\psi_{01}(x) = \frac{1}{\sqrt{2}}\left[\psi_0(x) + \psi_1(x)\right]$$

The symbol $\langle q \rangle_s$ denotes the expectation value of the quantity q for the state $\psi_s(x)$. Calculate the expectation values (a) $\langle x \rangle_0$, (b) $\langle x \rangle_1$, and (c) $\langle x \rangle_{01}$.

58. A two-slit electron diffraction experiment is done with slits of *unequal* widths. When only slit 1 is open, the number of electrons reaching the screen per second is 25.0 times the number of electrons reaching the screen per second when only slit 2 is open. When both slits are open, an interference pattern results in which the destructive interference is not complete. Find the ratio of the probability of an electron arriving at an interference maximum to the probability of an electron arriving at an adjacent interference minimum. *Suggestion:* Use the superposition principle.

Challenge Problems

59. Particles incident from the left in Figure P41.59 are confronted with a step in potential energy. The step has a height U at $x = 0$. The particles have energy $E > U$. Classically, all the particles would continue moving forward with reduced speed. According to quantum mechanics, however, a fraction of the particles are reflected at the step. (a) Prove that the reflection coefficient R for this case is

$$R = \frac{(k_1 - k_2)^2}{(k_1 + k_2)^2}$$

where $k_1 = 2\pi/\lambda_1$ and $k_2 = 2\pi/\lambda_2$ are the wave numbers for the incident and transmitted particles, respectively. Proceed as follows. Show that the wave function $\psi_1 = Ae^{ik_1x} + Be^{-ik_1x}$ satisfies the Schrödinger equation in region 1, for $x < 0$. Here Ae^{ik_1x} represents the incident beam and Be^{-ik_1x} represents the reflected particles. Show that $\psi_2 = Ce^{ik_2x}$ satisfies the Schrödinger equation in region 2, for $x > 0$. Impose the boundary conditions $\psi_1 = \psi_2$ and $d\psi_1/dx = d\psi_2/dx$, at $x = 0$, to find the relationship between B and A. Then evaluate $R = B^2/A^2$. A particle that has kinetic energy $E = 7.00$ eV is incident from a region where the potential energy is zero onto one where $U = 5.00$ eV. Find (b) its probability of being reflected and (c) its probability of being transmitted.

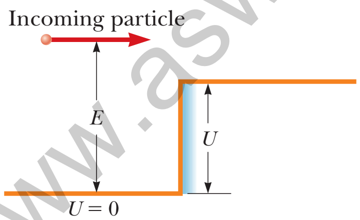


Figure P41.59

60. Consider a "crystal" consisting of two fixed ions of charge $+e$ and two electrons as shown in Figure P41.60. (a) Taking into account all the pairs of interactions, find the potential energy of the system as a function of d. (b) Assuming the electrons to be restricted to a one-dimensional box of length $3d$, find the minimum kinetic energy of the two electrons. (c) Find the value of d for which the total energy is a minimum. (d) State how this value of d compares with the spacing of atoms in lithium, which has a density of 0.530 g/cm^3 and a molar mass of 6.94 g/mol.

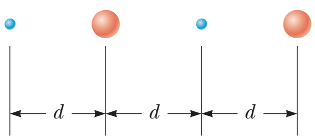


Figure P41.60

61. An electron is trapped in a quantum dot. The quantum dot may be modeled as a one-dimensional, rigid-walled box of length 1.00 nm. (a) Taking $x = 0$ as the left side of the box, calculate the probability of finding the electron between $x_1 = 0.150$ nm and $x_2 = 0.350$ nm for the $n = 1$ state. (b) Repeat part (a) for the $n = 2$ state. Calculate the energies in electron volts of (c) the $n = 1$ state and (d) the $n = 2$ state.

62. An electron is represented by the time-independent wave function

$$\psi(x) = \begin{cases} Ae^{-\alpha x} & \text{for } x > 0 \\ Ae^{+\alpha x} & \text{for } x < 0 \end{cases}$$

(a) Sketch the wave function as a function of x. (b) Sketch the probability density representing the likelihood that the electron is found between x and $x + dx$. (c) Only an infinite value of potential energy could produce the discontinuity in the derivative of the wave function at $x = 0$. Aside from this feature, argue that $\psi(x)$ can be a physically reasonable wave function. (d) Normalize the wave function. (e) Determine the probability of finding the electron somewhere in the range

$$-\frac{1}{2\alpha} \leq x \leq \frac{1}{2\alpha}$$

63. The wave function

$$\psi(x) = Bxe^{-(m\omega/2\hbar)x^2}$$

is a solution to the simple harmonic oscillator problem. (a) Find the energy of this state. (b) At what position are you least likely to find the particle? (c) At what positions are you most likely to find the particle? (d) Determine the value of B required to normalize the wave function. (e) **What If?** Determine the classical probability of finding the particle in an interval of small length δ centered at the position $x = 2(\hbar/m\omega)^{1/2}$. (f) What is the actual probability of finding the particle in this interval?

64. (a) Find the normalization constant A for a wave function made up of the two lowest states of a quantum particle in a box extending from $x = 0$ to $x = L$:

$$\psi(x) = A\left[\sin\left(\frac{\pi x}{L}\right) + 4\sin\left(\frac{2\pi x}{L}\right)\right]$$

(b) A particle is described in the space $-a \leq x \leq a$ by the wave function

$$\psi(x) = A\cos\left(\frac{\pi x}{2a}\right) + B\sin\left(\frac{\pi x}{a}\right)$$

Determine the relationship between the values of A and B required for normalization.

CHAPTER

42

Atomic Physics

This street in the Ginza district in Tokyo displays many signs formed from neon lamps of varying bright colors. The light from these lamps has its origin in transitions between quantized energy states in the atoms contained in the lamps. In this chapter, we investigate those transitions. *(© Ken Straiton/Corbis)*

In Chapter 41, we introduced some basic concepts and techniques used in quantum mechanics along with their applications to various one-dimensional systems. In this chapter, we apply quantum mechanics to atomic systems. A large portion of the chapter is focused on the application of quantum mechanics to the study of the hydrogen atom. Understanding the hydrogen atom, the simplest atomic system, is important for several reasons:

- The hydrogen atom is the only atomic system that can be solved exactly.
- Much of what was learned in the 20th century about the hydrogen atom, with its single electron, can be extended to such single-electron ions as He^{+} and Li^{2+}.
- The hydrogen atom is an ideal system for performing precise tests of theory against experiment and for improving our overall understanding of atomic structure.

- The quantum numbers that are used to characterize the allowed states of hydrogen can also be used to investigate more complex atoms, and such a description enables us to understand the periodic table of the elements. This understanding is one of the greatest triumphs of quantum mechanics.
- The basic ideas about atomic structure must be well understood before we attempt to deal with the complexities of molecular structures and the electronic structure of solids.

The full mathematical solution of the Schrödinger equation applied to the hydrogen atom gives a complete and beautiful description of the atom's properties. Because the mathematical procedures involved are beyond the scope of this text, however, many details are omitted. The solutions for some states of hydrogen are discussed, together with the quantum numbers used to characterize various allowed states. We also discuss the physical significance of the quantum numbers and the effect of a magnetic field on certain quantum states.

A new physical idea, the *exclusion principle*, is presented in this chapter. This principle is extremely important for understanding the properties of multielectron atoms and the arrangement of elements in the periodic table.

Finally, we apply our knowledge of atomic structure to describe the mechanisms involved in the production of x-rays and in the operation of a laser.

42.1 Atomic Spectra of Gases

As pointed out in Section 40.1, all objects emit thermal radiation characterized by a *continuous* distribution of wavelengths. In sharp contrast to this continuous-distribution spectrum is the *discrete* **line spectrum** observed when a low-pressure gas undergoes an electric discharge. (Electric discharge occurs when the gas is subject to a potential difference that creates an electric field greater than the dielectric strength of the gas.) Observation and analysis of these spectral lines is called **emission spectroscopy.**

When the light from a gas discharge is examined using a spectrometer (see Fig. 38.15), it is found to consist of a few bright lines of color on a generally dark background. This discrete line spectrum contrasts sharply with the continuous rainbow of colors seen when a glowing solid is viewed through the same instrument. Figure 42.1a (page 1298) shows that the wavelengths contained in a given line spectrum are characteristic of the element emitting the light. The simplest line spectrum is that for atomic hydrogen, and we describe this spectrum in detail. Because no two elements have the same line spectrum, this phenomenon represents a practical and sensitive technique for identifying the elements present in unknown samples.

Another form of spectroscopy very useful in analyzing substances is **absorption spectroscopy.** An absorption spectrum is obtained by passing white light from a continuous source through a gas or a dilute solution of the element being analyzed. The absorption spectrum consists of a series of dark lines superimposed on the continuous spectrum of the light source as shown in Figure 42.1b for atomic hydrogen.

The absorption spectrum of an element has many practical applications. For example, the continuous spectrum of radiation emitted by the Sun must pass through the cooler gases of the solar atmosphere. The various absorption lines observed in the solar spectrum have been used to identify elements in the solar atmosphere. In early studies of the solar spectrum, experimenters found some lines that did not correspond to any known element. A new element had been discovered!

Pitfall Prevention 42.1

Why Lines? The phrase "spectral lines" is often used when discussing the radiation from atoms. Lines are seen because the light passes through a long and very narrow slit before being separated by wavelength. You will see many references to these "lines" in both physics and chemistry.

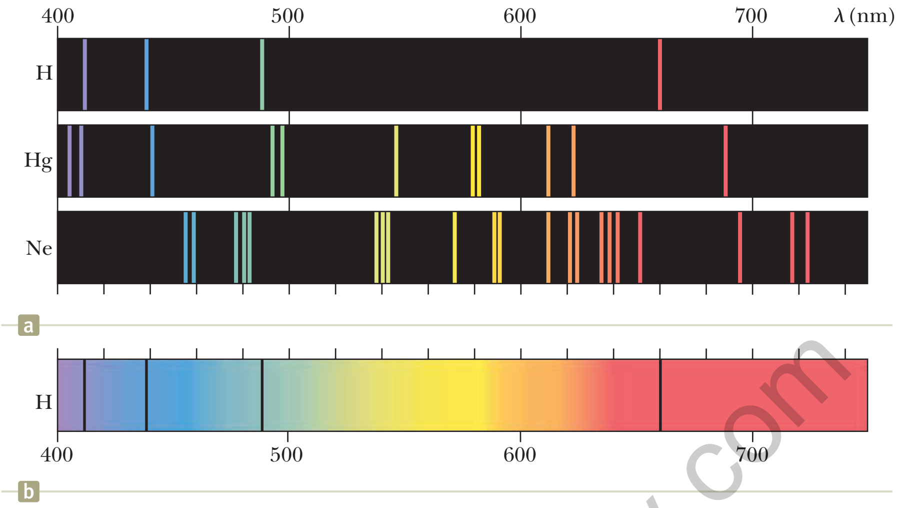


Figure 42.1 (a) Emission line spectra for hydrogen, mercury, and neon. (b) The absorption spectrum for hydrogen. Notice that the dark absorption lines occur at the same wavelengths as the hydrogen emission lines in (a). (K. W. Whitten, R. E. Davis, M. L. Peck, and G. G. Stanley, *General Chemistry,* 7th ed., Belmont, CA, Brooks/Cole, 2004.)

The new element was named helium, after the Greek word for Sun, *helios.* Helium was subsequently isolated from subterranean gas on the Earth.

Using this technique, scientists have examined the light from stars other than our Sun and have never detected elements other than those present on the Earth. Absorption spectroscopy has also been useful in analyzing heavy-metal contamination of the food chain. For example, the first determination of high levels of mercury in tuna was made with the use of atomic absorption spectroscopy.

The discrete emissions of light from gas discharges are used in "neon" signs such as those in the opening photograph of this chapter. Neon, the first gas used in these types of signs and the gas after which these signs are named, emits strongly in the red region. As a result, a glass tube filled with neon gas emits bright red light when an applied voltage causes a continuous discharge. Early signs used different gases to provide different colors, although the brightness of these signs was generally very low. Many present-day "neon" signs contain mercury vapor, which emits strongly in the ultraviolet range of the electromagnetic spectrum. The inside of a present-day sign's glass tube is coated with a material that emits a particular color when it absorbs ultraviolet radiation from the mercury. The color of the light from the tube results from the particular material chosen. A household fluorescent light operates in the same manner, with a white-emitting material coating the inside of the glass tube.

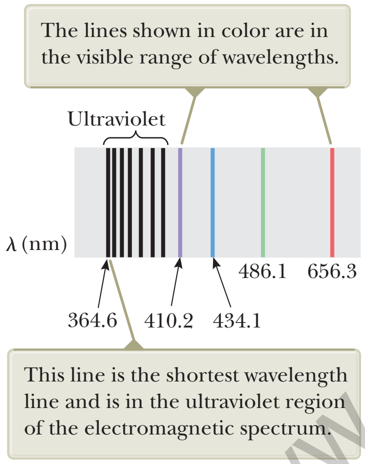


Figure 42.2 The Balmer series of spectral lines for atomic hydrogen, with several lines marked with the wavelength in nanometers. (The horizontal wavelength axis is not to scale.)

From 1860 to 1885, scientists accumulated a great deal of data on atomic emissions using spectroscopic measurements. In 1885, a Swiss schoolteacher, Johann Jacob Balmer (1825–1898), found an empirical equation that correctly predicted the wavelengths of four visible emission lines of hydrogen: H_α (red), H_β (blue-green), H_γ (blue-violet), and H_δ (violet). Figure 42.2 shows these and other lines (in the ultraviolet) in the emission spectrum of hydrogen. The four visible lines occur at the wavelengths 656.3 nm, 486.1 nm, 434.1 nm, and 410.2 nm. The complete set of lines is called the **Balmer series.** The wavelengths of these lines can be described by the following equation, which is a modification made by Johannes Rydberg (1854–1919) of Balmer's original equation:

Balmer series ▶

$$\frac{1}{\lambda} = R_H\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \quad n = 3, 4, 5, \ldots \tag{42.1}$$

where R_H is a constant now called the **Rydberg constant** with a value of $1.097\,373\,2 \times 10^7 \text{ m}^{-1}$. The integer values of n from 3 to 6 give the four visible lines from 656.3 nm (red) down to 410.2 nm (violet). Equation 42.1 also describes the ultraviolet spectral lines in the Balmer series if n is carried out beyond $n = 6$. The **series limit** is the shortest wavelength in the series and corresponds to $n \to \infty$, with a wavelength of 364.6 nm as in Figure 42.2. The measured spectral lines agree with the empirical equation, Equation 42.1, to within 0.1%.

Other lines in the spectrum of hydrogen were found following Balmer's discovery. These spectra are called the Lyman, Paschen, and Brackett series after their discoverers. The wavelengths of the lines in these series can be calculated through the use of the following empirical equations:

$$\frac{1}{\lambda} = R_H\left(1 - \frac{1}{n^2}\right) \quad n = 2, 3, 4, \ldots \tag{42.2}$$

◀ **Lyman series**

$$\frac{1}{\lambda} = R_H\left(\frac{1}{3^2} - \frac{1}{n^2}\right) \quad n = 4, 5, 6, \ldots \tag{42.3}$$

◀ **Paschen series**

$$\frac{1}{\lambda} = R_H\left(\frac{1}{4^2} - \frac{1}{n^2}\right) \quad n = 5, 6, 7, \ldots \tag{42.4}$$

◀ **Brackett series**

No theoretical basis existed for these equations; they simply worked. The same constant R_H appears in each equation, and all equations involve small integers. In Section 42.3, we shall discuss the remarkable achievement of a theory for the hydrogen atom that provided an explanation for these equations.

42.2 Early Models of the Atom

The model of the atom in the days of Newton was a tiny, hard, indestructible sphere. Although this model provided a good basis for the kinetic theory of gases (Chapter 21), new models had to be devised when experiments revealed the electrical nature of atoms. In 1897, J. J. Thomson established the charge-to-mass ratio for electrons. (See Fig. 29.15 in Section 29.3.) The following year, he suggested a model that describes the atom as a region in which positive charge is spread out in space with electrons embedded throughout the region, much like the seeds in a watermelon or raisins in thick pudding (Fig. 42.3). The atom as a whole would then be electrically neutral.

In 1911, Ernest Rutherford (1871–1937) and his students Hans Geiger and Ernest Marsden performed a critical experiment that showed that Thomson's model could not be correct. In this experiment, a beam of positively charged alpha particles (helium nuclei) was projected into a thin metallic foil such as the target in Figure 42.4a (page 1300). Most of the particles passed through the foil as if it were empty space, but some of the results of the experiment were astounding. Many of the particles deflected from their original direction of travel were scattered through *large* angles. Some particles were even deflected backward, completely reversing their direction of travel! When Geiger informed Rutherford that some alpha particles were scattered backward, Rutherford wrote, "It was quite the most incredible event that has ever happened to me in my life. It was almost as incredible as if you fired a 15-inch [artillery] shell at a piece of tissue paper and it came back and hit you."

Such large deflections were not expected on the basis of Thomson's model. According to that model, the positive charge of an atom in the foil is spread out over such a great volume (the entire atom) that there is no concentration of positive charge strong enough to cause any large-angle deflections of the positively charged alpha particles. Furthermore, the electrons are so much less massive than the alpha particles that they would not cause large-angle scattering either. Rutherford explained his astonishing results by developing a new atomic model, one that assumed the positive charge in the atom was concentrated in a region that was small relative to the size of the atom. He called this concentration of positive charge the **nucleus** of the atom. Any electrons belonging to the atom were assumed to be in the relatively large volume outside the nucleus. To explain why these electrons were not pulled into the nucleus by the attractive electric force, Rutherford modeled them as moving in orbits around the nucleus in the same manner as the planets orbit the Sun (Fig. 42.4b). For this reason, this model is often referred to as the planetary model of the atom.

Two basic difficulties exist with Rutherford's planetary model. As we saw in Section 42.1, an atom emits (and absorbs) certain characteristic frequencies of

Stock Montage, Inc.

Joseph John Thomson
English physicist (1856–1940)
The recipient of a Nobel Prize in Physics in 1906, Thomson is usually considered the discoverer of the electron. He opened up the field of subatomic particle physics with his extensive work on the deflection of cathode rays (electrons) in an electric field.

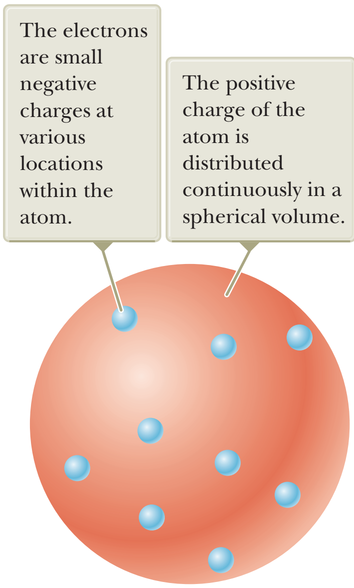


Figure 42.3 Thomson's model of the atom.

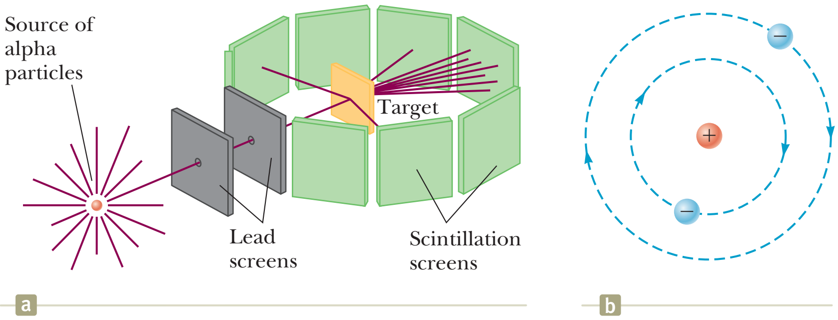


Figure 42.4 (a) Rutherford's technique for observing the scattering of alpha particles from a thin foil target. The source is a naturally occurring radioactive substance, such as radium. (b) Rutherford's planetary model of the atom.

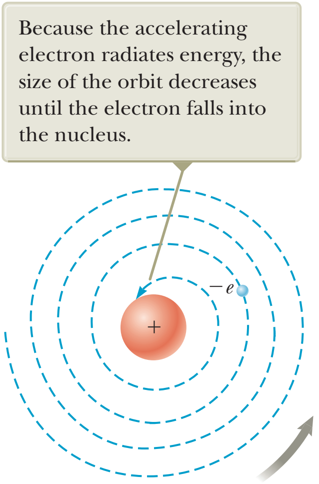


Figure 42.5 The classical model of the nuclear atom predicts that the atom decays.

electromagnetic radiation and no others, but the Rutherford model cannot explain this phenomenon. A second difficulty is that Rutherford's electrons are described by the particle in uniform circular motion model; they have a centripetal acceleration. According to Maxwell's theory of electromagnetism, centripetally accelerated charges revolving with frequency f should radiate electromagnetic waves of frequency f. Unfortunately, this classical model leads to a prediction of self-destruction when applied to the atom. Identifying the electron and the proton as a nonisolated system for energy, Equation 8.2 becomes $\Delta K + \Delta U = T_{ER}$, where K is the kinetic energy of the electron, U is the electric potential energy of the electron–nucleus system, and T_{ER} represents the outgoing electromagnetic radiation. As energy leaves the system, the radius of the electron's orbit steadily decreases (Fig. 42.5). The system is an isolated system for angular momentum because there is no torque on the system. Therefore, as the electron moves closer to the nucleus, the angular speed of the electron will increase, just like the spinning skater in Figure 11.10 in Section 11.4. This process leads to an ever-increasing frequency of emitted radiation and an ultimate collapse of the atom as the electron plunges into the nucleus.

42.3 Bohr's Model of the Hydrogen Atom

Given the situation described at the end of Section 42.2, the stage was set for Niels Bohr in 1913 when he presented a new model of the hydrogen atom that circumvented the difficulties of Rutherford's planetary model. Bohr applied Planck's ideas of quantized energy levels (Section 40.1) to Rutherford's orbiting atomic electrons. Bohr's theory was historically important to the development of quantum physics, and it appeared to explain the spectral line series described by Equations 42.1 through 42.4. Although Bohr's model is now considered obsolete and has been completely replaced by a probabilistic quantum-mechanical theory, we can use the Bohr model to develop the notions of energy quantization and angular momentum quantization as applied to atomic-sized systems.

Bohr combined ideas from Planck's original quantum theory, Einstein's concept of the photon, Rutherford's planetary model of the atom, and Newtonian mechanics to arrive at a semiclassical structural model based on some revolutionary ideas. The structural model of the Bohr theory as it applies to the hydrogen atom has the following properties:

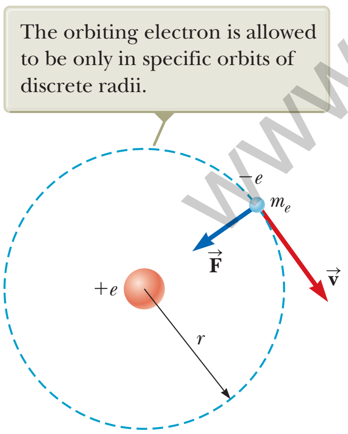


Figure 42.6 Diagram representing Bohr's model of the hydrogen atom.

1. *Physical components:*
   The electron moves in circular orbits around the proton under the influence of the electric force of attraction as shown in Figure 42.6.

2. *Behavior of the components:*

(a) Only certain electron orbits are stable. When in one of these **stationary states,** as Bohr called them, the electron does not emit energy in the form of radiation, even though it is accelerating. Hence, the total energy of the atom remains constant and classical mechanics can be used to describe the electron's motion. Bohr's model claims that the centripetally accelerated electron does not continuously emit radiation, losing energy and eventually spiraling into the nucleus, as predicted by classical physics in the form of Rutherford's planetary model.

(b) The atom emits radiation when the electron makes a transition from a more energetic initial stationary state to a lower-energy stationary state. This transition cannot be visualized or treated classically. In particular, the frequency f of the photon emitted in the transition is related to the change in the atom's energy and is not equal to the frequency of the electron's orbital motion. The frequency of the emitted radiation is found from the energy-conservation expression

$$E_i - E_f = hf \tag{42.5}$$

where E_i is the energy of the initial state, E_f is the energy of the final state, and $E_i > E_f$. In addition, energy of an incident photon can be absorbed by the atom, but only if the photon has an energy that exactly matches the difference in energy between an allowed state of the atom and a higher-energy state. Upon absorption, the photon disappears and the atom makes a transition to the higher-energy state.

(c) The size of an allowed electron orbit is determined by a condition imposed on the electron's orbital angular momentum: the allowed orbits are those for which the electron's orbital angular momentum about the nucleus is quantized and equal to an integral multiple of $\hbar = h/2\pi$,

$$m_e v r = n\hbar \quad n = 1, 2, 3, \ldots \tag{42.6}$$

where m_e is the electron mass, v is the electron's speed in its orbit, and r is the orbital radius.

Niels Bohr

Danish Physicist (1885–1962)

Bohr was an active participant in the early development of quantum mechanics and provided much of its philosophical framework. During the 1920s and 1930s, he headed the Institute for Advanced Studies in Copenhagen. The institute was a magnet for many of the world's best physicists and provided a forum for the exchange of ideas. Bohr was awarded the 1922 Nobel Prize in Physics for his investigation of the structure of atoms and the radiation emanating from them. When Bohr visited the United States in 1939 to attend a scientific conference, he brought news that the fission of uranium had been observed by Hahn and Strassman in Berlin. The results were the foundations of the nuclear weapon developed in the United States during World War II.

These postulates are a mixture of established principles and completely new and untested ideas at the time. Property 1, from classical mechanics, treats the electron in orbit around the nucleus in the same way we treat a planet in a circular orbit around a star, using the particle in uniform circular motion analysis model. Property 2(a) was a radical new idea in 1913 that was completely at odds with the understanding of electromagnetism at the time. Property 2(b) represents the principle of conservation of energy as described by the nonisolated system model for energy. Property 2(c) is another new idea that had no basis in classical physics.

Property 2(b) implies qualitatively the existence of a characteristic discrete emission line spectrum *and also* a corresponding absorption line spectrum of the kind shown in Figure 42.1 for hydrogen. Using these postulates, let's calculate the allowed energy levels and find quantitative values of the emission wavelengths of the hydrogen atom.

The electric potential energy of the system shown in Figure 42.6 is given by Equation 25.13, $U = k_e q_1 q_2/r = -k_e e^2/r$, where k_e is the Coulomb constant and the negative sign arises from the charge $-e$ on the electron. Therefore, the *total* energy of the atom, which consists of the electron's kinetic energy and the system's potential energy, is

$$E = K + U = \tfrac{1}{2}m_e v^2 - k_e \frac{e^2}{r} \tag{42.7}$$

The electron is modeled as a particle in uniform circular motion, so the electric force $k_e e^2/r^2$ exerted on the electron must equal the product of its mass and its centripetal acceleration ($a_c = v^2/r$):

$$\frac{k_e e^2}{r^2} = \frac{m_e v^2}{r}$$

$$v^2 = \frac{k_e e^2}{m_e r} \qquad \textbf{(42.8)}$$

From Equation 42.8, we find that the kinetic energy of the electron is

$$K = \tfrac{1}{2} m_e v^2 = \frac{k_e e^2}{2r}$$

Substituting this value of K into Equation 42.7 gives the following expression for the total energy of the atom:[1]

$$E = -\frac{k_e e^2}{2r} \qquad \textbf{(42.9)}$$

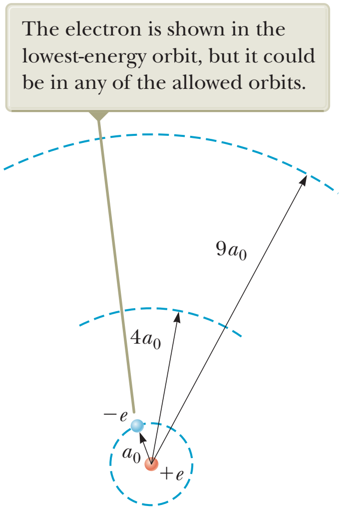


Figure 42.7 The first three circular orbits predicted by the Bohr model of the hydrogen atom.

Because the total energy is *negative,* which indicates a bound electron–proton system, energy in the amount of $k_e e^2/2r$ must be added to the atom to remove the electron and make the total energy of the system zero.

We can obtain an expression for r, the radius of the allowed orbits, by solving Equation 42.6 for v^2 and equating it to Equation 42.8:

$$v^2 = \frac{n^2\hbar^2}{m_e^2 r^2} = \frac{k_e e^2}{m_e r}$$

$$r_n = \frac{n^2\hbar^2}{m_e k_e e^2} \qquad n = 1, 2, 3, \ldots \qquad \textbf{(42.10)}$$

Equation 42.10 shows that the radii of the allowed orbits have discrete values: they are quantized. The result is based on the *assumption* that the electron can exist only in certain allowed orbits determined by the integer n (Bohr's Property 2(c)).

The orbit with the smallest radius, called the **Bohr radius** a_0, corresponds to $n = 1$ and has the value

Bohr radius ▶

$$a_0 = \frac{\hbar^2}{m_e k_e e^2} = 0.052\,9 \text{ nm} \qquad \textbf{(42.11)}$$

Substituting Equation 42.11 into Equation 42.10 gives a general expression for the radius of any orbit in the hydrogen atom:

Radii of Bohr orbits in hydrogen ▶

$$r_n = n^2 a_0 = n^2(0.052\,9 \text{ nm}) \qquad n = 1, 2, 3, \ldots \qquad \textbf{(42.12)}$$

Bohr's theory predicts a value for the radius of a hydrogen atom on the right order of magnitude, based on experimental measurements. This result was a striking triumph for Bohr's theory. The first three Bohr orbits are shown to scale in Figure 42.7.

The quantization of orbit radii leads to energy quantization. Substituting $r_n = n^2 a_0$ into Equation 42.9 gives

$$\boxed{E_n = -\frac{k_e e^2}{2a_0}\left(\frac{1}{n^2}\right) \qquad n = 1, 2, 3, \ldots} \qquad \textbf{(42.13)}$$

Inserting numerical values into this expression, we find that

$$\boxed{E_n = -\frac{13.606 \text{ eV}}{n^2} \qquad n = 1, 2, 3, \ldots} \qquad \textbf{(42.14)}$$

[1]Compare Equation 42.9 with its gravitational counterpart, Equation 13.19.

Only energies satisfying this equation are permitted. The lowest allowed energy level, the ground state, has $n = 1$ and energy $E_1 = -13.606$ eV. The next energy level, the first excited state, has $n = 2$ and energy $E_2 = E_1/2^2 = -3.401$ eV. Figure 42.8 is an energy-level diagram showing the energies of these discrete energy states and the corresponding quantum numbers n. The uppermost level corresponds to $n = \infty$ (or $r = \infty$) and $E = 0$.

Notice how the allowed energies of the hydrogen atom differ from those of the particle in a box. The particle-in-a-box energies (Eq. 41.14) increase as n^2, so they become farther apart in energy as n increases. On the other hand, the energies of the hydrogen atom (Eq. 42.14) are inversely proportional to n^2, so their separation in energy becomes smaller as n increases. The separation between energy levels approaches zero as n approaches infinity and the energy approaches zero.

Zero energy represents the boundary between a bound system of an electron and a proton and an unbound system. If the energy of the atom is raised from that of the ground state to any energy larger than zero, the atom is **ionized.** The minimum energy required to ionize the atom in its ground state is called the **ionization energy.** As can be seen from Figure 42.8, the ionization energy for hydrogen in the ground state, based on Bohr's calculation, is 13.6 eV. This finding constituted another major achievement for the Bohr theory because the ionization energy for hydrogen had already been measured to be 13.6 eV.

Equations 42.5 and 42.13 can be used to calculate the frequency of the photon emitted when the electron makes a transition from an outer orbit to an inner orbit:

$$f = \frac{E_i - E_f}{h} = \frac{k_e e^2}{2a_0 h}\left(\frac{1}{n_f^{\,2}} - \frac{1}{n_i^2}\right) \tag{42.15}$$

Because the quantity measured experimentally is wavelength, it is convenient to use $c = f\lambda$ to express Equation 42.15 in terms of wavelength:

$$\frac{1}{\lambda} = \frac{f}{c} = \frac{k_e e^2}{2a_0 hc}\left(\frac{1}{n_f^{\,2}} - \frac{1}{n_i^2}\right) \tag{42.16}$$

Remarkably, this expression, which is purely theoretical, is *identical* to the general form of the empirical relationships discovered by Balmer and Rydberg and given by Equations 42.1 to 42.4:

$$\frac{1}{\lambda} = R_{\mathrm{H}}\left(\frac{1}{n_f^{\,2}} - \frac{1}{n_i^2}\right) \tag{42.17}$$

provided the constant $k_e e^2/2a_0 hc$ is equal to the experimentally determined Rydberg constant. Soon after Bohr demonstrated that these two quantities agree to within approximately 1%, this work was recognized as the crowning achievement of his new quantum theory of the hydrogen atom. Furthermore, Bohr showed that all the spectral series for hydrogen have a natural interpretation in his theory. The different series correspond to transitions to different final states characterized by the quantum number n_f. Figure 42.8 shows the origin of these spectral series as transitions between energy levels.

Bohr extended his model for hydrogen to other elements in which all but one electron had been removed. These systems have the same structure as the hydrogen atom except that the nuclear charge is larger. Ionized elements such as He^+, Li^{2+}, and Be^{3+} were suspected to exist in hot stellar atmospheres, where atomic collisions frequently have enough energy to completely remove one or more atomic electrons. Bohr showed that many mysterious lines observed in the spectra of the Sun and several other stars could not be due to hydrogen but were correctly predicted by his theory if attributed to singly ionized helium. In general, the number of protons in the nucleus of an atom is called the **atomic number** of the element

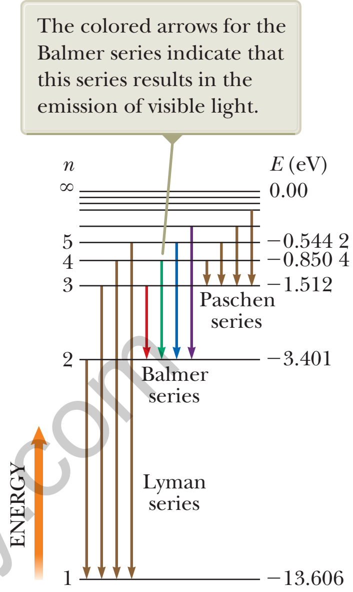


Figure 42.8 An energy-level diagram for the hydrogen atom. Quantum numbers are given on the left, and energies (in electron volts) are given on the right. Vertical arrows represent the four lowest-energy transitions for each of the spectral series shown.

Pitfall Prevention 42.2

The Bohr Model Is Great, but . . . The Bohr model correctly predicts the ionization energy and general features of the spectrum for hydrogen, but it cannot account for the spectra of more complex atoms and is unable to predict many subtle spectral details of hydrogen and other simple atoms. Scattering experiments show that the electron in a hydrogen atom does not move in a flat circle around the nucleus. Instead, the atom is spherical. The ground-state angular momentum of the atom is zero and not $\hbar$.

and is given the symbol Z. To describe a single electron orbiting a fixed nucleus of charge $+Ze$, Bohr's theory gives

$$r_n = (n^2)\frac{a_0}{Z} \tag{42.18}$$

$$E_n = -\frac{k_e e^2}{2a_0}\left(\frac{Z^2}{n^2}\right) \quad n = 1, 2, 3, \ldots \tag{42.19}$$

Although the Bohr theory was triumphant in its agreement with some experimental results on the hydrogen atom, it suffered from some difficulties. One of the first indications that the Bohr theory needed to be modified arose when improved spectroscopic techniques were used to examine the spectral lines of hydrogen. It was found that many of the lines in the Balmer and other series were not single lines at all. Instead, each was a group of lines spaced very close together. An additional difficulty arose when it was observed that in some situations certain single spectral lines were split into three closely spaced lines when the atoms were placed in a strong magnetic field. Efforts to explain these and other deviations from the Bohr model led to modifications in the theory and ultimately to a replacement theory that will be discussed in Section 42.4.

Bohr's Correspondence Principle

In our study of relativity, we found that Newtonian mechanics is a special case of relativistic mechanics and is usable only for speeds much less than c. Similarly,

> quantum physics agrees with classical physics when the difference between quantized levels becomes vanishingly small.

This principle, first set forth by Bohr, is called the **correspondence principle.**[2]

For example, consider an electron orbiting the hydrogen atom with $n > 10\,000$. For such large values of n, the energy differences between adjacent levels approach zero; therefore, the levels are nearly continuous. Consequently, the classical model is reasonably accurate in describing the system for large values of n. According to the classical picture, the frequency of the light emitted by the atom is equal to the frequency of revolution of the electron in its orbit about the nucleus. Calculations show that for $n > 10\,000$, this frequency is different from that predicted by quantum mechanics by less than 0.015%.

Quick Quiz 42.1 A hydrogen atom is in its ground state. Incident on the atom is a photon having an energy of 10.5 eV. What is the result? **(a)** The atom is excited to a higher allowed state. **(b)** The atom is ionized. **(c)** The photon passes by the atom without interaction.

Quick Quiz 42.2 A hydrogen atom makes a transition from the $n = 3$ level to the $n = 2$ level. It then makes a transition from the $n = 2$ level to the $n = 1$ level. Which transition results in emission of the longer-wavelength photon? **(a)** the first transition **(b)** the second transition **(c)** neither transition because the wavelengths are the same for both

Example 42.1 Electronic Transitions in Hydrogen

(A) The electron in a hydrogen atom makes a transition from the $n = 2$ energy level to the ground level ($n = 1$). Find the wavelength and frequency of the emitted photon.

[2]In reality, the correspondence principle is the starting point for Bohr's property 2(c) on angular momentum quantization. To see how property 2(c) arises from the correspondence principle, see J. W. Jewett Jr., *Physics Begins with Another M . . . Mysteries, Magic, Myth, and Modern Physics* (Boston: Allyn & Bacon, 1996), pp. 353–356.

42.1 continued

SOLUTION

Conceptualize Imagine the electron in a circular orbit about the nucleus as in the Bohr model in Figure 42.6. When the electron makes a transition to a lower stationary state, it emits a photon with a given frequency and drops to a circular orbit of smaller radius.

Categorize We evaluate the results using equations developed in this section, so we categorize this example as a substitution problem.

Use Equation 42.17 to obtain λ, with $n_i = 2$ and $n_f = 1$:

$$\frac{1}{\lambda} = R_H\left(\frac{1}{1^2} - \frac{1}{2^2}\right) = \frac{3R_H}{4}$$

$$\lambda = \frac{4}{3R_H} = \frac{4}{3(1.097 \times 10^7\ \text{m}^{-1})} = 1.22 \times 10^{-7}\ \text{m} = \boxed{122\ \text{nm}}$$

Use Equation 34.20 to find the frequency of the photon:

$$f = \frac{c}{\lambda} = \frac{3.00 \times 10^8\ \text{m/s}}{1.22 \times 10^{-7}\ \text{m}} = \boxed{2.47 \times 10^{15}\ \text{Hz}}$$

(B) In interstellar space, highly excited hydrogen atoms called Rydberg atoms have been observed. Find the wavelength to which radio astronomers must tune to detect signals from electrons dropping from the $n = 273$ level to the $n = 272$ level.

SOLUTION

Use Equation 42.17, this time with $n_i = 273$ and $n_f = 272$:

$$\frac{1}{\lambda} = R_H\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) = R_H\left(\frac{1}{(272)^2} - \frac{1}{(273)^2}\right) = 9.88 \times 10^{-8}\ R_H$$

Solve for λ:

$$\lambda = \frac{1}{9.88 \times 10^{-8} R_H} = \frac{1}{(9.88 \times 10^{-8})(1.097 \times 10^7\ \text{m}^{-1})} = \boxed{0.922\ \text{m}}$$

(C) What is the radius of the electron orbit for a Rydberg atom for which $n = 273$?

SOLUTION

Use Equation 42.12 to find the radius of the orbit:

$$r_{273} = (273)^2\ (0.052\ 9\ \text{nm}) = \boxed{3.94\ \mu\text{m}}$$

This radius is large enough that the atom is on the verge of becoming macroscopic!

(D) How fast is the electron moving in a Rydberg atom for which $n = 273$?

SOLUTION

Solve Equation 42.8 for the electron's speed:

$$v = \sqrt{\frac{k_e e^2}{m_e r}} = \sqrt{\frac{(8.99 \times 10^9\ \text{N}\cdot\text{m}^2/\text{C}^2)(1.60 \times 10^{-19}\ \text{C})^2}{(9.11 \times 10^{-31}\ \text{kg})(3.94 \times 10^{-6}\ \text{m})}}$$

$$= \boxed{8.01 \times 10^3\ \text{m/s}}$$

WHAT IF? What if radiation from the Rydberg atom in part (B) is treated classically? What is the wavelength of radiation emitted by the atom in the $n = 273$ level?

Answer Classically, the frequency of the emitted radiation is that of the rotation of the electron around the nucleus.

Calculate this frequency using the period defined in Equation 4.15:

$$f = \frac{1}{T} = \frac{v}{2\pi r}$$

Substitute the radius and speed from parts (C) and (D):

$$f = \frac{v}{2\pi r} = \frac{8.02 \times 10^3\ \text{m/s}}{2\pi(3.94 \times 10^{-6}\ \text{m})} = 3.24 \times 10^8\ \text{Hz}$$

Find the wavelength of the radiation from Equation 34.20:

$$\lambda = \frac{c}{f} = \frac{3.00 \times 10^8\ \text{m/s}}{3.24 \times 10^8\ \text{Hz}} = 0.927\ \text{m}$$

This value is about 0.5% different from the wavelength calculated in part (B). As indicated in the discussion of Bohr's correspondence principle, this difference becomes even smaller for higher values of n.

42.4 The Quantum Model of the Hydrogen Atom

In the preceding section, we described how the Bohr model views the electron as a particle orbiting the nucleus in nonradiating, quantized energy levels. This model combines both classical and quantum concepts. Although the model demonstrates excellent agreement with some experimental results, it cannot explain others. These difficulties are removed when a full quantum model involving the Schrödinger equation is used to describe the hydrogen atom.

The formal procedure for solving the problem of the hydrogen atom is to substitute the appropriate potential energy function into the Schrödinger equation, find solutions to the equation, and apply boundary conditions as we did for the particle in a box in Chapter 41. The potential energy function for the hydrogen atom is that due to the electrical interaction between the electron and the proton (see Section 25.3):

$$U(r) = -k_e \frac{e^2}{r} \tag{42.20}$$

where k_e is the Coulomb constant and r is the radial distance from the proton (situated at $r = 0$) to the electron.

The mathematics for the hydrogen atom is more complicated than that for the particle in a box for two primary reasons: (1) the atom is three-dimensional, and (2) U is not constant, but rather depends on the radial coordinate r. If the time-independent Schrödinger equation (Eq. 41.15) is extended to three-dimensional rectangular coordinates, the result is

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right) + U\psi = E\psi$$

It is easier to solve this equation for the hydrogen atom if rectangular coordinates are converted to spherical polar coordinates, an extension of the plane polar coordinates introduced in Section 3.1. In spherical polar coordinates, a point in space is represented by the three variables r, θ, and ϕ, where r is the radial distance from the origin, $r = \sqrt{x^2 + y^2 + z^2}$. With the point represented at the end of a position vector $\vec{\mathbf{r}}$ as shown in Figure 42.9, the angular coordinate θ specifies its angular position relative to the z axis. Once that position vector is projected onto the xy plane, the angular coordinate ϕ specifies the projection's (and therefore the point's) angular position relative to the x axis.

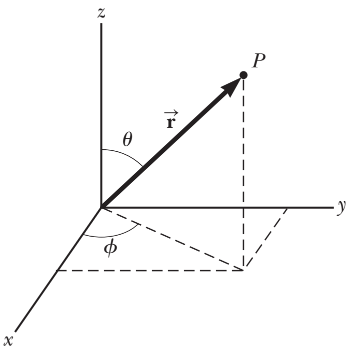


Figure 42.9 A point P in space is located by means of a position vector $\vec{\mathbf{r}}$. In Cartesian coordinates, the components of this vector are x, y, and z. In spherical polar coordinates, the point is described by r, the distance from the origin; θ, the angle between $\vec{\mathbf{r}}$ and the z axis; and ϕ, the angle between the x axis and a projection of $\vec{\mathbf{r}}$ onto the xy plane.

The conversion of the three-dimensional time-independent Schrödinger equation for $\psi(x, y, z)$ to the equivalent form for $\psi(r, \theta, \phi)$ is straightforward but very tedious, so we omit the details.[3] In Chapter 41, we separated the time dependence from the space dependence in the general wave function Ψ. In this case of the hydrogen atom, the three space variables in $\psi(r, \theta, \phi)$ can be similarly separated by writing the wave function as a product of functions of each single variable:

$$\psi(r, \theta, \phi) = R(r)f(\theta)g(\phi)$$

In this way, Schrödinger's equation, which is a three-dimensional partial differential equation, can be transformed into three separate ordinary differential equations: one for $R(r)$, one for $f(\theta)$, and one for $g(\phi)$. Each of these functions is subject to boundary conditions. For example, $R(r)$ must remain finite as $r \to 0$ and $r \to \infty$; furthermore, $g(\phi)$ must have the same value as $g(\phi + 2\pi)$.

The potential energy function given in Equation 42.20 depends *only* on the radial coordinate r and not on either of the angular coordinates; therefore, it appears only in the equation for $R(r)$. As a result, the equations for θ and ϕ are independent of the particular system and their solutions are valid for *any* system exhibiting rotation.

When the full set of boundary conditions is applied to all three functions, three different quantum numbers are found for each allowed state of the hydrogen atom,

[3] Descriptions of the solutions to the Schrödinger equation for the hydrogen atom are available in modern physics textbooks such as R. A. Serway, C. Moses, and C. A. Moyer, *Modern Physics,* 3rd ed. (Belmont, CA: Brooks/Cole, 2005).

one for each of the separate differential equations. These quantum numbers are restricted to integer values and correspond to the three independent degrees of freedom (three space dimensions).

The first quantum number, associated with the radial function $R(r)$ of the full wave function, is called the **principal quantum number** and is assigned the symbol n. The differential equation for $R(r)$ leads to functions giving the probability of finding the electron at a certain radial distance from the nucleus. In Section 42.5, we will describe two of these radial wave functions. From the boundary conditions, the energies of the allowed states for the hydrogen atom are found to be related to n as follows:

$$E_n = -\left(\frac{k_e e^2}{2a_0}\right)\frac{1}{n^2} = -\frac{13.606 \text{ eV}}{n^2} \qquad n = 1, 2, 3, \ldots \tag{42.21}$$

◀ Allowed energies of the quantum hydrogen atom

This result is in exact agreement with that obtained in the Bohr theory (Eqs. 42.13 and 42.14)! This agreement is *remarkable* because the Bohr theory and the full quantum theory arrive at the result from completely different starting points.

The **orbital quantum number,** symbolized ℓ, comes from the differential equation for $f(\theta)$ and is associated with the orbital angular momentum of the electron. The **orbital magnetic quantum number** m_ℓ arises from the differential equation for $g(\phi)$. Both ℓ and m_ℓ are integers. We will expand our discussion of these two quantum numbers in Section 42.6, where we also introduce a fourth (nonintegral) quantum number, resulting from a relativistic treatment of the hydrogen atom.

The application of boundary conditions on the three parts of the full wave function leads to important relationships among the three quantum numbers as well as certain restrictions on their values:

The values of n are integers that can range from 1 to ∞.

The values of ℓ are integers that can range from 0 to $n - 1$.

The values of m_ℓ are integers that can range from $-\ell$ to ℓ.

◀ Restrictions on the values of hydrogen-atom quantum numbers

For example, if $n = 1$, only $\ell = 0$ and $m_\ell = 0$ are permitted. If $n = 2$, then ℓ may be 0 or 1; if $\ell = 0$, then $m_\ell = 0$; but if $\ell = 1$, then m_ℓ may be 1, 0, or -1. Table 42.1 summarizes the rules for determining the allowed values of ℓ and m_ℓ for a given n.

For historical reasons, all states having the same principal quantum number are said to form a **shell.** Shells are identified by the letters K, L, M, . . . , which designate the states for which $n = 1, 2, 3, \ldots$. Likewise, all states having the same values of n and ℓ are said to form a **subshell.** The letters[4] *s, p, d, f, g, h,* . . . are used to designate the subshells for which $\ell = 0, 1, 2, 3, \ldots$. The state designated by $3p$, for example, has the quantum numbers $n = 3$ and $\ell = 1$; the $2s$ state has the quantum numbers $n = 2$ and $\ell = 0$. These notations are summarized in Tables 42.2 and 42.3 (page 1308).

States that violate the rules given in Table 42.1 do not exist. (They do not satisfy the boundary conditions on the wave function.) For instance, the $2d$ state, which

Pitfall Prevention 42.3

Energy Depends on *n* Only for Hydrogen The implication in Equation 42.21 that the energy depends only on the quantum number n is true only for the hydrogen atom. For more complicated atoms, we will use the same quantum numbers developed here for hydrogen. The energy levels for these atoms depend primarily on n, but they also depend to a lesser degree on other quantum numbers.

Pitfall Prevention 42.4

Quantum Numbers Describe a System It is common to assign the quantum numbers to an electron. Remember, however, that these quantum numbers arise from the Schrödinger equation, which involves a potential energy function for the *system* of the electron and the nucleus. Therefore, it is more *proper* to assign the quantum numbers to the atom, but it is more *popular* to assign them to an electron. We follow this latter usage because it is so common.

Table 42.1 Three Quantum Numbers for the Hydrogen Atom

Quantum Number	Name	Allowed Values	Number of Allowed States
n	Principal quantum number	$1, 2, 3, \ldots$	Any number
ℓ	Orbital quantum number	$0, 1, 2, \ldots, n - 1$	n
m_ℓ	Orbital magnetic quantum number	$-\ell, -\ell + 1, \ldots, 0, \ldots, \ell - 1, \ell$	$2\ell + 1$

Table 42.2 Atomic Shell Notations

n	Shell Symbol
1	K
2	L
3	M
4	N
5	O
6	P

[4]The first four of these letters come from early classifications of spectral lines: sharp, principal, diffuse, and fundamental. The remaining letters are in alphabetical order.

Table 42.3 Atomic Subshell Notations

ℓ	Subshell Symbol
0	s
1	p
2	d
3	f
4	g
5	h

would have $n = 2$ and $\ell = 2$, cannot exist because the highest allowed value of ℓ is $n - 1$, which in this case is 1. Therefore, for $n = 2$, the $2s$ and $2p$ states are allowed but $2d$, $2f$, . . . are not. For $n = 3$, the allowed subshells are $3s$, $3p$, and $3d$.

Quick Quiz 42.3 How many possible subshells are there for the $n = 4$ level of hydrogen? **(a)** 5 **(b)** 4 **(c)** 3 **(d)** 2 **(e)** 1

Quick Quiz 42.4 When the principal quantum number is $n = 5$, how many different values of **(a)** ℓ and **(b)** m_ℓ are possible?

Example 42.2 The $n = 2$ Level of Hydrogen

For a hydrogen atom, determine the allowed states corresponding to the principal quantum number $n = 2$ and calculate the energies of these states.

SOLUTION

Conceptualize Think about the atom in the $n = 2$ quantum state. There is only one such state in the Bohr theory, but our discussion of the quantum theory allows for more states because of the possible values of ℓ and m_ℓ.

Categorize We evaluate the results using rules discussed in this section, so we categorize this example as a substitution problem.

From Table 42.1, we find that when $n = 2$, ℓ can be 0 or 1. Find the possible values of m_ℓ from Table 42.1:

$$\ell = 0 \rightarrow m_\ell = 0$$
$$\ell = 1 \rightarrow m_\ell = -1, 0, \text{ or } 1$$

Hence, we have one state, designated as the $2s$ state, that is associated with the quantum numbers $n = 2$, $\ell = 0$, and $m_\ell = 0$, and we have three states, designated as $2p$ states, for which the quantum numbers are $n = 2$, $\ell = 1$, and $m_\ell = -1$; $n = 2$, $\ell = 1$, and $m_\ell = 0$; and $n = 2$, $\ell = 1$, and $m_\ell = 1$.

Find the energy for all four of these states with $n = 2$ from Equation 42.21:

$$E_2 = -\frac{13.606 \text{ eV}}{2^2} = \boxed{-3.401 \text{ eV}}$$

42.5 The Wave Functions for Hydrogen

Because the potential energy of the hydrogen atom depends only on the radial distance r between nucleus and electron, some of the allowed states for this atom can be represented by wave functions that depend only on r. For these states, $f(\theta)$ and $g(\phi)$ are constants. The simplest wave function for hydrogen is the one that describes the $1s$ state and is designated $\psi_{1s}(r)$:

Wave function for hydrogen in its ground state

$$\psi_{1s}(r) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} \quad (42.22)$$

where a_0 is the Bohr radius. (In Problem 26, you can verify that this function satisfies the Schrödinger equation.) Note that ψ_{1s} approaches zero as r approaches ∞ and is normalized as presented (see Eq. 41.7). Furthermore, because ψ_{1s} depends only on r, it is *spherically symmetric*. This symmetry exists for all s states.

Recall that the probability of finding a particle in any region is equal to an integral of the probability density $|\psi|^2$ for the particle over the region. The probability density for the $1s$ state is

$$|\psi_{1s}|^2 = \left(\frac{1}{\pi a_0^3}\right) e^{-2r/a_0} \quad (42.23)$$

Because we imagine the nucleus to be fixed in space at $r = 0$, we can assign this probability density to the question of locating the electron. According to Equation 41.3, the probability of finding the electron in a volume element dV is $|\psi|^2\,dV$. It is convenient to define the *radial probability density function* $P(r)$ as the probability per unit radial length of finding the electron in a spherical shell of radius r and thickness dr. Therefore, $P(r)\;dr$ is the probability of finding the electron in this shell. The volume dV of such an infinitesimally thin shell equals its surface area $4\pi r^2$ multiplied by the shell thickness dr (Fig. 42.10), so we can write this probability as

$$P(r)\;dr = |\psi|^2\,dV = |\psi|^2\,4\pi r^2\,dr$$

Therefore, the radial probability density function for an s state is

$$P(r) = 4\pi r^2|\psi|^2 \tag{42.24}$$

Substituting Equation 42.23 into Equation 42.24 gives the radial probability density function for the hydrogen atom in its ground state:

$$P_{1s}(r) = \left(\frac{4r^2}{a_0^{\,3}}\right)e^{-2r/a_0} \tag{42.25}$$

Radial probability density for the $1s$ state of hydrogen

A plot of the function $P_{1s}(r)$ versus r is presented in Figure 42.11a. The peak of the curve corresponds to the most probable value of r for this particular state. We show in Example 42.3 that this peak occurs at the Bohr radius, the radial position of the electron when the hydrogen atom is in its ground state in the Bohr theory, another remarkable agreement between the Bohr theory and the quantum theory.

According to quantum mechanics, the atom has no sharply defined boundary as suggested by the Bohr theory. The probability distribution in Figure 42.11a suggests that the charge of the electron can be modeled as being extended throughout a region of space, commonly referred to as an *electron cloud*. Figure 42.11b shows the probability density of the electron in a hydrogen atom in the $1s$ state as a function of position in the xy plane. The darkness of the blue color corresponds to the value of the probability density. The darkest portion of the distribution appears at $r = a_0$, corresponding to the most probable value of r for the electron.

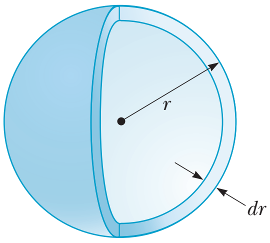


Figure 42.10 A spherical shell of radius r and infinitesimal thickness dr has a volume equal to $4\pi r^2\,dr$.

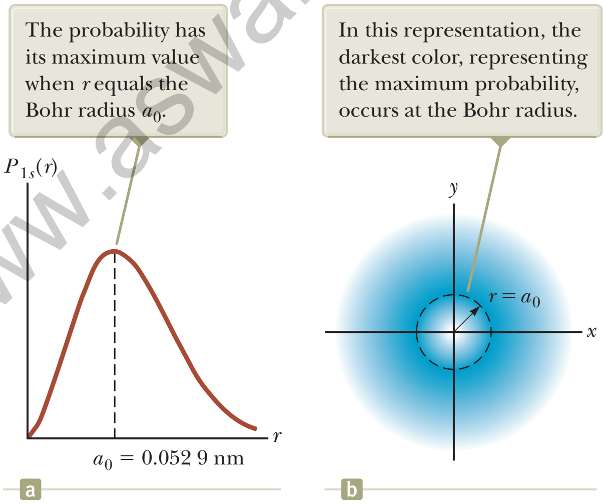


Figure 42.11 (a) The probability of finding the electron as a function of distance from the nucleus for the hydrogen atom in the $1s$ (ground) state. (b) The cross section in the xy plane of the spherical electronic charge distribution for the hydrogen atom in its $1s$ state.

Example 42.3 The Ground State of Hydrogen

(A) Calculate the most probable value of r for an electron in the ground state of the hydrogen atom.

continued

42.3 continued

SOLUTION

Conceptualize Do not imagine the electron in orbit around the proton as in the Bohr theory of the hydrogen atom. Instead, imagine the charge of the electron spread out in space around the proton in an electron cloud with spherical symmetry.

Categorize Because the statement of the problem asks for the "most probable value of *r*," we categorize this example as a problem in which the quantum approach is used. (In the Bohr atom, the electron moves in an orbit with an *exact* value of *r*.)

Analyze The most probable value of r corresponds to the maximum in the plot of $P_{1s}(r)$ versus r. We can evaluate the most probable value of r by setting $dP_{1s}/dr = 0$ and solving for r.

Differentiate Equation 42.25 and set the result equal to zero:

$$\frac{dP_{1s}}{dr} = \frac{d}{dr}\left[\left(\frac{4r^2}{a_0^3}\right)e^{-2r/a_0}\right] = 0$$

$$e^{-2r/a_0}\frac{d}{dr}(r^2) + r^2\frac{d}{dr}(e^{-2r/a_0}) = 0$$

$$2re^{-2r/a_0} + r^2(-2/a_0)e^{-2r/a_0} = 0$$

$$(1) \quad 2r[1 - (r/a_0)]e^{-2r/a_0} = 0$$

Set the bracketed expression equal to zero and solve for r:

$$1 - \frac{r}{a_0} = 0 \quad \rightarrow \quad r = \boxed{a_0}$$

Finalize The most probable value of r is the Bohr radius! Equation (1) is also satisfied at $r = 0$ and as $r \rightarrow \infty$. These points are locations of the *minimum* probability, which is equal to zero as seen in Figure 42.11a.

(B) Calculate the probability that the electron in the ground state of hydrogen will be found outside the Bohr radius.

SOLUTION

Analyze The probability is found by integrating the radial probability density function $P_{1s}(r)$ for this state from the Bohr radius a_0 to ∞.

Set up this integral using Equation 42.25:

$$P = \int_{a_0}^{\infty} P_{1s}(r)\,dr = \frac{4}{a_0^3}\int_{a_0}^{\infty} r^2 e^{-2r/a_0}\,dr$$

Put the integral in dimensionless form by changing variables from r to $z = 2r/a_0$, noting that $z = 2$ when $r = a_0$ and that $dr = (a_0/2)\,dz$:

$$P = \frac{4}{a_0^3}\int_2^{\infty}\left(\frac{za_0}{2}\right)^2 e^{-z}\left(\frac{a_0}{2}\right)dz = \frac{1}{2}\int_2^{\infty} z^2 e^{-z}\,dz$$

Evaluate the integral using partial integration (see Appendix B.7):

$$P = -\tfrac{1}{2}(z^2 + 2z + 2)e^{-z}\Big|_2^{\infty}$$

Evaluate between the limits:

$$P = 0 - [-\tfrac{1}{2}(4 + 4 + 2)e^{-2}] = 5e^{-2} = \boxed{0.677 \text{ or } 67.7\%}$$

Finalize This probability is larger than 50%. The reason for this value is the asymmetry in the radial probability density function (Fig. 42.11a), which has more area to the right of the peak than to the left.

WHAT IF? What if you were asked for the *average* value of r for the electron in the ground state rather than the most probable value?

Answer The average value of r is the same as the expectation value for r.

Use Equation 42.25 to evaluate the average value of r:

$$r_{\text{avg}} = \langle r\rangle = \int_0^{\infty} rP(r)\,dr = \int_0^{\infty} r\left(\frac{4r^2}{a_0^3}\right)e^{-2r/a_0}\,dr$$

$$= \left(\frac{4}{a_0^3}\right)\int_0^{\infty} r^3 e^{-2r/a_0}\,dr$$

42.3 continued

Evaluate the integral with the help of the first integral listed in Table B.6 in Appendix B:

$$r_{\text{avg}} = \left(\frac{4}{{a_0}^3}\right)\left(\frac{3!}{(2/a_0)^4}\right) = \tfrac{3}{2}a_0$$

Again, the average value is larger than the most probable value because of the asymmetry in the wave function as seen in Figure 42.11a.

The next-simplest wave function for the hydrogen atom is the one corresponding to the 2*s* state ($n = 2$, $\ell = 0$). The normalized wave function for this state is

$$\psi_{2s}(r) = \frac{1}{4\sqrt{2\pi}}\left(\frac{1}{a_0}\right)^{3/2}\left(2 - \frac{r}{a_0}\right)e^{-r/2a_0} \qquad \textbf{(42.26)}$$

◀ **Wave function for hydrogen in the 2*s* state**

Again notice that ψ_{2s} depends only on r and is spherically symmetric. The energy corresponding to this state is $E_2 = -(13.606/4)$ eV $= -3.401$ eV. This energy level represents the first excited state of hydrogen. A plot of the radial probability density function for this state in comparison to the 1*s* state is shown in Figure 42.12. The plot for the 2*s* state has two peaks. In this case, the most probable value corresponds to that value of r that has the highest value of P ($\approx 5a_0$). An electron in the 2*s* state would be much farther from the nucleus (on the average) than an electron in the 1*s* state.

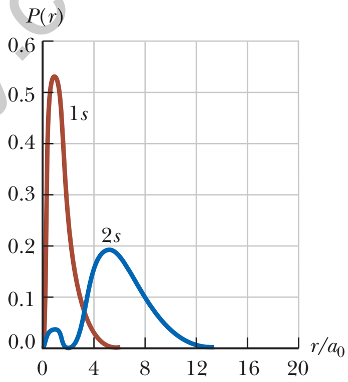


Figure 42.12 The radial probability density function versus r/a_0 for the 1*s* and 2*s* states of the hydrogen atom.

42.6 Physical Interpretation of the Quantum Numbers

The principal quantum number n of a particular state in the hydrogen atom determines the energy of the atom according to Equation 42.21. Now let's see what the other quantum numbers in our atomic model correspond to physically.

The Orbital Quantum Number ℓ

We begin this discussion by returning briefly to the Bohr model of the atom. If the electron moves in a circle of radius r, the magnitude of its angular momentum relative to the center of the circle is $L = m_e vr$. The direction of $\vec{\mathbf{L}}$ is perpendicular to the plane of the circle and is given by a right-hand rule. According to classical physics, the magnitude L of the orbital angular momentum can have any value. The Bohr model of hydrogen, however, postulates that the magnitude of the angular momentum of the electron is restricted to multiples of $\hbar$; that is, $L = n\hbar$. This model must be modified because it predicts (incorrectly) that the ground state of hydrogen has one unit of angular momentum. Furthermore, if L is taken to be zero in the Bohr model, the electron must be pictured as a particle oscillating along a straight line through the nucleus, which is a physically unacceptable situation.

These difficulties are resolved with the quantum-mechanical model of the atom, although we must give up the convenient mental representation of an electron orbiting in a well-defined circular path. Despite the absence of this representation, the atom does indeed possess an angular momentum and it is still called orbital angular momentum. According to quantum mechanics, an atom in a state whose principal quantum number is n can take on the following *discrete* values of the magnitude of the orbital angular momentum:[5]

$$L = \sqrt{\ell(\ell + 1)}\,\hbar \qquad \ell = 0, 1, 2, \ldots, n - 1 \qquad \textbf{(42.27)}$$

◀ **Allowed values of *L***

[5]Equation 42.27 is a direct result of the mathematical solution of the Schrödinger equation and the application of angular boundary conditions. This development, however, is beyond the scope of this book.

Given these allowed values of ℓ, we see that $L = 0$ (corresponding to $\ell = 0$) is an acceptable value of the magnitude of the angular momentum. That L can be zero in this model serves to point out the inherent difficulties in any attempt to describe results based on quantum mechanics in terms of a purely particle-like (classical) model. In the quantum-mechanical interpretation, the electron cloud for the $L = 0$ state is spherically symmetric and has no fundamental rotation axis.

The Orbital Magnetic Quantum Number m_ℓ

Because angular momentum is a vector, its direction must be specified. Recall from Chapter 29 that a current loop has a corresponding magnetic moment $\vec{\boldsymbol{\mu}} = I\vec{\mathbf{A}}$ (Eq. 29.15), where I is the current in the loop and $\vec{\mathbf{A}}$ is a vector perpendicular to the loop whose magnitude is the area of the loop. Such a moment placed in a magnetic field $\vec{\mathbf{B}}$ interacts with the field. Suppose a weak magnetic field applied along the z axis defines a direction in space. According to classical physics, the energy of the loop–field system depends on the direction of the magnetic moment of the loop with respect to the magnetic field as described by Equation 29.18, $U_B = -\vec{\boldsymbol{\mu}} \cdot \vec{\mathbf{B}}$. Any energy between $-\mu B$ and $+\mu B$ is allowed by classical physics.

In the Bohr theory, the circulating electron represents a current loop. In the quantum-mechanical approach to the hydrogen atom, we abandon the circular orbit viewpoint of the Bohr theory, but the atom still possesses an orbital angular momentum. Therefore, there is some sense of rotation of the electron around the nucleus and a magnetic moment is present due to this angular momentum.

As mentioned in Section 42.3, spectral lines from some atoms are observed to split into groups of three closely spaced lines when the atoms are placed in a magnetic field. Suppose the hydrogen atom is located in a magnetic field. According to quantum mechanics, there are *discrete* directions allowed for the magnetic moment vector $\vec{\boldsymbol{\mu}}$ with respect to the magnetic field vector $\vec{\mathbf{B}}$. This situation is very different from that in classical physics, in which all directions are allowed.

Because the magnetic moment $\vec{\boldsymbol{\mu}}$ of the atom can be related[6] to the angular momentum vector $\vec{\mathbf{L}}$, the discrete directions of $\vec{\boldsymbol{\mu}}$ translate to the direction of $\vec{\mathbf{L}}$ being quantized. This quantization means that L_z (the projection of $\vec{\mathbf{L}}$ along the z axis) can have only discrete values. The orbital magnetic quantum number m_ℓ specifies the allowed values of the z component of the orbital angular momentum according to the expression[7]

Allowed values of L_z

$$L_z = m_\ell \hbar \tag{42.28}$$

The quantization of the possible orientations of $\vec{\mathbf{L}}$ with respect to an external magnetic field is often referred to as **space quantization.**

Let's look at the possible magnitudes and orientations of $\vec{\mathbf{L}}$ for a given value of ℓ. Recall that m_ℓ can have values ranging from $-\ell$ to ℓ. If $\ell = 0$, then $L = 0$; the only allowed value of m_ℓ is $m_\ell = 0$ and $L_z = 0$. If $\ell = 1$, then $L = \sqrt{2}\hbar$ from Equation 42.27. The possible values of m_ℓ are -1, 0, and 1, so Equation 42.28 tells us that L_z may be $-\hbar$, 0, or $\hbar$. If $\ell = 2$, the magnitude of the orbital angular momentum is $\sqrt{6}\hbar$. The value of m_ℓ can be -2, -1, 0, 1, or 2, corresponding to L_z values of $-2\hbar$, $-\hbar$, 0, $\hbar$, or $2\hbar$, and so on.

Figure 42.13a shows a **vector model** that describes space quantization for the case $\ell = 2$. Notice that $\vec{\mathbf{L}}$ can never be aligned parallel or antiparallel to $\vec{\mathbf{B}}$ because the maximum value of L_z is $\ell\hbar$, which is less than the magnitude of the angular momentum $L = \sqrt{\ell(\ell + 1)}\,\hbar$. The angular momentum vector $\vec{\mathbf{L}}$ is allowed to be perpendicular to $\vec{\mathbf{B}}$, which corresponds to the case of $L_z = 0$ and $\ell = 0$.

[6]See Equation 30.22 for this relationship as derived from a classical viewpoint. Quantum mechanics arrives at the same result.

[7]As with Equation 42.27, the relationship expressed in Equation 42.28 arises from the solution to the Schrödinger equation and application of boundary conditions.

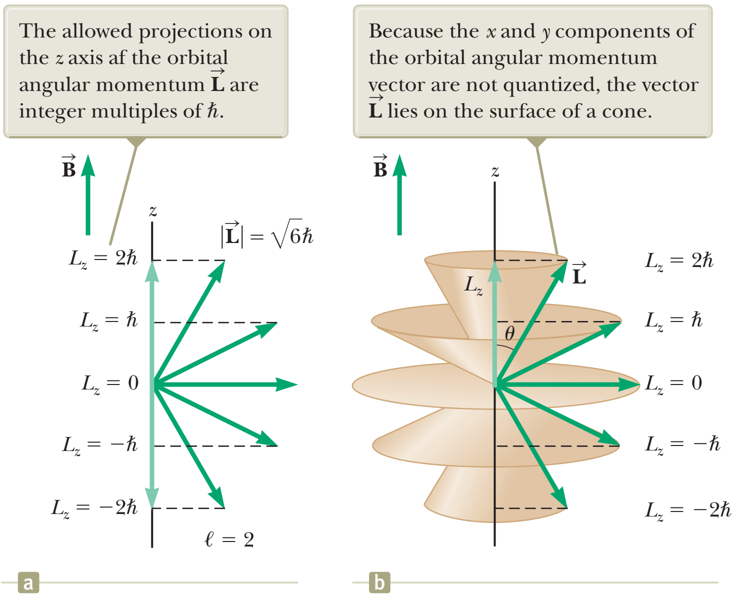


Figure 42.13 A vector model for $\ell = 2$.

The vector $\vec{\mathbf{L}}$ does not point in one specific direction. If $\vec{\mathbf{L}}$ were known exactly, all three components L_x, L_y, and L_z would be specified, which is inconsistent with an angular momentum version of the uncertainty principle. How can the magnitude and z component of a vector be specified, but the vector not be completely specified? To answer, imagine that L_x and L_y are completely unspecified so that $\vec{\mathbf{L}}$ lies anywhere on the surface of a cone that makes an angle θ with the z axis as shown in Figure 42.13b. From the figure, we see that θ is also quantized and that its values are specified through the relationship

$$\cos\theta = \frac{L_z}{L} = \frac{m_\ell}{\sqrt{\ell(\ell+1)}} \qquad (42.29)$$

◀ **Allowed directions of the orbital angular momentum vector**

If the atom is placed in a magnetic field, the energy $U_B = -\vec{\boldsymbol{\mu}} \cdot \vec{\mathbf{B}}$ is additional energy for the atom–field system beyond that described in Equation 42.21. Because the directions of $\vec{\boldsymbol{\mu}}$ are quantized, there are discrete total energies for the system corresponding to different values of m_ℓ. Figure 42.14a shows a transition between two atomic levels in the absence of a magnetic field. In Figure 42.14b, a magnetic

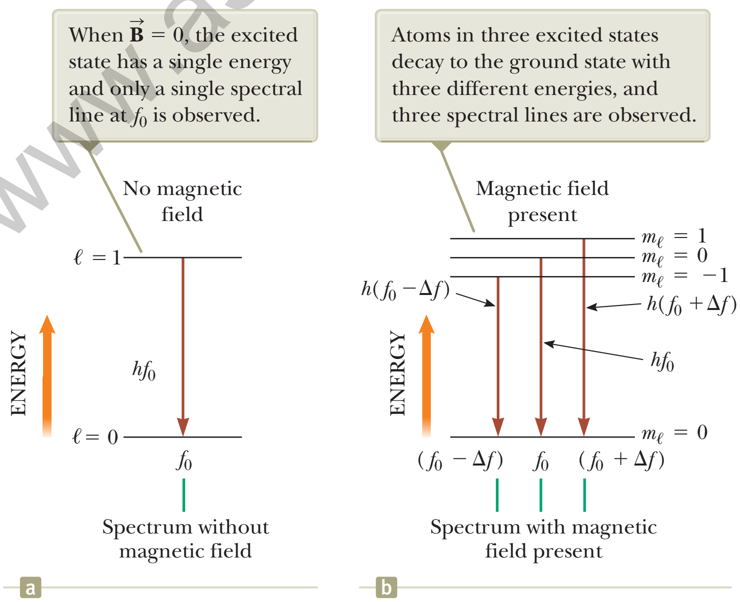


Figure 42.14 The Zeeman effect. (a) Energy levels for the ground and first excited states of a hydrogen atom. (b) When the atom is immersed in a magnetic field $\vec{\mathbf{B}}$, the state with $\ell = 1$ splits into three states, giving rise to emission lines at f_0, $f_0 + \Delta f$, and $f_0 - \Delta f$, where Δf is the frequency shift of the emission caused by the magnetic field.

field is applied and the upper level, with $\ell = 1$, splits into three levels corresponding to the different directions of $\vec{\boldsymbol{\mu}}$. There are now three possible transitions from the $\ell = 1$ subshell to the $\ell = 0$ subshell. Therefore, in a collection of atoms, there are atoms in all three states and the single spectral line in Figure 42.14a splits into three spectral lines. This phenomenon is called the *Zeeman effect.*

The Zeeman effect can be used to measure extraterrestrial magnetic fields. For example, the splitting of spectral lines in light from hydrogen atoms in the surface of the Sun can be used to calculate the magnitude of the magnetic field at that location. The Zeeman effect is one of many phenomena that cannot be explained with the Bohr model but are successfully explained by the quantum model of the atom.

Example 42.4 Space Quantization for Hydrogen

Consider the hydrogen atom in the $\ell = 3$ state. Calculate the magnitude of $\vec{\mathbf{L}}$, the allowed values of L_z, and the corresponding angles θ that $\vec{\mathbf{L}}$ makes with the z axis.

SOLUTION

Conceptualize Consider Figure 42.13a, which is a vector model for $\ell = 2$. Draw such a vector model for $\ell = 3$ to help with this problem.

Categorize We evaluate results using equations developed in this section, so we categorize this example as a substitution problem.

Calculate the magnitude of the orbital angular momentum using Equation 42.27:

$$L = \sqrt{\ell(\ell + 1)}\,\hbar = \sqrt{3(3 + 1)}\,\hbar = \boxed{2\sqrt{3}\,\hbar}$$

Calculate the allowed values of L_z using Equation 42.28 with $m_\ell = -3, -2, -1, 0, 1, 2$, and 3:

$$L_z = \boxed{-3\hbar, -2\hbar, -\hbar, 0, \hbar, 2\hbar, 3\hbar}$$

Calculate the allowed values of $\cos\theta$ using Equation 42.29:

$$\cos\theta = \frac{\pm 3}{2\sqrt{3}} = \pm 0.866 \qquad \cos\theta = \frac{\pm 2}{2\sqrt{3}} = \pm 0.577$$

$$\cos\theta = \frac{\pm 1}{2\sqrt{3}} = \pm 0.289 \qquad \cos\theta = \frac{0}{2\sqrt{3}} = 0$$

Find the angles corresponding to these values of $\cos\theta$:

$$\theta = \boxed{30.0^\circ, 54.7^\circ, 73.2^\circ, 90.0^\circ, 107^\circ, 125^\circ, 150^\circ}$$

WHAT IF? What if the value of ℓ is an arbitrary integer? For an arbitrary value of ℓ, how many values of m_ℓ are allowed?

Answer For a given value of ℓ, the values of m_ℓ range from $-\ell$ to $+\ell$ in steps of 1. Therefore, there are 2ℓ nonzero values of m_ℓ (specifically, $\pm 1, \pm 2, \ldots, \pm\ell$). In addition, one more value of $m_\ell = 0$ is possible, for a total of $2\ell + 1$ values of m_ℓ. This result is critical in understanding the results of the Stern–Gerlach experiment described below with regard to spin.

MARGARETHE BOHR COLLECTION/ Photo Researchers, Inc.

Wolfgang Pauli and Niels Bohr watch a spinning top. The spin of the electron is analogous to the spin of the top but is different in many ways.

The Spin Magnetic Quantum Number m_s

The three quantum numbers n, ℓ, and m_ℓ discussed so far are generated by applying boundary conditions to solutions of the Schrödinger equation, and we can assign a physical interpretation to each quantum number. Let's now consider **electron spin,** which does *not* come from the Schrödinger equation.

In Example 42.2, we found four quantum states corresponding to $n = 2$. In reality, however, eight such states occur. The additional four states can be explained by requiring a fourth quantum number for each state, the **spin magnetic quantum number m_s.**

The need for this new quantum number arises because of an unusual feature observed in the spectra of certain gases, such as sodium vapor. Close examination of one prominent line in the emission spectrum of sodium reveals that the line is, in fact, two closely spaced lines called a *doublet*.[8] The wavelengths of these lines occur in the yellow region of the electromagnetic spectrum at 589.0 nm and 589.6 nm. In 1925, when this doublet was first observed, it could not be explained with the existing atomic theory. To resolve this dilemma, Samuel Goudsmit (1902–1978) and George Uhlenbeck (1900–1988), following a suggestion made by Austrian physicist Wolfgang Pauli, proposed the spin quantum number.

To describe this new quantum number, it is convenient (but technically incorrect) to imagine the electron spinning about its axis as it orbits the nucleus as described in Section 30.6. As illustrated in Figure 42.15, only two directions exist for the electron spin. If the direction of spin is as shown in Figure 42.15a, the electron is said to have *spin up*. If the direction of spin is as shown in Figure 42.15b, the electron is said to have *spin down*. In the presence of a magnetic field, the energy associated with the electron is slightly different for the two spin directions. This energy difference accounts for the sodium doublet.

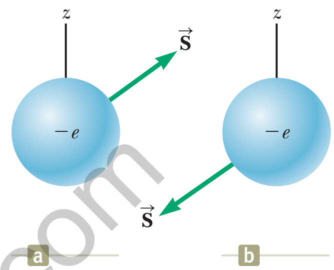


Figure 42.15 The spin of an electron can be either (a) up or (b) down relative to a specified z axis. As in the case of orbital angular momentum, the x and y components of the spin angular momentum vector are not quantized.

The classical description of electron spin—as resulting from a spinning electron—is incorrect. More recent theory indicates that the electron is a point particle, without spatial extent. Therefore, the electron is not modeled as a rigid object and cannot be considered to be spinning. Despite this conceptual difficulty, all experimental evidence supports the idea that an electron does have some intrinsic angular momentum that can be described by the spin magnetic quantum number. Paul Dirac (1902–1984) showed that this fourth quantum number originates in the relativistic properties of the electron.

In 1921, Otto Stern (1888–1969) and Walter Gerlach (1889–1979) performed an experiment that demonstrated space quantization. Their results, however, were not in quantitative agreement with the atomic theory that existed at that time. In their experiment, a beam of silver atoms sent through a nonuniform magnetic field was split into two discrete components (Fig. 42.16). Stern and Gerlach repeated the experiment using other atoms, and in each case the beam split into two or more components. The classical argument is as follows. If the z direction is chosen to be the direction of the maximum nonuniformity of $\vec{\mathbf{B}}$, the net magnetic force on the atoms is along the z axis and is proportional to the component of the magnetic moment $\vec{\boldsymbol{\mu}}$ of the atom in the z direction. Classically, $\vec{\boldsymbol{\mu}}$ can have any orientation, so the deflected beam should be spread out continuously. According to quantum mechanics, however, the deflected beam has an integral number of discrete components and the number of components determines the number of possible values of μ_z. Therefore, because the Stern–Gerlach experiment showed split beams, space quantization was at least qualitatively verified.

Pitfall Prevention 42.5

The Electron Is Not Spinning Although the concept of a spinning electron is conceptually useful, it should not be taken literally. The spin of the Earth is a mechanical rotation. On the other hand, electron spin is a purely quantum effect that gives the electron an angular momentum as if it were physically spinning.

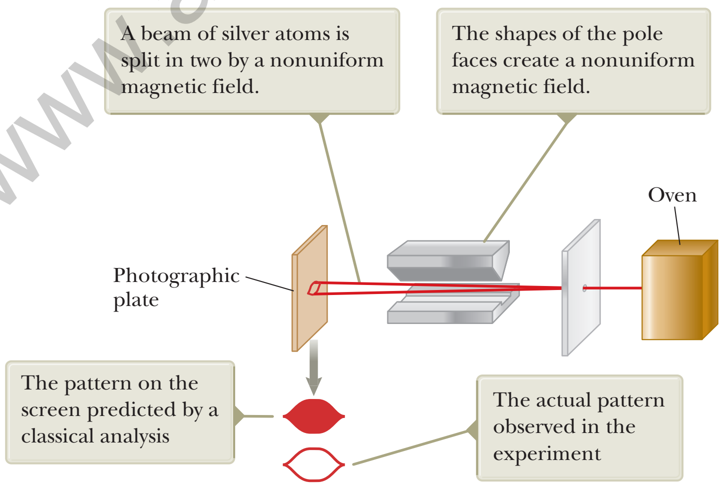


Figure 42.16 The technique used by Stern and Gerlach to verify space quantization.

[8]This phenomenon is a Zeeman effect for spin and is identical in nature to the Zeeman effect for orbital angular momentum discussed before Example 42.4 except that no external magnetic field is required. The magnetic field for this Zeeman effect is internal to the atom and arises from the relative motion of the electron and the nucleus.

For the moment, let's assume the magnetic moment of the atom is due to the orbital angular momentum. Because μ_z is proportional to m_ℓ, the number of possible values of μ_z is $2\ell + 1$ as found in the What If? section of Example 42.4. Furthermore, because ℓ is an integer, the number of values of μ_z is always odd. This prediction is not consistent with Stern and Gerlach's observation of two components (an *even* number) in the deflected beam of silver atoms. Hence, either quantum mechanics is incorrect or the model is in need of refinement.

In 1927, T. E. Phipps and J. B. Taylor repeated the Stern–Gerlach experiment using a beam of hydrogen atoms. Their experiment was important because it involved an atom containing a single electron in its ground state, for which the quantum theory makes reliable predictions. Recall that $\ell = 0$ for hydrogen in its ground state, so $m_\ell = 0$. Therefore, we would not expect the beam to be deflected by the magnetic field at all because the magnetic moment $\vec{\boldsymbol{\mu}}$ of the atom is zero. The beam in the Phipps–Taylor experiment, however, was again split into two components! On the basis of that result, we must conclude that something other than the electron's orbital motion is contributing to the atomic magnetic moment.

As we learned earlier, Goudsmit and Uhlenbeck had proposed that the electron has an intrinsic angular momentum, spin, apart from its orbital angular momentum. In other words, the total angular momentum of the electron in a particular electronic state contains both an orbital contribution $\vec{\mathbf{L}}$ and a spin contribution $\vec{\mathbf{S}}$. The Phipps–Taylor result confirmed the hypothesis of Goudsmit and Uhlenbeck.

In 1929, Dirac used the relativistic form of the total energy of a system to solve the relativistic wave equation for the electron in a potential well. His analysis confirmed the fundamental nature of electron spin. (Spin, like mass and charge, is an *intrinsic* property of a particle, independent of its surroundings.) Furthermore, the analysis showed that electron spin[9] can be described by a single quantum number s, whose value can be only $s = \frac{1}{2}$. The spin angular momentum of the electron *never changes.* This notion contradicts classical laws, which dictate that a rotating charge slows down in the presence of an applied magnetic field because of the Faraday emf that accompanies the changing field (Chapter 31). Furthermore, if the electron is viewed as a spinning ball of charge subject to classical laws, parts of the electron near its surface would be rotating with speeds exceeding the speed of light. Therefore, the classical picture must not be pressed too far; ultimately, spin of an electron is a quantum entity defying any simple classical description.

Because spin is a form of angular momentum, it must follow the same quantum rules as orbital angular momentum. In accordance with Equation 42.27, the magnitude of the **spin angular momentum** $\vec{\mathbf{S}}$ for the electron is

Magnitude of the spin angular momentum of an electron

$$S = \sqrt{s(s+1)}\,\hbar = \frac{\sqrt{3}}{2}\hbar \qquad \textbf{(42.30)}$$

Like orbital angular momentum $\vec{\mathbf{L}}$, spin angular momentum $\vec{\mathbf{S}}$ exhibits space quantization as described in Figure 42.17. The spin vector $\vec{\mathbf{S}}$ can have two orientations relative to a z axis, specified by the **spin magnetic quantum number** $m_s = \pm\frac{1}{2}$. Similar to Equation 42.28 for orbital angular momentum, the z component of spin angular momentum is

Allowed values of S_z

$$S_z = m_s\hbar = \pm\tfrac{1}{2}\hbar \qquad \textbf{(42.31)}$$

The two values $\pm\hbar/2$ for S_z correspond to the two possible orientations for $\vec{\mathbf{S}}$ shown in Figure 42.17. The value $m_s = +\frac{1}{2}$ refers to the spin-up case, and $m_s = -\frac{1}{2}$ refers to the spin-down case. Notice that Equations 42.30 and 42.31 do not allow the spin vector to lie along the z axis. The actual direction of $\vec{\mathbf{S}}$ is at a relatively large angle with respect to the z axis as shown in Figures 42.15 and 42.17.

[9]Scientists often use the word *spin* when referring to the spin angular momentum quantum number. For example, it is common to say, "The electron has a spin of one half."

As discussed in the What If? feature of Example 42.4, there are $2\ell + 1$ possible values of m_ℓ for orbital angular momentum. Similarly, for spin angular momentum, there are $2s + 1$ values of m_s. For a spin of $s = \frac{1}{2}$, the number of values of m_s is $2s + 1 = 2$. These two possibilities for m_s lead to the splitting of the beams into two components in the Stern–Gerlach and Phipps–Taylor experiments.

The spin magnetic moment $\vec{\boldsymbol{\mu}}_{\text{spin}}$ of the electron is related to its spin angular momentum $\vec{\mathbf{S}}$ by the expression

$$\vec{\boldsymbol{\mu}}_{\text{spin}} = -\frac{e}{m_e}\vec{\mathbf{S}} \tag{42.32}$$

where e is the electronic charge and m_e is the mass of the electron. Because $S_z = \pm\frac{1}{2}\hbar$, the z component of the spin magnetic moment can have the values

$$\vec{\boldsymbol{\mu}}_{\text{spin},z} = \pm\frac{e\hbar}{2m_e} \tag{42.33}$$

As we learned in Section 30.6, the quantity $e\hbar/2m_e$ is the Bohr magneton $\mu_B = 9.27 \times 10^{-24}$ J/T. The ratio of magnetic moment to angular momentum is twice as great for spin angular momentum (Eq. 42.32) as it is for orbital angular momentum (Eq. 30.22). The factor of 2 is explained in a relativistic treatment first carried out by Dirac.

Today, physicists explain the Stern–Gerlach and Phipps–Taylor experiments as follows. The observed magnetic moments for both silver and hydrogen are due to spin angular momentum only, with no contribution from orbital angular momentum. In the Phipps–Taylor experiment, the single electron in the hydrogen atom has its electron spin quantized in the magnetic field in such a way that the z component of spin angular momentum is either $\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$, corresponding to $m_s = \pm\frac{1}{2}$. Electrons with spin $+\frac{1}{2}$ are deflected downward, and those with spin $-\frac{1}{2}$ are deflected upward. In the Stern–Gerlach experiment, 46 of a silver atom's 47 electrons are in filled subshells with paired spins. Therefore, these 46 electrons have a net zero contribution to both orbital and spin angular momentum for the atom. The angular momentum of the atom is due to only the 47th electron. This electron lies in the 5s subshell, so there is no contribution from orbital angular momentum. As a result, the silver atoms have angular momentum due to just the spin of one electron and behave in the same way in a nonuniform magnetic field as the hydrogen atoms in the Phipps–Taylor experiment.

The Stern–Gerlach experiment provided two important results. First, it verified the concept of space quantization. Second, it showed that spin angular momentum exists, even though this property was not recognized until four years after the experiments were performed.

As mentioned earlier, there are eight quantum states corresponding to $n = 2$ in the hydrogen atom, not four as found in Example 42.2. Each of the four states in Example 42.2 is actually two states because of the two possible values of m_s. Table 42.4 shows the quantum numbers corresponding to these eight states.

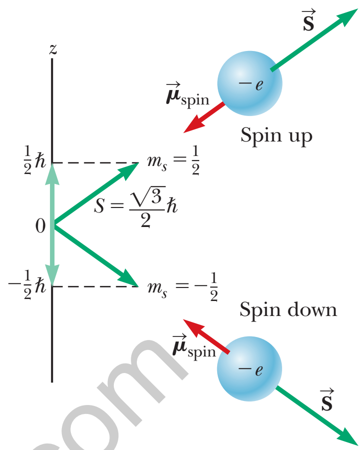


Figure 42.17 Spin angular momentum $\vec{\mathbf{S}}$ exhibits space quantization. This figure shows the two allowed orientations of the spin angular momentum vector $\vec{\mathbf{S}}$ and the spin magnetic moment $\vec{\boldsymbol{\mu}}_{\text{spin}}$ for a spin-$\frac{1}{2}$ particle, such as the electron.

Table 42.4 Quantum Numbers for the $n = 2$ State of Hydrogen

n	ℓ	m_ℓ	m_s	Subshell	Shell	Number of States in Subshell
2	0	0	$\frac{1}{2}$	2s	L	2
2	0	0	$-\frac{1}{2}$			
2	1	1	$\frac{1}{2}$	2p	L	6
2	1	1	$-\frac{1}{2}$			
2	1	0	$\frac{1}{2}$			
2	1	0	$-\frac{1}{2}$			
2	1	−1	$\frac{1}{2}$			
2	1	−1	$-\frac{1}{2}$			

Keystone/Getty Images

Wolfgang Pauli
Austrian Theoretical Physicist (1900–1958)
An extremely talented theoretician who made important contributions in many areas of modern physics, Pauli gained public recognition at the age of 21 with a masterful review article on relativity that is still considered one of the finest and most comprehensive introductions to the subject. His other major contributions were the discovery of the exclusion principle, the explanation of the connection between particle spin and statistics, theories of relativistic quantum electrodynamics, the neutrino hypothesis, and the hypothesis of nuclear spin.

Pitfall Prevention 42.6

The Exclusion Principle Is More General A more general form of the exclusion principle, discussed in Chapter 46, states that no two *fermions* can be in the same quantum state. Fermions are particles with half-integral spin ($\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$, and so on).

42.7 The Exclusion Principle and the Periodic Table

We have found that the state of a hydrogen atom is specified by four quantum numbers: n, ℓ, m_ℓ, and m_s. As it turns out, the number of states available to other atoms may also be predicted by this same set of quantum numbers. In fact, these four quantum numbers can be used to describe all the electronic states of an atom, regardless of the number of electrons in its structure.

For our discussion of atoms with many electrons, it is often easiest to assign the quantum numbers to the electrons in the atom as opposed to the entire atom. An obvious question that arises here is, "How many electrons can be in a particular quantum state?" Pauli answered this important question in 1925, in a statement known as the **exclusion principle:**

> No two electrons can ever be in the same quantum state; therefore, no two electrons in the same atom can have the same set of quantum numbers.

If this principle were not valid, an atom could radiate energy until every electron in the atom is in the lowest possible energy state and therefore the chemical behavior of the elements would be grossly modified. Nature as we know it would not exist.

In reality, we can view the electronic structure of complex atoms as a succession of filled levels increasing in energy. As a general rule, the order of filling of an atom's subshells is as follows. Once a subshell is filled, the next electron goes into the lowest-energy vacant subshell. We can understand this behavior by recognizing that if the atom were not in the lowest energy state available to it, it would radiate energy until it reached this state. This tendency of a quantum system to achieve the lowest energy state is consistent with the second law of thermodynamics discussed in Chapter 22. The entropy of the Universe is increased by the system emitting photons, so that energy is spread out over a larger volume of space.

Before we discuss the electronic configuration of various elements, it is convenient to define an *orbital* as the atomic state characterized by the quantum numbers n, ℓ, and m_ℓ. The exclusion principle tells us that only two electrons can be present in any orbital. One of these electrons has a spin magnetic quantum number $m_s = +\frac{1}{2}$, and the other has $m_s = -\frac{1}{2}$. Because each orbital is limited to two electrons, the number of electrons that can occupy the various shells is also limited.

Table 42.5 shows the allowed quantum states for an atom up to $n = 3$. The arrows pointing upward indicate an electron described by $m_s = +\frac{1}{2}$, and those pointing downward indicate that $m_s = -\frac{1}{2}$. The $n = 1$ shell can accommodate only two electrons because $m_\ell = 0$ means that only one orbital is allowed. (The three quantum numbers describing this orbital are $n = 1$, $\ell = 0$, and $m_\ell = 0$.) The $n = 2$ shell has two subshells, one for $\ell = 0$ and one for $\ell = 1$. The $\ell = 0$ subshell is limited to two electrons because $m_\ell = 0$. The $\ell = 1$ subshell has three allowed orbitals, corresponding to $m_\ell = 1$, 0, and -1. Because each orbital can accommodate two electrons, the $\ell = 1$ subshell can hold six electrons. Therefore, the $n = 2$ shell can contain eight electrons as shown in Table 42.4. The $n = 3$ shell has three subshells ($\ell = 0, 1, 2$) and nine orbitals, accommodating up to 18 electrons. In general, each shell can accommodate up to $2n^2$ electrons.

Table 42.5 **Allowed Quantum States for an Atom Up to $n = 3$**

Shell	n	1	2				3								
Subshell	ℓ	0	0	1			0	1			2				
Orbital	m_ℓ	0	0	1	0	−1	0	1	0	−1	2	1	0	−1	−2
	m_s	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓	↑↓

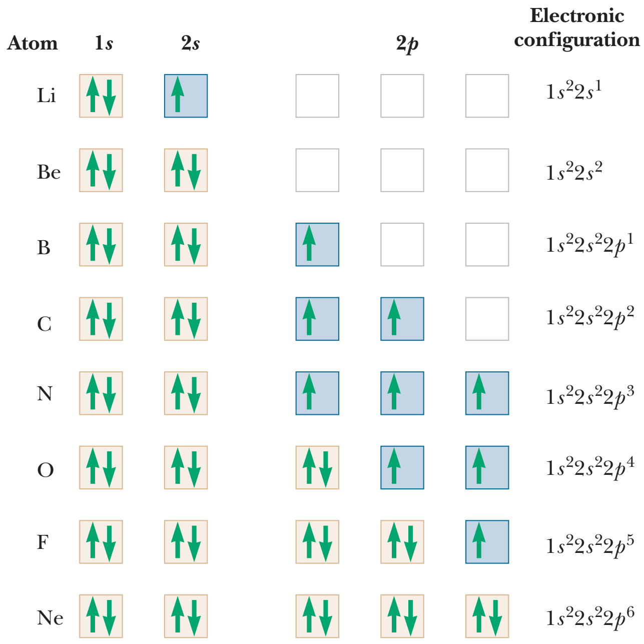


Figure 42.18 The filling of electronic states must obey both the exclusion principle and Hund's rule (page 1320).

The exclusion principle can be illustrated by examining the electronic arrangement in a few of the lighter atoms. The atomic number Z of any element is the number of protons in the nucleus of an atom of that element. A neutral atom of that element has Z electrons. Hydrogen ($Z = 1$) has only one electron, which, in the ground state of the atom, can be described by either of two sets of quantum numbers n, ℓ, m_ℓ, m_s: $1, 0, 0, \frac{1}{2}$ or $1, 0, 0, -\frac{1}{2}$. This electronic configuration is often written $1s^1$. The notation $1s$ refers to a state for which $n = 1$ and $\ell = 0$, and the superscript indicates that one electron is present in the s subshell.

Helium ($Z = 2$) has two electrons. In the ground state, their quantum numbers are $1, 0, 0, \frac{1}{2}$ and $1, 0, 0, -\frac{1}{2}$. No other possible combinations of quantum numbers exist for this level, and we say that the K shell is filled. This electronic configuration is written $1s^2$.

Lithium ($Z = 3$) has three electrons. In the ground state, two of them are in the $1s$ subshell. The third is in the $2s$ subshell because this subshell is slightly lower in energy than the $2p$ subshell.[10] Hence, the electronic configuration for lithium is $1s^22s^1$.

The electronic configurations of lithium and the next several elements are provided in Figure 42.18. The electronic configuration of beryllium ($Z = 4$), with its four electrons, is $1s^22s^2$, and boron ($Z = 5$) has a configuration of $1s^22s^22p^1$. The $2p$ electron in boron may be described by any of the six equally probable sets of quantum numbers listed in Table 42.4. In Figure 42.18, we show this electron in the leftmost $2p$ box with spin up, but it is equally likely to be in any $2p$ box with spin either up or down.

Carbon ($Z = 6$) has six electrons, giving rise to a question concerning how to assign the two $2p$ electrons. Do they go into the same orbital with paired spins ($\uparrow\downarrow$), or do they occupy different orbitals with unpaired spins ($\uparrow\uparrow$)? Experimental data show that the most stable configuration (that is, the one with the lowest energy) is the latter, in which the spins are unpaired. Hence, the two $2p$ electrons in carbon and the three $2p$ electrons in nitrogen ($Z = 7$) have unpaired spins as

[10]To a first approximation, energy depends only on the quantum number n, as we have discussed. Because of the effect of the electronic charge shielding the nuclear charge, however, energy depends on ℓ also in multielectron atoms. We shall discuss these shielding effects in Section 42.8.

Figure 42.18 shows. The general rule that governs such situations, called **Hund's rule,** states that

> when an atom has orbitals of equal energy, the order in which they are filled by electrons is such that a maximum number of electrons have unpaired spins.

Some exceptions to this rule occur in elements having subshells that are close to being filled or half-filled.

In 1871, long before quantum mechanics was developed, the Russian chemist Dmitri Mendeleev (1834–1907) made an early attempt at finding some order among the chemical elements. He was trying to organize the elements for the table of contents of a book he was writing. He arranged the atoms in a table similar to that shown in Figure 42.19, according to their atomic masses and chemical similarities. The first table Mendeleev proposed contained many blank spaces, and he boldly stated that the gaps were there only because the elements had not yet been discovered. By noting the columns in which some missing elements should be located, he was able to make rough predictions about their chemical properties. Within 20 years of this announcement, most of these elements were indeed discovered.

The elements in the **periodic table** (Fig. 42.19) are arranged so that all those in a column have similar chemical properties. For example, consider the elements in the last column, which are all gases at room temperature: He (helium), Ne (neon), Ar (argon), Kr (krypton), Xe (xenon), and Rn (radon). The outstanding characteristic of all these elements is that they do not normally take part in chemical reactions; that is, they do not readily join with other atoms to form molecules. They are therefore called *inert gases* or *noble gases.*

Group I	Group II	Transition elements										Group III	Group IV	Group V	Group VI	Group VII	Group 0
H 1 $1s^1$																H 1 $1s^1$	He 2 $1s^2$
Li 3 $2s^1$	Be 4 $2s^2$											B 5 $2p^1$	C 6 $2p^2$	N 7 $2p^3$	O 8 $2p^4$	F 9 $2p^5$	Ne 10 $2p^6$
Na 11 $3s^1$	Mg 12 $3s^2$											Al 13 $3p^1$	Si 14 $3p^2$	P 15 $3p^3$	S 16 $3p^4$	Cl 17 $3p^5$	Ar 18 $3p^6$
K 19 $4s^1$	Ca 20 $4s^2$	Sc 21 $3d^14s^2$	Ti 22 $3d^24s^2$	V 23 $3d^34s^2$	Cr 24 $3d^54s^1$	Mn 25 $3d^54s^2$	Fe 26 $3d^64s^2$	Co 27 $3d^74s^2$	Ni 28 $3d^84s^2$	Cu 29 $3d^{10}4s^1$	Zn 30 $3d^{10}4s^2$	Ga 31 $4p^1$	Ge 32 $4p^2$	As 33 $4p^3$	Se 34 $4p^4$	Br 35 $4p^5$	Kr 36 $4p^6$
Rb 37 $5s^1$	Sr 38 $5s^2$	Y 39 $4d^15s^2$	Zr 40 $4d^25s^2$	Nb 41 $4d^45s^1$	Mo 42 $4d^55s^1$	Tc 43 $4d^55s^2$	Ru 44 $4d^75s^1$	Rh 45 $4d^85s^1$	Pd 46 $4d^{10}$	Ag 47 $4d^{10}5s^1$	Cd 48 $4d^{10}5s^2$	In 49 $5p^1$	Sn 50 $5p^2$	Sb 51 $5p^3$	Te 52 $5p^4$	I 53 $5p^5$	Xe 54 $5p^6$
Cs 55 $6s^1$	Ba 56 $6s^2$	57–71*	Hf 72 $5d^26s^2$	Ta 73 $5d^36s^2$	W 74 $5d^46s^2$	Re 75 $5d^56s^2$	Os 76 $5d^66s^2$	Ir 77 $5d^76s^2$	Pt 78 $5d^96s^1$	Au 79 $5d^{10}6s^1$	Hg 80 $5d^{10}6s^2$	Tl 81 $6p^1$	Pb 82 $6p^2$	Bi 83 $6p^3$	Po 84 $6p^4$	At 85 $6p^5$	Rn 86 $6p^6$
Fr 87 $7s^1$	Ra 88 $7s^2$	89–103**	Rf 104 $6d^27s^2$	Db 105 $6d^37s^2$	Sg 106 $6d^47s^2$	Bh 107 $6d^57s^2$	Hs 108 $6d^67s^2$	Mt 109 $6d^77s^2$	Ds 110 $6d^97s^1$	Rg 111	Cn 112	113	Fl 114	115	Lv 116	117	118

*Lanthanide series	La 57 $5d^16s^2$	Ce 58 $5d^14f^16s^2$	Pr 59 $4f^36s^2$	Nd 60 $4f^46s^2$	Pm 61 $4f^56s^2$	Sm 62 $4f^66s^2$	Eu 63 $4f^76s^2$	Gd 64 $5d^14f^76s^2$	Tb 65 $5d^14f^86s^2$	Dy 66 $4f^{10}6s^2$	Ho 67 $4f^{11}6s^2$	Er 68 $4f^{12}6s^2$	Tm 69 $4f^{13}6s^2$	Yb 70 $4f^{14}6s^2$	Lu 71 $5d^14f^{14}6s^2$
**Actinide series	Ac 89 $6d^17s^2$	Th 90 $6d^27s^2$	Pa 91 $5f^26d^17s^2$	U 92 $5f^36d^17s^2$	Np 93 $5f^46d^17s^2$	Pu 94 $5f^67s^2$	Am 95 $5f^77s^2$	Cm 96 $5f^76d^17s^2$	Bk 97 $5f^86d^17s^2$	Cf 98 $5f^{10}7s^2$	Es 99 $5f^{11}7s^2$	Fm 100 $5f^{12}7s^2$	Md 101 $5f^{13}7s^2$	No 102 $5f^{14}7s^2$	Lr 103 $5f^{14}6d^17s^2$

Figure 42.19 The periodic table of the elements is an organized tabular representation of the elements that shows their periodic chemical behavior. Elements in a given column have similar chemical behavior. This table shows the chemical symbol for the element, the atomic number, and the electron configuration. A more complete periodic table is available in Appendix C.

We can partially understand this behavior by looking at the electronic configurations in Figure 42.19. The chemical behavior of an element depends on the outermost shell that contains electrons. The electronic configuration for helium is $1s^2$, and the $n = 1$ shell (which is the outermost shell because it is the only shell) is filled. Also, the energy of the atom in this configuration is considerably lower than the energy for the configuration in which an electron is in the next available level, the $2s$ subshell. Next, look at the electronic configuration for neon, $1s^22s^22p^6$. Again, the outermost shell ($n = 2$ in this case) is filled and a wide gap in energy occurs between the filled $2p$ subshell and the next available one, the $3s$ subshell. Argon has the configuration $1s^22s^22p^63s^23p^6$. Here, it is only the $3p$ subshell that is filled, but again a wide gap in energy occurs between the filled $3p$ subshell and the next available one, the $3d$ subshell. This pattern continues through all the noble gases. Krypton has a filled $4p$ subshell, xenon a filled $5p$ subshell, and radon a filled $6p$ subshell.

The column to the left of the noble gases in the periodic table consists of a group of elements called the *halogens:* fluorine, chlorine, bromine, iodine, and astatine. At room temperature, fluorine and chlorine are gases, bromine is a liquid, and iodine and astatine are solids. In each of these atoms, the outer subshell is one electron short of being filled. As a result, the halogens are chemically very active, readily accepting an electron from another atom to form a closed shell. The halogens tend to form strong ionic bonds with atoms at the other side of the periodic table. (We shall discuss ionic bonds in Chapter 43.) In a halogen lightbulb, bromine or iodine atoms combine with tungsten atoms evaporated from the filament and return them to the filament, resulting in a longer-lasting lightbulb. In addition, the filament can be operated at a higher temperature than in ordinary lightbulbs, giving a brighter and whiter light.

At the left side of the periodic table, the Group I elements consist of hydrogen and the *alkali metals:* lithium, sodium, potassium, rubidium, cesium, and francium. Each of these atoms contains one electron in a subshell outside of a closed subshell. Therefore, these elements easily form positive ions because the lone electron is bound with a relatively low energy and is easily removed. Therefore, the alkali metal atoms are chemically active and form very strong bonds with halogen atoms. For example, table salt, NaCl, is a combination of an alkali metal and a halogen. Because the outer electron is weakly bound, pure alkali metals tend to be good electrical conductors. Because of their high chemical activity, however, they are not generally found in nature in pure form.

It is interesting to plot ionization energy versus atomic number Z as in Figure 42.20. Notice the pattern of $\Delta Z = 2, 8, 8, 18, 18, 32$ for the various peaks. This pattern follows from the exclusion principle and helps explain why the elements repeat their chemical properties in groups. For example, the peaks at $Z = 2, 10, 18,$

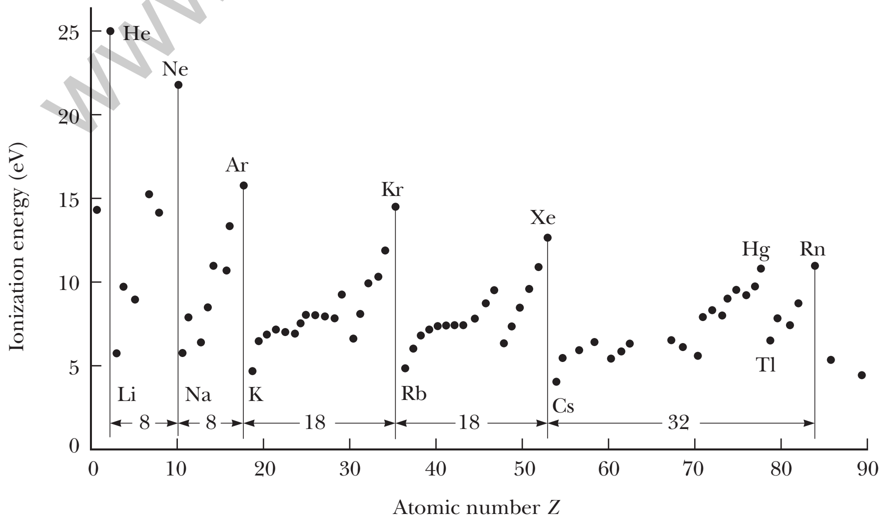


Figure 42.20 Ionization energy of the elements versus atomic number.

and 36 correspond to the noble gases helium, neon, argon, and krypton, respectively, which, as we have mentioned, all have filled outermost shells. These elements have relatively high ionization energies and similar chemical behavior.

42.8 More on Atomic Spectra: Visible and X-Ray

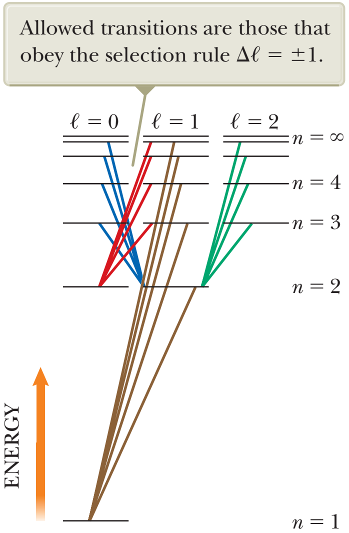


Figure 42.21 Some allowed electronic transitions for hydrogen, represented by the colored lines.

In Section 42.1, we discussed the observation and early interpretation of visible spectral lines from gases. These spectral lines have their origin in transitions between quantized atomic states. We shall investigate these transitions more deeply in these final three sections of this chapter.

A modified energy-level diagram for hydrogen is shown in Figure 42.21. In this diagram, the allowed values of ℓ for each shell are separated horizontally. Figure 42.21 shows only those states up to $\ell = 2$; the shells from $n = 4$ upward would have more sets of states to the right, which are not shown. Transitions for which ℓ does not change are very unlikely to occur and are called *forbidden transitions.* (Such transitions actually can occur, but their probability is very low relative to the probability of "allowed" transitions.) The various diagonal lines represent allowed transitions between stationary states. Whenever an atom makes a transition from a higher energy state to a lower one, a photon of light is emitted. The frequency of this photon is $f = \Delta E/h$, where ΔE is the energy difference between the two states and h is Planck's constant. The **selection rules** for the *allowed transitions* are

Selection rules for allowed atomic transitions

$$\Delta\ell = \pm 1 \quad \text{and} \quad \Delta m_\ell = 0, \pm 1 \tag{42.34}$$

Figure 42.21 shows that the orbital angular momentum of an atom *changes* when it makes a transition to a lower energy state. Therefore, the atom alone is a *nonisolated* system for angular momentum. If we consider the atom–photon system, however, it must be an *isolated* system for angular momentum because nothing else is interacting with this system. The photon involved in the process must carry angular momentum away from the atom when the transition occurs. In fact, the photon has an angular momentum equivalent to that of a particle having a spin of 1. We have now determined over several chapters that a photon has energy, linear momentum, and angular momentum, and each of these is conserved in atomic processes.

Recall from Equation 42.19 that the allowed energies for one-electron atoms and ions, such as hydrogen and He$^+$, are

$$E_n = -\frac{k_e e^2}{2a_0}\left(\frac{Z^2}{n^2}\right) = -\frac{(13.6\text{ eV})Z^2}{n^2} \tag{42.35}$$

This equation was developed from the Bohr theory, but it serves as a good first approximation in quantum theory as well. For multielectron atoms, the positive nuclear charge Ze is largely shielded by the negative charge of the inner-shell electrons. Therefore, the outer electrons interact with a net charge that is smaller than the nuclear charge. The expression for the allowed energies for multielectron atoms has the same form as Equation 42.35 with Z replaced by an effective atomic number Z_{eff}:

$$E_n = -\frac{(13.6\text{ eV})Z_{\text{eff}}^2}{n^2} \tag{42.36}$$

where Z_{eff} depends on n and ℓ.

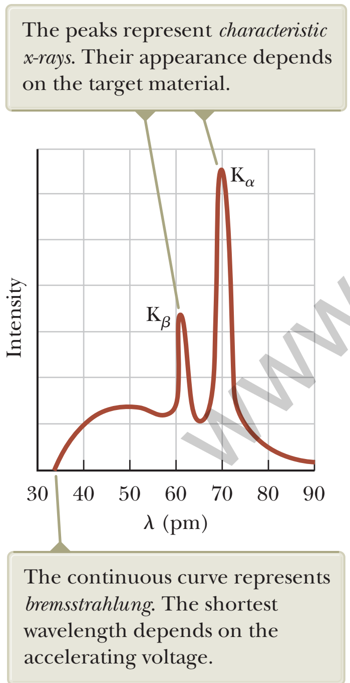


Figure 42.22 The x-ray spectrum of a metal target. The data shown were obtained when 37-keV electrons bombarded a molybdenum target.

X-Ray Spectra

X-rays are emitted when high-energy electrons or any other charged particles bombard a metal target. The x-ray spectrum typically consists of a broad continuous band containing a series of sharp lines as shown in Figure 42.22. In Section 34.6, we mentioned that an accelerated electric charge emits electromagnetic radiation. The x-rays in Figure 42.22 are the result of the slowing down of high-energy electrons as they strike the target. It may take several interactions with the atoms of the target before the electron gives up all its kinetic energy. The amount of kinetic energy given up in any interaction can vary from zero up to the entire kinetic energy of the electron. Therefore, the wavelength of radiation from these interactions lies in a continuous range from some minimum value up to infinity. It is this general slowing down of the electrons that provides the continuous curve in Figure 42.22, which shows the cutoff of x-rays below a minimum wavelength value that depends on the kinetic energy of the incoming electrons. X-ray radiation with its origin in the slowing down of electrons is called **bremsstrahlung,** the German word for "braking radiation."

Extremely high-energy bremsstrahlung can be used for the treatment of cancerous tissues. Figure 42.23 shows a machine that uses a linear accelerator to accelerate electrons up to 18 MeV and smash them into a tungsten target. The result is a beam of photons, up to a maximum energy of 18 MeV, which is actually in the gamma-ray range in Figure 34.13. This radiation is directed at the tumor in the patient.

Figure 42.23 Bremsstrahlung is created by this machine and used to treat cancer in a patient.

The discrete lines in Figure 42.22, called **characteristic x-rays** and discovered in 1908, have a different origin. Their origin remained unexplained until the details of atomic structure were understood. The first step in the production of characteristic x-rays occurs when a bombarding electron collides with a target atom. The electron must have sufficient energy to remove an inner-shell electron from the atom. The vacancy created in the shell is filled when an electron in a higher level drops down into the level containing the vacancy. The existence of characteristic lines in an x-ray spectrum is further direct evidence of the quantization of energy in atomic systems.

The time interval for atomic transitions to happen is very short, less than 10^{-9} s. This transition is accompanied by the emission of a photon whose energy equals the difference in energy between the two levels. Typically, the energy of such transitions is greater than 1 000 eV and the emitted x-ray photons have wavelengths in the range of 0.01 nm to 1 nm.

Let's assume the incoming electron has dislodged an atomic electron from the innermost shell, the K shell. If the vacancy is filled by an electron dropping from the next higher shell—the L shell—the photon emitted has an energy corresponding to the K_α characteristic x-ray line on the curve of Figure 42.22. In this notation, K refers to the final level of the electron and the subscript α, as the *first* letter of the Greek alphabet, refers to the initial level as the *first* one above the final level. Figure 42.24 shows this transition as well as others discussed below. If the vacancy in the K shell is filled by an electron dropping from the M shell, the K_β line in Figure 42.22 is produced.

Other characteristic x-ray lines are formed when electrons drop from upper levels to vacancies other than those in the K shell. For example, L lines are produced when vacancies in the L shell are filled by electrons dropping from higher shells. An L_α line is produced as an electron drops from the M shell to the L shell, and an L_β line is produced by a transition from the N shell to the L shell.

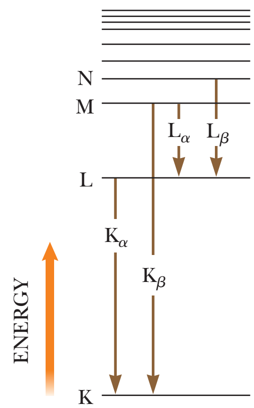


Figure 42.24 Transitions between higher and lower atomic energy levels that give rise to x-ray photons from heavy atoms when they are bombarded with high-energy electrons.

Although multielectron atoms cannot be analyzed exactly with either the Bohr model or the Schrödinger equation, we can apply Gauss's law from Chapter 24 to make some surprisingly accurate estimates of expected x-ray energies and wavelengths. Consider an atom of atomic number Z in which one of the two electrons in the K shell has been ejected. Imagine drawing a gaussian sphere immediately inside the most probable radius of the L electrons. The electric field at the position of the L electrons is a combination of the fields created by the nucleus, the single K electron, the other L electrons, and the outer electrons. The wave functions of the outer electrons are such that the electrons have a very high probability of being farther from the nucleus than the L electrons are. Therefore, the outer electrons are much more likely to be outside the gaussian surface than inside and, on average, do not contribute significantly to the electric field at the position of the L electrons. The effective charge inside the gaussian surface is the positive nuclear charge and one negative charge due to the single K electron. Ignoring the interactions between L electrons, a single L electron behaves as if it experiences an electric field due to a charge $(Z - 1)e$ enclosed by the gaussian surface. The nuclear charge is shielded by the electron in the K shell such that Z_{eff} in Equation 42.36 is $Z - 1$. For higher-level shells, the nuclear charge is shielded by electrons in all the inner shells.

We can now use Equation 42.36 to estimate the energy associated with an electron in the L shell:

$$E_L = -(Z-1)^2 \frac{13.6\text{ eV}}{2^2}$$

After the atom makes the transition, there are two electrons in the K shell. We can approximate the energy associated with one of these electrons as that of a one-electron atom. (In reality, the nuclear charge is reduced somewhat by the negative charge of the other electron, but let's ignore this effect.) Therefore,

$$E_K \approx -Z^2(13.6\text{ eV}) \tag{42.37}$$

As Example 42.5 shows, the energy of the atom with an electron in an M shell can be estimated in a similar fashion. Taking the energy difference between the initial and final levels, we can then calculate the energy and wavelength of the emitted photon.

In 1914, Henry G. J. Moseley (1887–1915) plotted $\sqrt{1/\lambda}$ versus the Z values for a number of elements where λ is the wavelength of the K_α line of each element. He found that the plot is a straight line as in Figure 42.25, which is consistent with rough calculations of the energy levels given by Equation 42.37. From this plot, Moseley determined the Z values of elements that had not yet been discovered and produced a periodic table in excellent agreement with the known chemical properties of the elements. Until that experiment, atomic numbers had been merely placeholders for the elements that appeared in the periodic table, the elements being ordered according to mass.

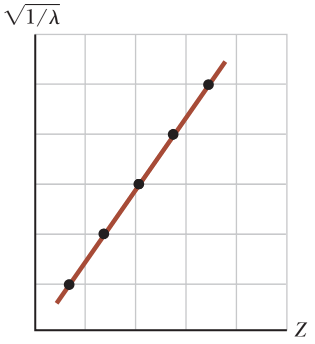


Figure 42.25 A Moseley plot of $\sqrt{1/\lambda}$ versus Z, where λ is the wavelength of the K_α x-ray line of the element of atomic number Z.

Quick Quiz 42.5 In an x-ray tube, as you increase the energy of the electrons striking the metal target, do the wavelengths of the characteristic x-rays **(a)** increase, **(b)** decrease, or **(c)** remain constant?

Quick Quiz 42.6 True or False: It is possible for an x-ray spectrum to show the continuous spectrum of x-rays without the presence of the characteristic x-rays.

Example 42.5 Estimating the Energy of an X-Ray

Estimate the energy of the characteristic x-ray emitted from a tungsten target when an electron drops from an M shell ($n = 3$ state) to a vacancy in the K shell ($n = 1$ state). The atomic number for tungsten is $Z = 74$.

SOLUTION

Conceptualize Imagine an accelerated electron striking a tungsten atom and ejecting an electron from the K shell ($n = 1$). Subsequently, an electron in the M shell ($n = 3$) drops down to fill the vacancy and the energy difference between the states is emitted as an x-ray photon.

Categorize We estimate the results using equations developed in this section, so we categorize this example as a substitution problem.

Use Equation 42.37 and $Z = 74$ for tungsten to estimate the energy associated with the electron in the K shell:

$$E_K \approx -(74)^2(13.6\text{ eV}) = -7.4 \times 10^4\text{ eV}$$

42.5 continued

Use Equation 42.36 and that nine electrons shield the nuclear charge (eight electrons in the $n = 2$ state and one electron in the $n = 1$ state) to estimate the energy of the M shell:

$$E_M \approx -\frac{(13.6\text{ eV})(74-9)^2}{(3)^2} \approx -6.4 \times 10^3\text{ eV}$$

Find the energy of the emitted x-ray photon:

$$hf = E_M - E_K \approx -6.4 \times 10^3\text{ eV} - (-7.4 \times 10^4\text{ eV})$$
$$\approx 6.8 \times 10^4\text{ eV} = \boxed{68\text{ keV}}$$

Consultation of x-ray tables shows that the M–K transition energies in tungsten vary from 66.9 keV to 67.7 keV, where the range of energies is due to slightly different energy values for states of different ℓ. Therefore, our estimate differs from the midpoint of this experimentally measured range by approximately 1%.

42.9 Spontaneous and Stimulated Transitions

We have seen that an atom absorbs and emits electromagnetic radiation only at frequencies that correspond to the energy differences between allowed states. Let's now examine more details of these processes. Consider an atom having the allowed energy levels labeled E_1, E_2, E_3, ... When radiation is incident on the atom, only those photons whose energy hf matches the energy separation ΔE between two energy levels can be absorbed by the atom as represented in Figure 42.26. This process is called **stimulated absorption** because the photon stimulates the atom to make the upward transition. At ordinary temperatures, most of the atoms in a sample are in the ground state. If a vessel containing many atoms of a gaseous element is illuminated with radiation of all possible photon frequencies (that is, a continuous spectrum), only those photons having energy $E_2 - E_1$, $E_3 - E_1$, $E_4 - E_1$, and so on are absorbed by the atoms. As a result of this absorption, some of the atoms are raised to excited states.

Once an atom is in an excited state, the excited atom can make a transition back to a lower energy level, emitting a photon in the process as in Figure 42.27. This process is known as **spontaneous emission** because it happens naturally, without requiring an event to trigger the transition. Typically, an atom remains in an excited state for only about 10^{-8} s.

In addition to spontaneous emission, **stimulated emission** occurs. Suppose an atom is in an excited state E_2 as in Figure 42.28 (page 1326). If the excited state is a *metastable state*—that is, if its lifetime is much longer than the typical 10^{-8} s lifetime of

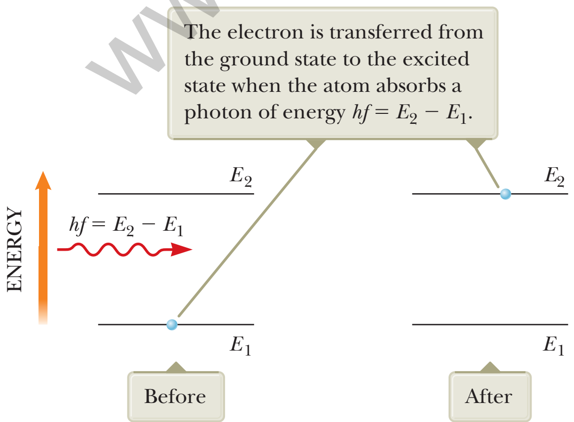


Figure 42.26 Stimulated absorption of a photon.

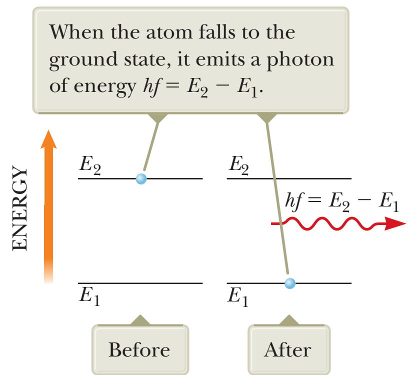


Figure 42.27 Spontaneous emission of a photon by an atom that is initially in the excited state E_2.

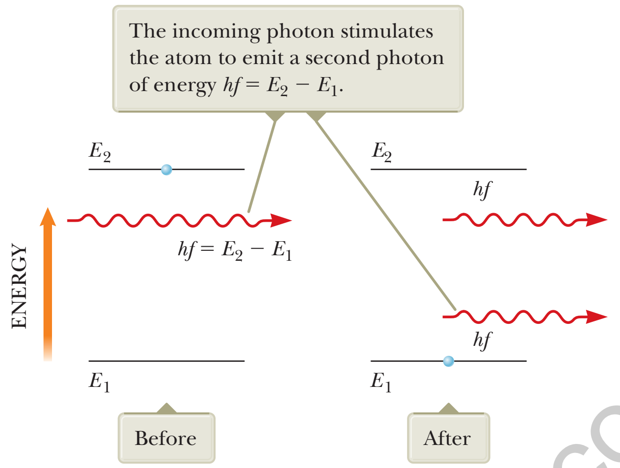


Figure 42.28 Stimulated emission of a photon by an incoming photon of energy $hf = E_2 - E_1$. Initially, the atom is in the excited state.

excited states—the time interval until spontaneous emission occurs is relatively long. Let's imagine that during that interval a photon of energy $hf = E_2 - E_1$ is incident on the atom. One possibility is that the photon energy is sufficient for the photon to ionize the atom. Another possibility is that the interaction between the incoming photon and the atom causes the atom to return to the ground state[11] and thereby emit a second photon with energy $hf = E_2 - E_1$. In this process, the incident photon is not absorbed; therefore, after the stimulated emission, two photons with identical energy exist: the incident photon and the emitted photon. The two are in phase and travel in the same direction, which is an important consideration in lasers, discussed next.

42.10 Lasers

In this section, we explore the nature of laser light and a variety of applications of lasers in our technological society. The primary properties of laser light that make it useful in these technological applications are the following:

- Laser light is coherent. The individual rays of light in a laser beam maintain a fixed phase relationship with one another.
- Laser light is monochromatic. Light in a laser beam has a very narrow range of wavelengths.
- Laser light has a small angle of divergence. The beam spreads out very little, even over large distances.

To understand the origin of these properties, let's combine our knowledge of atomic energy levels from this chapter with some special requirements for the atoms that emit laser light.

We have described how an incident photon can cause atomic energy transitions either upward (stimulated absorption) or downward (stimulated emission). The two processes are equally probable. When light is incident on a collection of atoms, a net absorption of energy usually occurs because when the system is in thermal equilibrium, many more atoms are in the ground state than in excited states. If the situation can be inverted so that more atoms are in an excited state than in the ground state, however, a net emission of photons can result. Such a condition is called **population inversion.**

Population inversion is, in fact, the fundamental principle involved in the operation of a **laser** (an acronym for *l*ight *a*mplification by *s*timulated *e*mission of *r*adia-

[11]This phenomenon is fundamentally due to *resonance.* The incoming photon has a frequency and drives the system of the atom at that frequency. Because the driving frequency matches that associated with a transition between states—one of the natural frequencies of the atom—there is a large response: the atom makes the transition.

tion). The full name indicates one of the requirements for laser light: to achieve laser action, the process of stimulated emission must occur.

Suppose an atom is in the excited state E_2 as in Figure 42.28 and a photon with energy $hf = E_2 - E_1$ is incident on it. As described in Section 42.9, the incoming photon can stimulate the excited atom to return to the ground state and thereby emit a second photon having the same energy hf and traveling in the same direction. The incident photon is not absorbed, so after the stimulated emission, there are two identical photons: the incident photon and the emitted photon. The emitted photon is in phase with the incident photon. These photons can stimulate other atoms to emit photons in a chain of similar processes. The many photons produced in this fashion are the source of the intense, coherent light in a laser.

For the stimulated emission to result in laser light, there must be a buildup of photons in the system. The following three conditions must be satisfied to achieve this buildup:

- The system must be in a state of population inversion: there must be more atoms in an excited state than in the ground state. That must be true because the number of photons emitted must be greater than the number absorbed.
- The excited state of the system must be a *metastable state,* meaning that its lifetime must be long compared with the usually short lifetimes of excited states, which are typically 10^{-8} s. In this case, the population inversion can be established and stimulated emission is likely to occur before spontaneous emission.
- The emitted photons must be confined in the system long enough to enable them to stimulate further emission from other excited atoms. That is achieved by using reflecting mirrors at the ends of the system. One end is made totally reflecting, and the other is partially reflecting. A fraction of the light intensity passes through the partially reflecting end, forming the beam of laser light (Fig. 42.29).

One device that exhibits stimulated emission of radiation is the helium–neon gas laser. Figure 42.30 is an energy-level diagram for the neon atom in this system. The mixture of helium and neon is confined to a glass tube that is sealed at the ends by mirrors. A voltage applied across the tube causes electrons to sweep through the tube, colliding with the atoms of the gases and raising them into excited states. Neon atoms are excited to state E_3* through this process (the asterisk indicates a metastable state) and also as a result of collisions with excited helium atoms. Stimulated emission occurs, causing neon atoms to make transitions to state E_2. Neighboring excited atoms are also stimulated. The result is the production of coherent light at a wavelength of 632.8 nm.

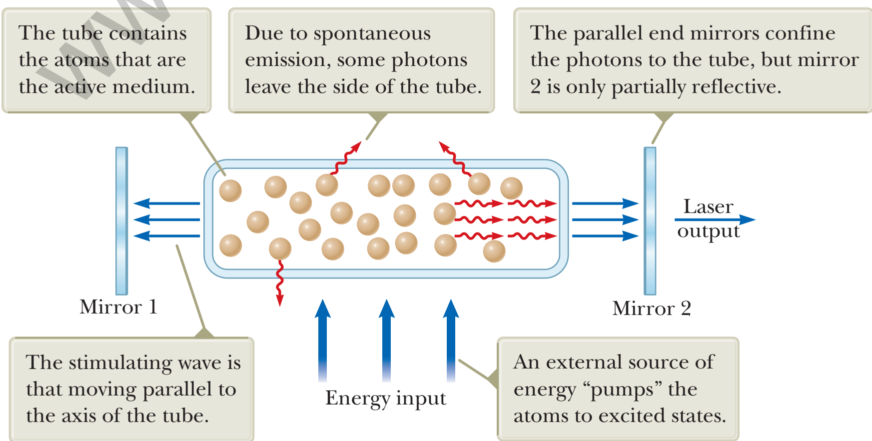


Figure 42.29 Schematic diagram of a laser design.

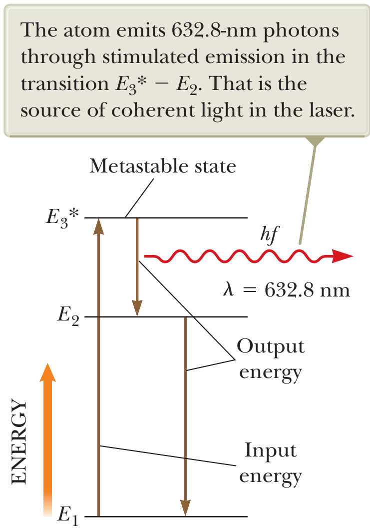


Figure 42.30 Energy-level diagram for a neon atom in a helium–neon laser.

Figure 42.31 This robot carrying laser scissors, which can cut up to 50 layers of fabric at a time, is one of the many applications of laser technology.

Applications

Since the development of the first laser in 1960, tremendous growth has occurred in laser technology. Lasers that cover wavelengths in the infrared, visible, and ultraviolet regions are now available. *Laser diodes* are used as laser pointers, and in surveying and construction rangefinders, fiber optic communication, DVD and Blu-ray players, and bar code readers. *Carbon dioxide lasers* are used in industry for welding and cutting, such as the process shown to cut fabric in Figure 42.31. *Excimer lasers* are used in Lasik eye surgery. A variety of other types of lasers exist and are used in various applications. These applications are possible because of the unique characteristics of laser light. In addition to being highly monochromatic, laser light is also highly directional and can be sharply focused to produce regions of extremely intense light energy (with energy densities 10^{12} times the density in the flame of a typical cutting torch).

Lasers are used in precision long-range distance measurement (range finding). In recent years, it has become important in astronomy and geophysics to measure as precisely as possible the distances from various points on the surface of the Earth to a point on the Moon's surface. To facilitate these measurements, the *Apollo* astronauts set up a 0.5-m square of reflector prisms on the Moon, which enables laser pulses directed from an Earth-based station to be retroreflected to the same station (see Fig. 35.8a). Using the known speed of light and the measured round-trip travel time of a laser pulse, the Earth–Moon distance can be determined to a precision of better than 10 cm.

Because various laser wavelengths can be absorbed in specific biological tissues, lasers have a number of medical applications. For example, certain laser procedures have greatly reduced blindness in patients with glaucoma and diabetes. Glaucoma is a widespread eye condition characterized by a high fluid pressure in the eye, a condition that can lead to destruction of the optic nerve. A simple laser operation (iridectomy) can "burn" open a tiny hole in a clogged membrane, relieving the destructive pressure. A serious side effect of diabetes is neovascularization, the proliferation of weak blood vessels, which often leak blood. When neovascularization occurs in the retina, vision deteriorates (diabetic retinopathy) and finally is destroyed. Today, it is possible to direct the green light from an argon ion laser through the clear eye lens and eye fluid, focus on the retina edges, and photocoagulate the leaky vessels. Even people who have only minor vision defects such as nearsightedness are benefiting from the use of lasers to reshape the cornea, changing its focal length and reducing the need for eyeglasses.

Laser surgery is now an everyday occurrence at hospitals and medical clinics around the world. Infrared light at 10 μm from a carbon dioxide laser can cut through muscle tissue, primarily by vaporizing the water contained in cellular material. Laser power of approximately 100 W is required in this technique. The advantage of the "laser knife" over conventional methods is that laser radiation cuts tissue and coagulates blood at the same time, leading to a substantial reduction in blood loss. In addition, the technique virtually eliminates cell migration, an important consideration when tumors are being removed.

A laser beam can be trapped in fine optical fiber light guides (endoscopes) by means of total internal reflection. An endoscope can be introduced through natural orifices, conducted around internal organs, and directed to specific interior body locations, eliminating the need for invasive surgery. For example, bleeding in the gastrointestinal tract can be optically cauterized by endoscopes inserted through the patient's mouth.

In biological and medical research, it is often important to isolate and collect unusual cells for study and growth. A laser cell separator exploits the tagging of specific cells with fluorescent dyes. All cells are then dropped from a tiny charged nozzle and laser-scanned for the dye tag. If triggered by the correct light-emitting tag, a small voltage applied to parallel plates deflects the falling electrically charged cell into a collection beaker.

An exciting area of research and technological applications arose in the 1990s with the development of *laser trapping* of atoms. One scheme, called *optical molasses* and developed by Steven Chu of Stanford University and his colleagues, involves focusing six laser beams onto a small region in which atoms are to be trapped. Each pair of lasers is along one of the x, y, and z axes and emits light in opposite directions (Fig. 42.32). The frequency of the laser light is tuned to be slightly below the absorption frequency of the subject atom. Imagine that an atom has been placed into the trap region and moves along the positive x axis toward the laser that is emitting light toward it (the rightmost laser on the x axis in Fig. 42.32). Because the atom is moving, the light from the laser appears Doppler-shifted upward in frequency in the reference frame of the atom. Therefore, a match between the Doppler-shifted laser frequency and the absorption frequency of the atom exists and the atom absorbs photons.[12] The momentum carried by these photons results in the atom being pushed back to the center of the trap. By incorporating six lasers, the atoms are pushed back into the trap regardless of which way they move along any axis.

In 1986, Chu developed *optical tweezers,* a device that uses a single tightly focused laser beam to trap and manipulate small particles. In combination with microscopes, optical tweezers have opened up many new possibilities for biologists. Optical tweezers have been used to manipulate live bacteria without damage, move chromosomes within a cell nucleus, and measure the elastic properties of a single DNA molecule. Chu shared the 1997 Nobel Prize in Physics with two of his colleagues for the development of the techniques of optical trapping.

An extension of laser trapping, *laser cooling,* is possible because the normal high speeds of the atoms are reduced when they are restricted to the region of the trap. As a result, the temperature of the collection of atoms can be reduced to a few microkelvins. The technique of laser cooling allows scientists to study the behavior of atoms at extremely low temperatures (Fig. 42.33).

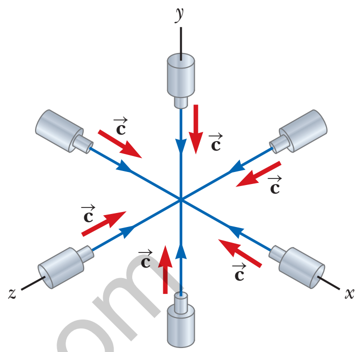
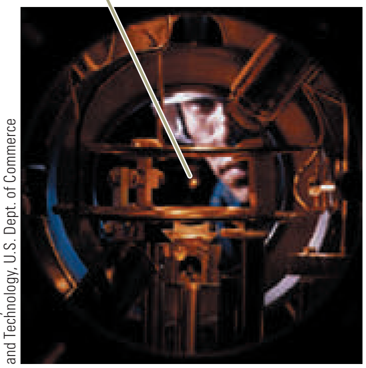


Figure 42.32 An optical trap for atoms is formed at the intersection point of six counterpropagating laser beams along mutually perpendicular axes.

The orange dot is the sample of trapped sodium atoms.



Courtesy of National Institute of Standards and Technology, U.S. Dept. of Commerce

Figure 42.33 A staff member of the National Institute of Standards and Technology views a sample of trapped sodium atoms cooled to a temperature of less than 1 mK.

Summary

Concepts and Principles

The wavelengths of spectral lines from hydrogen, called the **Balmer series,** can be described by the equation

$$\frac{1}{\lambda} = R_H\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \quad n = 3, 4, 5, \ldots \tag{42.1}$$

where R_H is the **Rydberg constant.** The spectral lines corresponding to values of n from 3 to 6 are in the visible range of the electromagnetic spectrum. Values of n higher than 6 correspond to spectral lines in the ultraviolet region of the spectrum.

continued

[12]The laser light traveling in the same direction as the atom is Doppler-shifted further downward in frequency, so there is no absorption. Therefore, the atom is not pushed out of the trap by the diametrically opposed laser.

The Bohr model of the atom is successful in describing some details of the spectra of atomic hydrogen and hydrogen-like ions. One basic assumption of the model is that the electron can exist only in discrete orbits such that the angular momentum of the electron is an integral multiple of $h/2\pi = \hbar$. When we assume circular orbits and a simple Coulomb attraction between electron and proton, the energies of the quantum states for hydrogen are calculated to be

$$E_n = -\frac{k_e e^2}{2a_0}\left(\frac{1}{n^2}\right) \quad n = 1, 2, 3, \ldots \tag{42.13}$$

where n is an integer called the **quantum number,** k_e is the Coulomb constant, e is the electronic charge, and $a_0 = 0.052\,9$ nm is the **Bohr radius.**

If the electron in a hydrogen atom makes a transition from an orbit whose quantum number is n_i to one whose quantum number is n_f, where $n_f < n_i$, a photon is emitted by the atom. The frequency of this photon is

$$f = \frac{k_e e^2}{2a_0 h}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \tag{42.15}$$

Quantum mechanics can be applied to the hydrogen atom by the use of the potential energy function $U(r) = -k_e e^2/r$ in the Schrödinger equation. The solution to this equation yields wave functions for allowed states and allowed energies:

$$E_n = -\left(\frac{k_e e^2}{2a_0}\right)\frac{1}{n^2} = -\frac{13.606 \text{ eV}}{n^2} \quad n = 1, 2, 3, \ldots \tag{42.21}$$

where n is the **principal quantum number.** The allowed wave functions depend on three quantum numbers: n, ℓ, and m_ℓ, where ℓ is the **orbital quantum number** and m_ℓ is the **orbital magnetic quantum number.** The restrictions on the quantum numbers are

$$n = 1, 2, 3, \ldots$$
$$\ell = 0, 1, 2, \ldots, n - 1$$
$$m_\ell = -\ell, -\ell + 1, \ldots, \ell - 1, \ell$$

All states having the same principal quantum number n form a **shell,** identified by the letters K, L, M, . . . (corresponding to $n = 1, 2, 3, \ldots$). All states having the same values of n and ℓ form a **subshell,** designated by the letters $s, p, d, f, \ldots$ (corresponding to $\ell = 0, 1, 2, 3, \ldots$).

An atom in a state characterized by a specific value of n can have the following values of L, the magnitude of the atom's orbital angular momentum $\vec{\mathbf{L}}$:

$$L = \sqrt{\ell(\ell + 1)}\,\hbar$$
$$\ell = 0, 1, 2, \ldots, n - 1 \tag{42.27}$$

The allowed values of the projection of $\vec{\mathbf{L}}$ along the z axis are

$$L_z = m_\ell \hbar \tag{42.28}$$

Only discrete values of L_z are allowed as determined by the restrictions on m_ℓ. This quantization of L_z is referred to as **space quantization.**

The electron has an intrinsic angular momentum called the **spin angular momentum.** Electron spin can be described by a single quantum number $s = \frac{1}{2}$. To describe a quantum state completely, it is necessary to include a fourth quantum number m_s, called the **spin magnetic quantum number.** This quantum number can have only two values, $\pm\frac{1}{2}$. The magnitude of the spin angular momentum is

$$S = \frac{\sqrt{3}}{2}\hbar \tag{42.30}$$

and the z component of $\vec{\mathbf{S}}$ is

$$S_z = m_s \hbar = \pm\tfrac{1}{2}\hbar \tag{42.31}$$

That is, the spin angular momentum is also quantized in space, as specified by the spin magnetic quantum number $m_s = \pm\frac{1}{2}$.

The **exclusion principle** states that **no two electrons in an atom can be in the same quantum state.** In other words, no two electrons can have the same set of quantum numbers n, ℓ, m_ℓ, and m_s. Using this principle, the electronic configurations of the elements can be determined. This principle serves as a basis for understanding atomic structure and the chemical properties of the elements.

The magnetic moment $\vec{\boldsymbol{\mu}}_{\text{spin}}$ associated with the spin angular momentum of an electron is

$$\vec{\boldsymbol{\mu}}_{\text{spin}} = -\frac{e}{m_e}\vec{\mathbf{S}} \tag{42.32}$$

The z component of $\vec{\boldsymbol{\mu}}_{\text{spin}}$ can have the values

$$\mu_{\text{spin},z} = \pm\frac{e\hbar}{2m_e} \tag{42.33}$$

The x-ray spectrum of a metal target consists of a set of sharp characteristic lines superimposed on a broad continuous spectrum. **Bremsstrahlung** is x-radiation with its origin in the slowing down of high-energy electrons as they encounter the target. **Characteristic x-rays** are emitted by atoms when an electron undergoes a transition from an outer shell to a vacancy in an inner shell.

Atomic transitions can be described with three processes: **stimulated absorption,** in which an incoming photon raises the atom to a higher energy state; **spontaneous emission,** in which the atom makes a transition to a lower energy state, emitting a photon; and **stimulated emission,** in which an incident photon causes an excited atom to make a downward transition, emitting a photon identical to the incident one.

Objective Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. **(i)** What is the principal quantum number of the initial state of an atom as it emits an M_β line in an x-ray spectrum? (a) 1 (b) 2 (c) 3 (d) 4 (e) 5 **(ii)** What is the principal quantum number of the final state for this transition? Choose from the same possibilities as in part (i).

2. If an electron in an atom has the quantum numbers $n = 3$, $\ell = 2$, $m_\ell = 1$, and $m_s = \frac{1}{2}$, what state is it in? (a) $3s$ (b) $3p$ (c) $3d$ (d) $4d$ (e) $3f$

3. An electron in the $n = 5$ energy level of hydrogen undergoes a transition to the $n = 3$ energy level. What is the wavelength of the photon the atom emits in this process? (a) 2.28×10^{-6} m (b) 8.20×10^{-7} m (c) 3.64×10^{-7} m (d) 1.28×10^{-6} m (e) 5.92×10^{-5} m

4. Consider the $n = 3$ energy level in a hydrogen atom. How many electrons can be placed in this level? (a) 1 (b) 2 (c) 8 (d) 9 (e) 18

5. Which of the following is *not* one of the basic assumptions of the Bohr model of hydrogen? (a) Only certain electron orbits are stable and allowed. (b) The electron moves in circular orbits about the proton under the influence of the Coulomb force. (c) The charge on the electron is quantized. (d) Radiation is emitted by the atom when the electron moves from a higher energy state to a lower energy state. (e) The angular momentum associated with the electron's orbital motion is quantized.

6. Let $-E$ represent the energy of a hydrogen atom. **(i)** What is the kinetic energy of the electron? (a) $2E$ (b) E (c) 0 (d) $-E$ (e) $-2E$ **(ii)** What is the potential energy of the atom? Choose from the same possibilities (a) through (e).

7. The periodic table is based on which of the following principles? (a) The uncertainty principle. (b) All electrons in an atom must have the same set of quantum numbers. (c) Energy is conserved in all interactions. (d) All electrons in an atom are in orbitals having the same energy. (e) No two electrons in an atom can have the same set of quantum numbers.

8. (a) Can a hydrogen atom in the ground state absorb a photon of energy less than 13.6 eV? (b) Can this atom absorb a photon of energy greater than 13.6 eV?

9. Which of the following electronic configurations are *not* allowed for an atom? Choose all correct answers. (a) $2s^22p^6$ (b) $3s^23p^7$ (c) $3d^74s^2$ (d) $3d^{10}4s^24p^6$ (e) $1s^22s^22d^1$

10. What can be concluded about a hydrogen atom with its electron in the d state? (a) The atom is ionized. (b) The orbital quantum number is $\ell = 1$. (c) The principal quantum number is $n = 2$. (d) The atom is in its ground state. (e) The orbital angular momentum of the atom is not zero.

11. **(i)** Rank the following transitions for a hydrogen atom from the transition with the greatest gain in energy to that with the greatest loss, showing any cases of equality. (a) $n_i = 2$; $n_f = 5$ (b) $n_i = 5$; $n_f = 3$ (c) $n_i = 7$; $n_f = 4$ (d) $n_i = 4$; $n_f = 7$ **(ii)** Rank the same transitions as in part (i) according to the wavelength of the photon absorbed or emitted by an otherwise isolated atom from greatest wavelength to smallest.

12. When an atom emits a photon, what happens? (a) One of its electrons leaves the atom. (b) The atom moves to a state of higher energy. (c) The atom moves to a state of lower energy. (d) One of its electrons collides with another particle. (e) None of those events occur.

13. (a) In the hydrogen atom, can the quantum number n increase without limit? (b) Can the frequency of possible discrete lines in the spectrum of hydrogen increase without limit? (c) Can the wavelength of possible discrete lines in the spectrum of hydrogen increase without limit?

14. Consider the quantum numbers (a) n, (b) ℓ, (c) m_ℓ, and (d) m_s. **(i)** Which of these quantum numbers are fractional as opposed to being integers? **(ii)** Which can sometimes attain negative values? **(iii)** Which can be zero?

15. When an electron collides with an atom, it can transfer all or some of its energy to the atom. A hydrogen atom is in its ground state. Incident on the atom are several electrons, each having a kinetic energy of 10.5 eV. What is the result? (a) The atom can be excited to a higher allowed state. (b) The atom is ionized. (c) The electrons pass by the atom without interaction.

Conceptual Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. Why is stimulated emission so important in the operation of a laser?
2. An energy of about 21 eV is required to excite an electron in a helium atom from the 1*s* state to the 2*s* state. The same transition for the He^+ ion requires approximately twice as much energy. Explain.
3. Why are three quantum numbers needed to describe the state of a one-electron atom (ignoring spin)?
4. Compare the Bohr theory and the Schrödinger treatment of the hydrogen atom, specifically commenting on their treatment of total energy and orbital angular momentum of the atom.
5. Could the Stern–Gerlach experiment be performed with ions rather than neutral atoms? Explain.
6. Why is a *nonuniform* magnetic field used in the Stern–Gerlach experiment?
7. Discuss some consequences of the exclusion principle.
8. (a) According to Bohr's model of the hydrogen atom, what is the uncertainty in the radial coordinate of the electron? (b) What is the uncertainty in the radial component of the velocity of the electron? (c) In what way does the model violate the uncertainty principle?
9. Why do lithium, potassium, and sodium exhibit similar chemical properties?
10. It is easy to understand how two electrons (one spin up, one spin down) fill the $n = 1$ or K shell for a helium atom. How is it possible that eight more electrons are allowed in the $n = 2$ shell, filling the K and L shells for a neon atom?
11. Suppose the electron in the hydrogen atom obeyed classical mechanics rather than quantum mechanics. Why should a gas of such hypothetical atoms emit a continuous spectrum rather than the observed line spectrum?
12. Does the intensity of light from a laser fall off as $1/r^2$? Explain.

Problems

Enhanced WebAssign The problems found in this chapter may be assigned online in Enhanced WebAssign

1. straightforward; 2. intermediate; 3. challenging

1. full solution available in the *Student Solutions Manual/Study Guide*

AMT Analysis Model tutorial available in Enhanced WebAssign

GP Guided Problem

M Master It tutorial available in Enhanced WebAssign

W Watch It video solution available in Enhanced WebAssign

Section 42.1 Atomic Spectra of Gases

1. The wavelengths of the Lyman series for hydrogen are given by

$$\frac{1}{\lambda} = R_H\left(1 - \frac{1}{n^2}\right) \qquad n = 2, 3, 4, \ldots$$

(a) Calculate the wavelengths of the first three lines in this series. (b) Identify the region of the electromagnetic spectrum in which these lines appear.

2. The wavelengths of the Paschen series for hydrogen are given by

$$\frac{1}{\lambda} = R_H\left(\frac{1}{3^2} - \frac{1}{n^2}\right) \qquad n = 4, 5, 6, \ldots$$

(a) Calculate the wavelengths of the first three lines in this series. (b) Identify the region of the electromagnetic spectrum in which these lines appear.

3. An isolated atom of a certain element emits light of wavelength 520 nm when the atom falls from its fifth excited state into its second excited state. The atom emits a photon of wavelength 410 nm when it drops from its sixth excited state into its second excited state. Find the wavelength of the light radiated when the atom makes a transition from its sixth to its fifth excited state.

4. An isolated atom of a certain element emits light of wavelength λ_{m1} when the atom falls from its state with quantum number *m* into its ground state of quantum number 1. The atom emits a photon of wavelength λ_{n1} when the atom falls from its state with quantum number *n* into its ground state. (a) Find the wavelength of the light radiated when the atom makes a transition from the *m* state to the *n* state. (b) Show that $k_{mn} = |k_{m1} - k_{n1}|$, where $k_{ij} = 2\pi/\lambda_{ij}$ is the wave number of the photon. This problem exemplifies the *Ritz combination principle*, an empirical rule formulated in 1908.

5. (a) What value of n_i is associated with the 94.96-nm spectral line in the Lyman series of hydrogen? (b) **What If?** Could this wavelength be associated with the Paschen series? (c) Could this wavelength be associated with the Balmer series?

Section 42.2 Early Models of the Atom

6. According to classical physics, a charge *e* moving with an acceleration *a* radiates energy at a rate

$$\frac{dE}{dt} = -\frac{1}{6\pi\epsilon_0}\frac{e^2a^2}{c^3}$$

(a) Show that an electron in a classical hydrogen atom (see Fig. 42.5) spirals into the nucleus at a rate

$$\frac{dr}{dt} = -\frac{e^4}{12\pi^2\epsilon_0^2 m_e^2 c^3}\left(\frac{1}{r^2}\right)$$

(b) Find the time interval over which the electron reaches $r = 0$, starting from $r_0 = 2.00 \times 10^{-10}$ m.

7. **Review.** In the Rutherford scattering experiment, 4.00-MeV alpha particles scatter off gold nuclei (containing 79 protons and 118 neutrons). Assume a particular alpha particle moves directly toward the gold nucleus and scatters backward at 180°, and that the gold nucleus remains fixed throughout the entire process. Determine (a) the distance of closest approach of the alpha particle to the gold nucleus and (b) the maximum force exerted on the alpha particle.

Section 42.3 Bohr's Model of the Hydrogen Atom

Note: In this section, unless otherwise indicated, assume the hydrogen atom is treated with the Bohr model.

8. Show that the speed of the electron in the nth Bohr orbit in hydrogen is given by

$$v_n = \frac{k_e e^2}{n\hbar}$$

9. How much energy is required to ionize hydrogen (a) when it is in the ground state and (b) when it is in the state for which $n = 3$?

10. M What is the energy of a photon that, when absorbed by a hydrogen atom, could cause an electronic transition from (a) the $n = 2$ state to the $n = 5$ state and (b) the $n = 4$ state to the $n = 6$ state?

11. A photon is emitted when a hydrogen atom undergoes a transition from the $n = 5$ state to the $n = 3$ state. Calculate (a) the energy (in electron volts), (b) the wavelength, and (c) the frequency of the emitted photon.

12. The Balmer series for the hydrogen atom corresponds to electronic transitions that terminate in the state with quantum number $n = 2$ as shown in Figure P42.12. Consider the photon of longest wavelength corresponding to a transition shown in the figure. Determine (a) its energy and (b) its wavelength. Consider the spectral line of shortest wavelength corresponding to a transition shown in the figure. Find (c) its photon energy and (d) its wavelength. (e) What is the shortest possible wavelength in the Balmer series?

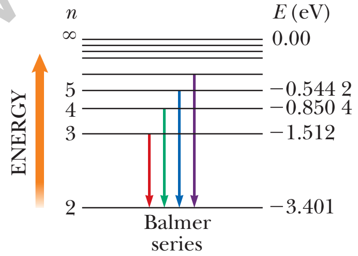


Figure P42.12

13. For a hydrogen atom in its ground state, compute (a) the orbital speed of the electron, (b) the kinetic energy of the electron, and (c) the electric potential energy of the atom.

14. AMT Two hydrogen atoms collide head-on and end up with zero kinetic energy. Each atom then emits light with a wavelength of 121.6 nm ($n = 2$ to $n = 1$ transition). At what speed were the atoms moving before the collision?

15. (a) Calculate the angular momentum of the Moon due to its orbital motion about the Earth. In your calculation, use 3.84×10^8 m as the average Earth–Moon distance and 2.36×10^6 s as the period of the Moon in its orbit. (b) Assume that the Moon's angular momentum is described by Bohr's assumption $mvr = n\hbar$. Determine the corresponding quantum number. (c) By what fraction would the Earth–Moon distance have to be increased to raise the quantum number by 1?

16. W A monochromatic beam of light is absorbed by a collection of ground-state hydrogen atoms in such a way that six different wavelengths are observed when the hydrogen relaxes back to the ground state. (a) What is the wavelength of the incident beam? Explain the steps in your solution. (b) What is the longest wavelength in the emission spectrum of these atoms? (c) To what portion of the electromagnetic spectrum and (d) to what series does it belong? (e) What is the shortest wavelength? (f) To what portion of the electromagnetic spectrum and (g) to what series does it belong?

17. A hydrogen atom is in its second excited state, corresponding to $n = 3$. Find (a) the radius of the electron's Bohr orbit and (b) the de Broglie wavelength of the electron in this orbit.

18. M W A hydrogen atom is in its first excited state ($n = 2$). Calculate (a) the radius of the orbit, (b) the linear momentum of the electron, (c) the angular momentum of the electron, (d) the kinetic energy of the electron, (e) the potential energy of the system, and (f) the total energy of the system.

19. A photon with energy 2.28 eV is absorbed by a hydrogen atom. Find (a) the minimum n for a hydrogen atom that can be ionized by such a photon and (b) the speed of the electron released from the state in part (a) when it is far from the nucleus.

20. An electron is in the nth Bohr orbit of the hydrogen atom. (a) Show that the period of the electron is $T = n^3 t_0$ and determine the numerical value of t_0. (b) On average, an electron remains in the $n = 2$ orbit for approximately 10 μs before it jumps down to the $n = 1$ (ground-state) orbit. How many revolutions does the electron make in the excited state? (c) Define the period of one revolution as an electron year, analogous to an Earth year being the period of the Earth's motion around the Sun. Explain whether we should think of the electron in the $n = 2$ orbit as "living for a long time."

21. (a) Construct an energy-level diagram for the He^+ ion, for which $Z = 2$, using the Bohr model. (b) What is the ionization energy for He^+?

Section 42.4 The Quantum Model of the Hydrogen Atom

22. A general expression for the energy levels of one-electron atoms and ions is

$$E_n = -\frac{\mu k_e^2 q_1^2 q_2^2}{2\hbar^2 n^2}$$

Here μ is the reduced mass of the atom, given by $\mu = m_1m_2/(m_1 + m_2)$, where m_1 is the mass of the electron and m_2 is the mass of the nucleus; k_e is the Coulomb constant; and q_1 and q_2 are the charges of the electron and the nucleus, respectively. The wavelength for the $n = 3$ to $n = 2$ transition of the hydrogen atom is 656.3 nm (visible red light). What are the wavelengths for this same transition in (a) positronium, which consists of an electron and a positron, and (b) singly ionized helium? *Note:* A positron is a positively charged electron.

23. Atoms of the same element but with different numbers of neutrons in the nucleus are called *isotopes*. Ordinary hydrogen gas is a mixture of two isotopes containing either one- or two-particle nuclei. These isotopes are hydrogen-1, with a proton nucleus, and hydrogen-2, called deuterium, with a deuteron nucleus. A deuteron is one proton and one neutron bound together. Hydrogen-1 and deuterium have identical chemical properties, but they can be separated via an ultracentrifuge or by other methods. Their emission spectra show lines of the same colors at very slightly different wavelengths. (a) Use the equation given in Problem 22 to show that the difference in wavelength between the hydrogen-1 and deuterium spectral lines associated with a particular electron transition is given by

$$\lambda_H - \lambda_D = \left(1 - \frac{\mu_H}{\mu_D}\right)\lambda_H$$

(b) Find the wavelength difference for the Balmer alpha line of hydrogen, with wavelength 656.3 nm, emitted by an atom making a transition from an $n = 3$ state to an $n = 2$ state. Harold Urey observed this wavelength difference in 1931 and so confirmed his discovery of deuterium.

24. An electron of momentum p is at a distance r from a stationary proton. The electron has kinetic energy $K = p^2/2m_e$. The atom has potential energy $U = -k_ee^2/r$ and total energy $E = K + U$. If the electron is bound to the proton to form a hydrogen atom, its average position is at the proton but the uncertainty in its position is approximately equal to the radius r of its orbit. The electron's average vector momentum is zero, but its average squared momentum is approximately equal to the squared uncertainty in its momentum as given by the uncertainty principle. Treating the atom as a one-dimensional system, (a) estimate the uncertainty in the electron's momentum in terms of r. Estimate the electron's (b) kinetic energy and (c) total energy in terms of r. The actual value of r is the one that *minimizes the total energy,* resulting in a stable atom. Find (d) that value of r and (e) the resulting total energy. (f) State how your answers compare with the predictions of the Bohr theory.

Section 42.5 The Wave Functions for Hydrogen

25. Plot the wave function $\psi_{1s}(r)$ versus r (see Eq. 42.22) and the radial probability density function $P_{1s}(r)$ versus r (see Eq. 42.25) for hydrogen. Let r range from 0 to $1.5a_0$, where a_0 is the Bohr radius.

26. For a spherically symmetric state of a hydrogen atom, the Schrödinger equation in spherical coordinates is

$$-\frac{\hbar^2}{2m_e}\left(\frac{d^2\psi}{dr^2} + \frac{2}{r}\frac{d\psi}{dr}\right) - \frac{k_ee^2}{r}\psi = E\psi$$

(a) Show that the 1*s* wave function for an electron in hydrogen,

$$\psi_{1s}(r) = \frac{1}{\sqrt{\pi a_0^3}}e^{-r/a_0}$$

satisfies the Schrödinger equation. (b) What is the energy of the atom for this state?

27. The radial function $R(r)$ of the wave function for a hydrogen atom in the 2*p* state is

$$\psi_{2p} = \frac{1}{\sqrt{3}(2a_0)^{3/2}}\frac{r}{a_0}e^{-r/2a_0}$$

What is the most likely distance from the nucleus to find an electron in the 2*p* state?

28. The ground-state wave function for the electron in a hydrogen atom is

$$\psi_{1s}(r) = \frac{1}{\sqrt{\pi a_0^3}}e^{-r/a_0}$$

where r is the radial coordinate of the electron and a_0 is the Bohr radius. (a) Show that the wave function as given is normalized. (b) Find the probability of locating the electron between $r_1 = a_0/2$ and $r_2 = 3a_0/2$.

29. In an experiment, a large number of electrons are fired at a sample of neutral hydrogen atoms and observations are made of how the incident particles scatter. The electron in the ground state of a hydrogen atom is found to be momentarily at a distance $a_0/2$ from the nucleus in 1 000 of the observations. In this set of trials, how many times is the atomic electron observed at a distance $2a_0$ from the nucleus?

Section 42.6 Physical Interpretation of the Quantum Numbers

30. List the possible sets of quantum numbers for the hydrogen atom associated with (a) the 3*d* subshell and (b) the 3*p* subshell.

31. If a hydrogen atom has orbital angular momentum 4.714×10^{-34} J · s, what is the orbital quantum number for the state of the atom?

32. Find all possible values of (a) L, (b) L_z, and (c) θ for a hydrogen atom in a 3*d* state.

33. Calculate the magnitude of the orbital angular momentum for a hydrogen atom in (a) the 4*d* state and (b) the 6*f* state.

34. How many sets of quantum numbers are possible for a hydrogen atom for which (a) $n = 1$, (b) $n = 2$, (c) $n = 3$, (d) $n = 4$, and (e) $n = 5$?

35. An electron in a sodium atom is in the N shell. Determine the maximum value the z component of its angular momentum could have.

36. W (a) Find the mass density of a proton, modeling it as a solid sphere of radius 1.00×10^{-15} m. (b) **What If?** Consider a classical model of an electron as a uniform solid sphere with the same density as the proton. Find its radius. (c) Imagine that this electron possesses spin angular momentum $I\omega = \hbar/2$ because of classical rotation about the z axis. Determine the speed of a point on the equator of the electron. (d) State how this speed compares with the speed of light.

37. A hydrogen atom is in its fifth excited state, with principal quantum number 6. The atom emits a photon with a wavelength of 1 090 nm. Determine the maximum possible magnitude of the orbital angular momentum of the atom after emission.

38. *Why is the following situation impossible?* A photon of wavelength 88.0 nm strikes a clean aluminum surface, ejecting a photoelectron. The photoelectron then strikes a hydrogen atom in its ground state, transferring energy to it and exciting the atom to a higher quantum state.

39. M The ρ^- meson has a charge of $-e$, a spin quantum number of 1, and a mass 1 507 times that of the electron. The possible values for its spin magnetic quantum number are -1, 0, and 1. **What If?** Imagine that the electrons in atoms are replaced by ρ^- mesons. List the possible sets of quantum numbers for ρ^- mesons in the $3d$ subshell.

Section 42.7 The Exclusion Principle and the Periodic Table

40. (a) As we go down the periodic table, which subshell is filled first, the $3d$ or the $4s$ subshell? (b) Which electronic configuration has a lower energy, $[\text{Ar}]3d^44s^2$ or $[\text{Ar}]3d^54s^1$? *Note:* The notation [Ar] represents the filled configuration for argon. *Suggestion:* Which has the greater number of unpaired spins? (c) Identify the element with the electronic configuration in part (b).

41. (a) Write out the electronic configuration of the ground state for nitrogen ($Z = 7$). (b) Write out the values for the possible set of quantum numbers n, ℓ, m_ℓ, and m_s for the electrons in nitrogen.

42. Devise a table similar to that shown in Figure 42.18 for atoms containing 11 through 19 electrons. Use Hund's rule and educated guesswork.

43. A certain element has its outermost electron in a $3p$ subshell. It has valence +3 because it has three more electrons than a certain noble gas. What element is it?

44. Scanning through Figure 42.19 in order of increasing atomic number, notice that the electrons usually fill the subshells in such a way that those subshells with the lowest values of $n + \ell$ are filled first. If two subshells have the same value of $n + \ell$, the one with the lower value of n is generally filled first. Using these two rules, write the order in which the subshells are filled through $n + \ell = 7$.

45. Two electrons in the same atom both have $n = 3$ and $\ell = 1$. Assume the electrons are distinguishable, so that interchanging them defines a new state. (a) How many states of the atom are possible considering the quantum numbers these two electrons can have? (b) **What If?** How many states would be possible if the exclusion principle were inoperative?

46. For a neutral atom of element 110, what would be the probable ground-state electronic configuration?

47. **Review.** For an electron with magnetic moment $\vec{\boldsymbol{\mu}}_s$ in a magnetic field $\vec{\mathbf{B}}$, Section 29.5 showed the following. The electron–field system can be in a higher energy state with the z component of the electron's magnetic moment opposite the field or a lower energy state with the z component of the magnetic moment in the direction of the field. The difference in energy between the two states is $2\mu_B B$.

Under high resolution, many spectral lines are observed to be doublets. The most famous doublet is the pair of two yellow lines in the spectrum of sodium (the D lines), with wavelengths of 588.995 nm and 589.592 nm. Their existence was explained in 1925 by Goudsmit and Uhlenbeck, who postulated that an electron has intrinsic spin angular momentum. When the sodium atom is excited with its outermost electron in a $3p$ state, the orbital motion of the outermost electron creates a magnetic field. The atom's energy is somewhat different depending on whether the electron is itself spin-up or spin-down in this field. Then the photon energy the atom radiates as it falls back into its ground state depends on the energy of the excited state. Calculate the magnitude of the internal magnetic field, mediating this so-called spin-orbit coupling.

Section 42.8 More on Atomic Spectra: Visible and X-Ray

48. In x-ray production, electrons are accelerated through a high voltage ΔV and then decelerated by striking a target. Show that the shortest wavelength of an x-ray that can be produced is

$$\lambda_{\min} = \frac{1\,240\text{ nm}\cdot\text{V}}{\Delta V}$$

49. What minimum accelerating voltage would be required to produce an x-ray with a wavelength of 70.0 pm?

50. A tungsten target is struck by electrons that have been accelerated from rest through a 40.0-keV potential difference. Find the shortest wavelength of the radiation emitted.

51. A bismuth target is struck by electrons, and x-rays are emitted. Estimate (a) the M- to L-shell transitional energy for bismuth and (b) the wavelength of the x-ray emitted when an electron falls from the M shell to the L shell.

52. The $3p$ level of sodium has an energy of -3.0 eV, and the $3d$ level has an energy of -1.5 eV. (a) Determine Z_{eff} for each of these states. (b) Explain the difference.

53. (a) Determine the possible values of the quantum numbers ℓ and m_ℓ for the He^+ ion in the state corresponding to $n = 3$. (b) What is the energy of this state?

54. M The K series of the discrete spectrum of tungsten contains wavelengths of 0.018 5 nm, 0.020 9 nm, and 0.021 5 nm. The K-shell ionization energy is 69.5 keV. Determine the ionization energies of the L, M, and N shells.

55. M Use the method illustrated in Example 42.5 to calculate the wavelength of the x-ray emitted from a molybdenum target ($Z = 42$) when an electron moves from the L shell ($n = 2$) to the K shell ($n = 1$).

56. In x-ray production, electrons are accelerated through a high voltage and then decelerated by striking a target. (a) To make possible the production of x-rays of wavelength λ, what is the minimum potential difference ΔV through which the electrons must be accelerated? (b) State in words how the required potential difference depends on the wavelength. (c) Explain whether your result predicts the correct minimum wavelength in Figure 42.22. (d) Does the relationship from part (a) apply to other kinds of electromagnetic radiation besides x-rays? (e) What does the potential difference approach as λ goes to zero? (f) What does the potential difference approach as λ increases without limit?

57. When an electron drops from the M shell ($n = 3$) to a vacancy in the K shell ($n = 1$), the measured wavelength of the emitted x-ray is found to be 0.101 nm. Identify the element.

Section 42.9 Spontaneous and Stimulated Transitions

Section 42.10 Lasers

58. Figure P42.58 shows portions of the energy-level diagrams of the helium and neon atoms. An electrical discharge excites the He atom from its ground state (arbitrarily assigned the energy $E_1 = 0$) to its excited state of 20.61 eV. The excited He atom collides with a Ne atom in its ground state and excites this atom to the state at 20.66 eV. Lasing action takes place for electron transitions from $E_3{}^*$ to E_2 in the Ne atoms. From the data in the figure, show that the wavelength of the red He–Ne laser light is approximately 633 nm.

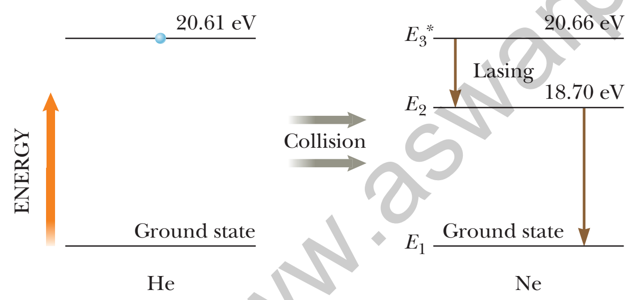


Figure P42.58

59. The carbon dioxide laser is one of the most powerful developed. The energy difference between the two laser levels is 0.117 eV. Determine (a) the frequency and (b) the wavelength of the radiation emitted by this laser. (c) In what portion of the electromagnetic spectrum is this radiation?

60. Review. A helium–neon laser can produce a green laser beam instead of a red one. Figure P42.60 shows the transitions involved to form the red beam and the green beam. After a population inversion is established, neon atoms make a variety of downward transitions in falling from the state labeled $E_4{}^*$ down eventually to level E_1 (arbitrarily assigned the energy $E_1 = 0$). The atoms emit both red light with a wavelength of 632.8 nm in a transition $E_4{}^* - E_3$ and green light with a wavelength of 543 nm in a competing transition $E_4{}^* - E_2$. (a) What is the energy E_2? Assume the atoms are in a cavity between mirrors designed to reflect the green light with high efficiency but to allow the red light to leave the cavity immediately. Then stimulated emission can lead to the buildup of a collimated beam of green light between the mirrors having a greater intensity than that of the red light. To constitute the radiated laser beam, a small fraction of the green light is permitted to escape by transmission through one mirror. The mirrors forming the resonant cavity can be made of layers of silicon dioxide (index of refraction $n = 1.458$) and titanium dioxide (index of refraction varies between 1.9 and 2.6). (b) How thick a layer of silicon dioxide, between layers of titanium dioxide, would minimize reflection of the red light? (c) What should be the thickness of a similar but separate layer of silicon dioxide to maximize reflection of the green light?

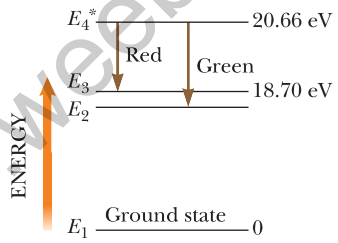


Figure P42.60 Problems 60 and 62.

61. M A ruby laser delivers a 10.0-ns pulse of 1.00-MW average power. If the photons have a wavelength of 694.3 nm, how many are contained in the pulse?

62. The number N of atoms in a particular state is called the population of that state. This number depends on the energy of that state and the temperature. In thermal equilibrium, the population of atoms in a state of energy E_n is given by a Boltzmann distribution expression

$$N = N_g e^{-(E_n - E_g)/k_B T}$$

where N_g is the population of the ground state of energy E_g, k_B is Boltzmann's constant, and T is the absolute temperature. For simplicity, assume each energy level has only one quantum state associated with it. (a) Before the power is switched on, the neon atoms in a laser are in thermal equilibrium at 27.0°C. Find the equilibrium ratio of the populations of the states $E_4{}^*$ and E_3 shown for the red transition in Figure P42.60. Lasers operate by a clever artificial production of a "population inversion" between the upper and lower atomic energy states involved in the lasing transition. This term means that more atoms are in the upper excited state than in the lower one. Consider the $E_4{}^*-E_3$ transition in Figure P42.60. Assume 2% more atoms occur in the upper state than in the lower. (b) To demonstrate how unnatural such a situation is, find the temperature for which the Boltzmann distribution describes a 2.00% population inversion. (c) Why does such a situation not occur naturally?

63. M W A neodymium–yttrium–aluminum garnet laser used in eye surgery emits a 3.00-mJ pulse in 1.00 ns, focused to a spot 30.0 μm in diameter on the retina. (a) Find (in SI units) the power per unit area at the retina. (In the optics industry, this quantity is called the *irradiance.*) (b) What energy is delivered by the pulse to an area of molecular size, taken as a circular area 0.600 nm in diameter?

64. AMT **Review.** Figure 42.29 represents the light bouncing between two mirrors in a laser cavity as two traveling waves. These traveling waves moving in opposite directions constitute a standing wave. If the reflecting surfaces are metallic films, the electric field has nodes at both ends. The electromagnetic standing wave is analogous to the standing string wave represented in Figure 18.10. (a) Assume that a helium–neon laser has precisely flat and parallel mirrors 35.124 103 cm apart. Assume that the active medium can efficiently amplify only light with wavelengths between 632.808 40 nm and 632.809 80 nm. Find the number of components that constitute the laser light, and the wavelength of each component, precise to eight digits. (b) Find the root-mean-square speed for a neon atom at 120°C. (c) Show that at this temperature the Doppler effect for light emission by moving neon atoms should realistically make the bandwidth of the light amplifier larger than the 0.001 40 nm assumed in part (a).

Additional Problems

65. How much energy is required to ionize a hydrogen atom when it is in (a) the $n = 2$ state and (b) the $n = 10$ state?

66. The force on a magnetic moment μ_z in a nonuniform magnetic field B_z is given by $F_z = \mu_z(dB_z/dz)$. If a beam of silver atoms travels a horizontal distance of 1.00 m through such a field and each atom has a speed of 100 m/s, how strong must be the field gradient dB_z/dz to deflect the beam 1.00 mm?

67. Suppose a hydrogen atom is in the 2*s* state, with its wave function given by Equation 42.26. Taking $r = a_0$, calculate values for (a) $\psi_{2s}(a_0)$, (b) $|\psi_{2s}(a_0)|^2$, and (c) $P_{2s}(a_0)$.

68. W **Review.** (a) How much energy is required to cause an electron in hydrogen to move from the $n = 1$ state to the $n = 2$ state? (b) Suppose the atom gains this energy through collisions among hydrogen atoms at a high temperature. At what temperature would the average atomic kinetic energy $\frac{3}{2}k_BT$ be great enough to excite the electron? Here k_B is Boltzmann's constant.

69. M In the technique known as electron spin resonance (ESR), a sample containing unpaired electrons is placed in a magnetic field. Consider a situation in which a single electron (*not* contained in an atom) is immersed in a magnetic field. In this simple situation, only two energy states are possible, corresponding to $m_s = \pm\frac{1}{2}$. In ESR, the absorption of a photon causes the electron's spin magnetic moment to flip from the lower energy state to the higher energy state. According to Section 29.5, the change in energy is $2\mu_B B$. (The lower energy state corresponds to the case in which the z component of the magnetic moment $\vec{\boldsymbol{\mu}}_{\text{spin}}$ is aligned with the magnetic field, and the higher energy state corresponds to the case in which the z component of $\vec{\boldsymbol{\mu}}_{\text{spin}}$ is aligned opposite to the field.) What is the photon frequency required to excite an ESR transition in a 0.350-T magnetic field?

70. An electron in chromium moves from the $n = 2$ state to the $n = 1$ state without emitting a photon. Instead, the excess energy is transferred to an outer electron (one in the $n = 4$ state), which is then ejected by the atom. In this Auger (pronounced "ohjay") process, the ejected electron is referred to as an Auger electron. Use the Bohr theory to find the kinetic energy of the Auger electron.

71. The states of matter are solid, liquid, gas, and plasma. Plasma can be described as a gas of charged particles or a gas of ionized atoms. Most of the matter in the Solar System is plasma (throughout the interior of the Sun). In fact, most of the matter in the Universe is plasma; so is a candle flame. Use the information in Figure 42.20 to make an order-of-magnitude estimate for the temperature to which a typical chemical element must be raised to turn into plasma by ionizing most of the atoms in a sample. Explain your reasoning.

72. Show that the wave function for a hydrogen atom in the 2*s* state

$$\psi_{2s}(r) = \frac{1}{4\sqrt{2\pi}}\left(\frac{1}{a_0}\right)^{3/2}\left(2 - \frac{r}{a_0}\right)e^{-r/2a_0}$$

satisfies the spherically symmetric Schrödinger equation given in Problem 26.

73. Find the average (expectation) value of $1/r$ in the 1*s* state of hydrogen. Note that the general expression is given by

$$\langle 1/r\rangle = \int_{\text{all space}} |\psi|^2(1/r)\,dV = \int_0^\infty P(r)(1/r)\,dr$$

Is the result equal to the inverse of the average value of r?

74. *Why is the following situation impossible?* An experiment is performed on an atom. Measurements of the atom when it is in a particular excited state show five possible values of the z component of orbital angular momentum, ranging between 3.16×10^{-34} kg · m^2/s and -3.16×10^{-34} kg · m^2/s.

75. In the Bohr model of the hydrogen atom, an electron travels in a circular path. Consider another case in which an electron travels in a circular path: a single electron moving perpendicular to a magnetic field $\vec{\mathbf{B}}$. Lev Davidovich Landau (1908–1968) solved the Schrödinger equation for such an electron. The electron can be considered as a model atom without a nucleus or as the irreducible quantum limit of the cyclotron. Landau proved its energy is quantized in uniform steps of $e\hbar B/m_e$. In 1999, a single electron was trapped by a Harvard University research team in an evacuated centimeter-size metal can cooled to a temperature of 80 mK. In a magnetic field of magnitude

5.26 T, the electron circulated for hours in its lowest energy level. (a) Evaluate the size of a quantum jump in the electron's energy. (b) For comparison, evaluate $k_B T$ as a measure of the energy available to the electron in blackbody radiation from the walls of its container. Microwave radiation was introduced to excite the electron. Calculate (c) the frequency and (d) the wavelength of the photon the electron absorbed as it jumped to its second energy level. Measurement of the resonant absorption frequency verified the theory and permitted precise determination of properties of the electron.

76. As the Earth moves around the Sun, its orbits are quantized. (a) Follow the steps of Bohr's analysis of the hydrogen atom to show that the allowed radii of the Earth's orbit are given by

$$r = \frac{n^2\hbar^2}{GM_S M_E^{\,2}}$$

where n is an integer quantum number, M_S is the mass of the Sun, and M_E is the mass of the Earth. (b) Calculate the numerical value of n for the Sun–Earth system. (c) Find the distance between the orbit for quantum number n and the next orbit out from the Sun corresponding to the quantum number $n + 1$. (d) Discuss the significance of your results from parts (b) and (c).

77. An elementary theorem in statistics states that the root-mean-square uncertainty in a quantity r is given by $\Delta r = \sqrt{\langle r^2\rangle - \langle r\rangle^2}$. Determine the uncertainty in the radial position of the electron in the ground state of the hydrogen atom. Use the average value of r found in Example 42.3: $\langle r\rangle = 3a_0/2$. The average value of the squared distance between the electron and the proton is given by

$$\langle r^2\rangle = \int_{\text{all space}} |\psi|^2 r^2\, dV = \int_0^\infty P(r) r^2\, dr$$

78. Example 42.3 calculates the most probable value and the average value for the radial coordinate r of the electron in the ground state of a hydrogen atom. For comparison with these modal and mean values, find the median value of r. Proceed as follows. (a) Derive an expression for the probability, as a function of r, that the electron in the ground state of hydrogen will be found outside a sphere of radius r centered on the nucleus. (b) Make a graph of the probability as a function of r/a_0. Choose values of r/a_0 ranging from 0 to 4.00 in steps of 0.250. (c) Find the value of r for which the probability of finding the electron outside a sphere of radius r is equal to the probability of finding the electron inside this sphere. You must solve a transcendental equation numerically, and your graph is a good starting point.

79. AMT M (a) For a hydrogen atom making a transition from the $n = 4$ state to the $n = 2$ state, determine the wavelength of the photon created in the process. (b) Assuming the atom was initially at rest, determine the recoil speed of the hydrogen atom when it emits this photon.

80. Astronomers observe a series of spectral lines in the light from a distant galaxy. On the hypothesis that the lines form the Lyman series for a (new?) one-electron atom, they start to construct the energy-level diagram shown in Figure P42.80, which gives the wavelengths of the first four lines and the short-wavelength limit of this series. Based on this information, calculate (a) the energies of the ground state and first four excited states for this one-electron atom and (b) the wavelengths of the first three lines and the short-wavelength limit in the Balmer series for this atom. (c) Show that the wavelengths of the first four lines and the short-wavelength limit of the Lyman series for the hydrogen atom are all 60.0% of the wavelengths for the Lyman series in the one-electron atom in the distant galaxy. (d) Based on this observation, explain why this atom could be hydrogen.

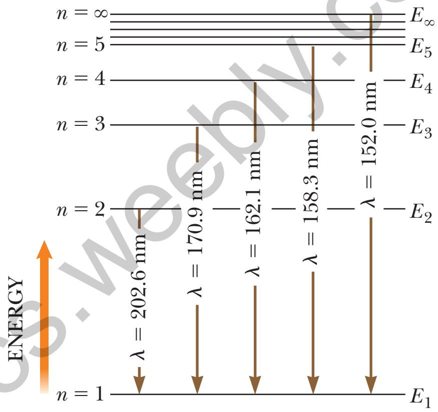


Figure P42.80

81. GP We wish to show that the most probable radial position for an electron in the 2s state of hydrogen is $r = 5.236a_0$. (a) Use Equations 42.24 and 42.26 to find the radial probability density for the 2s state of hydrogen. (b) Calculate the derivative of the radial probability density with respect to r. (c) Set the derivative in part (b) equal to zero and identify three values of r that represent minima in the function. (d) Find two values of r that represent maxima in the function. (e) Identify which of the values in part (c) represents the highest probability.

82. All atoms have the same size, to an order of magnitude. (a) To demonstrate this fact, estimate the atomic diameters for aluminum (with molar mass 27.0 g/mol and density 2.70 g/cm^3) and uranium (molar mass 238 g/mol and density 18.9 g/cm^3). (b) What do the results of part (a) imply about the wave functions for inner-shell electrons as we progress to higher and higher atomic mass atoms?

83. A pulsed ruby laser emits light at 694.3 nm. For a 14.0-ps pulse containing 3.00 J of energy, find (a) the physical length of the pulse as it travels through space and (b) the number of photons in it. (c) The beam has a circular cross section of diameter 0.600 cm. Find the number of photons per cubic millimeter.

84. A pulsed laser emits light of wavelength λ. For a pulse of duration Δt having energy T_{ER}, find (a) the physical length of the pulse as it travels through space and (b) the number of photons in it. (c) The beam has a circular cross section having diameter d. Find the number of photons per unit volume.

85. Assume three identical uncharged particles of mass m and spin $\frac{1}{2}$ are contained in a one-dimensional box of length L. What is the ground-state energy of this system?

86. Suppose the ionization energy of an atom is 4.10 eV. In the spectrum of this same atom, we observe emission lines with wavelengths 310 nm, 400 nm, and 1 377.8 nm. Use this information to construct the energy-level diagram with the fewest levels. Assume the higher levels are closer together.

87. For hydrogen in the 1s state, what is the probability of finding the electron farther than $2.50a_0$ from the nucleus?

88. For hydrogen in the 1s state, what is the probability of finding the electron farther than βa_0 from the nucleus, where β is an arbitrary number?

Challenge Problems

89. The positron is the antiparticle to the electron. It has the same mass and a positive electric charge of the same magnitude as that of the electron. Positronium is a hydrogen-like atom consisting of a positron and an electron revolving around each other. Using the Bohr model, find (a) the allowed distances between the two particles and (b) the allowed energies of the system.

90. Review. Steven Chu, Claude Cohen-Tannoudji, and William Phillips received the 1997 Nobel Prize in Physics for "the development of methods to cool and trap atoms with laser light." One part of their work was with a beam of atoms (mass $\sim 10^{-25}$ kg) that move at a speed on the order of 1 km/s, similar to the speed of molecules in air at room temperature. An intense laser light beam tuned to a visible atomic transition (assume 500 nm) is directed straight into the atomic beam; that is, the atomic beam and the light beam are traveling in opposite directions. An atom in the ground state immediately absorbs a photon. Total system momentum is conserved in the absorption process. After a lifetime on the order of 10^{-8} s, the excited atom radiates by spontaneous emission. It has an equal probability of emitting a photon in any direction. Therefore, the average "recoil" of the atom is zero over many absorption and emission cycles. (a) Estimate the average deceleration of the atomic beam. (b) What is the order of magnitude of the distance over which the atoms in the beam are brought to a halt?

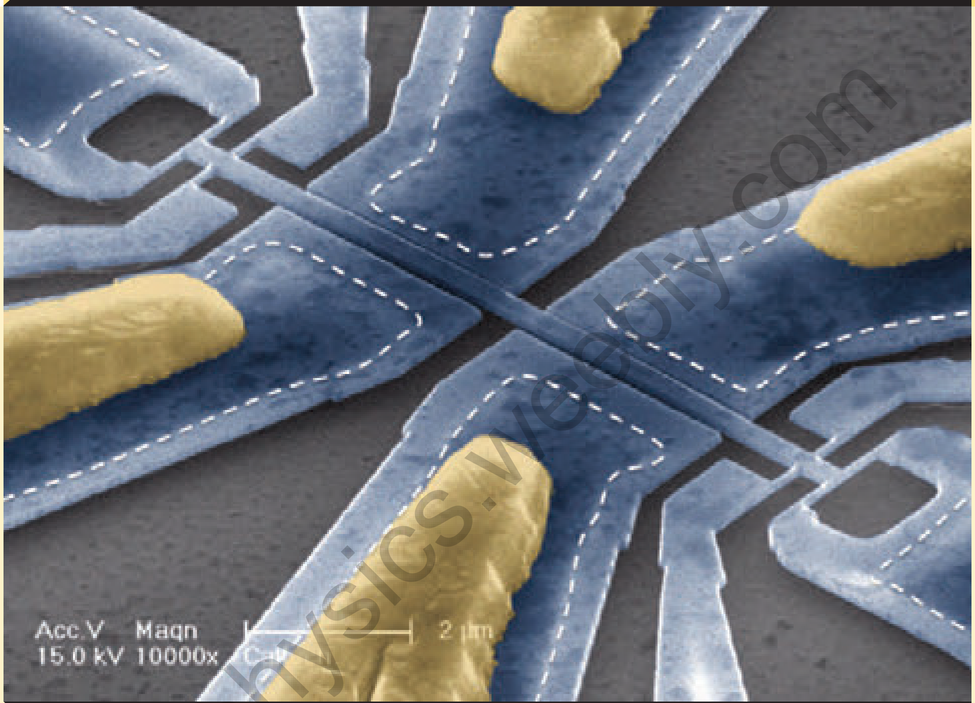
91. (a) Use Bohr's model of the hydrogen atom to show that when the electron moves from the n state to the $n-1$ state, the frequency of the emitted light is

$$f = \left(\frac{2\pi^2 m_e k_e^2 e^4}{h^3}\right)\frac{2n-1}{n^2(n-1)^2}$$

(b) Bohr's correspondence principle claims that quantum results should reduce to classical results in the limit of large quantum numbers. Show that as $n \to \infty$, this expression varies as $1/n^3$ and reduces to the classical frequency one expects the atom to emit. *Suggestion:* To calculate the classical frequency, note that the frequency of revolution is $v/2\pi r$, where v is the speed of the electron and r is given by Equation 42.10.

CHAPTER

43 Molecules and Solids



The photograph shows a *NEMS resonator,* where NEMS is an acronym for *nanoelectromechanical system.* The device employs a semiconductor bridge vibrating in a standing wave like the strings in Chapter 18. When a single molecule or other particle adheres to the bridge, the resonance frequencies of the normal modes shift in a measurable way. Scientists can determine the mass of the particle from the shifts in the frequencies. The new device shows promise in allowing the masses of molecules and many biological particles to be measured with great accuracy. *(Caltech/Scott Kelberg and Michael Roukes)*

The most random atomic arrangement, that of a gas, was well understood in the 1800s as discussed in our study of kinetic theory in Chapter 21. In a crystalline solid, the atoms are not randomly arranged; rather, they form a regular array. The symmetry of the arrangement of atoms both stimulated and allowed rapid progress in the field of solid-state physics in the 20th century. Recently, our understanding of liquids and amorphous solids has advanced. (In an amorphous solid such as glass or paraffin, the atoms do not form a regular array.) The recent interest in the physics of low-cost amorphous materials has been driven by their use in such devices as solar cells, memory elements, and fiber-optic waveguides. With the addition of liquids, amorphous solids, and some more exotic forms of matter, such as Bose–Einstein condensates, solid-state physics expanded in the middle of the 20th century to become known as *condensed matter physics.*

We begin this chapter by studying the aggregates of atoms known as molecules. We describe the bonding mechanisms in molecules, the various modes of molecular excitation, and the radiation emitted or absorbed by molecules. Next, we show how molecules combine to form solids. Then, by examining their energy-level structure, we explain the differences between insulating, conducting, semiconducting, and superconducting materials. The chapter also includes discussions of semiconducting junctions and several semiconductor devices.

43.1 Molecular Bonds

The bonding mechanisms in a molecule are fundamentally due to electric forces between atoms (or ions). Because the electric force is conservative, the forces between atoms in the system of a molecule are related to a potential energy function. A stable molecule is expected at a configuration for which the potential energy function for the molecule has its minimum value. (See Section 7.9.)

A potential energy function that can be used to model a molecule should account for two known features of molecular bonding:

1. The force between atoms is repulsive at very small separation distances. When two atoms are brought close to each other, some of their electron shells overlap, resulting in repulsion between the shells. This repulsion is partly electrostatic in origin and partly the result of the exclusion principle. Because all electrons must obey the exclusion principle, some electrons in the overlapping shells are forced into higher energy states and the system energy increases as if a repulsive force existed between the atoms.
2. At somewhat larger separations, the force between atoms is attractive. If that were not true, the atoms in a molecule would not be bound together.

Taking into account these two features, the potential energy for a system of two atoms can be represented by an expression of the form

$$U(r) = -\frac{A}{r^n} + \frac{B}{r^m} \tag{43.1}$$

where r is the internuclear separation distance between the two atoms and n and m are small integers. The parameter A is associated with the attractive force and B with the repulsive force. Example 7.9 gives one common model for such a potential energy function, the Lennard–Jones potential.

Potential energy versus internuclear separation distance for a two-atom system is graphed in Figure 43.1. At large separation distances between the two atoms, the slope of the curve is positive, corresponding to a net attractive force. At the equilibrium separation distance, the attractive and repulsive forces just balance. At this point, the potential energy has its minimum value and the slope of the curve is zero.

A complete description of the bonding mechanisms in molecules is highly complex because bonding involves the mutual interactions of many particles. In this section, we discuss only some simplified models.

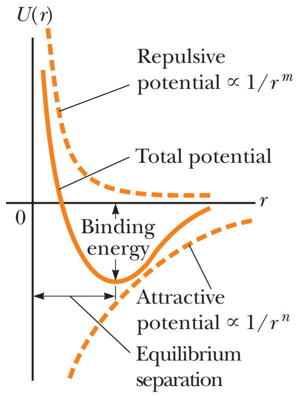


Figure 43.1 Total potential energy as a function of internuclear separation distance for a system of two atoms.

Ionic Bonding

When two atoms combine in such a way that one or more outer electrons are transferred from one atom to the other, the bond formed is called an **ionic bond.** Ionic bonds are fundamentally caused by the Coulomb attraction between oppositely charged ions.

A familiar example of an ionically bonded solid is sodium chloride, NaCl, which is common table salt. Sodium, which has the electronic configuration $1s^22s^22p^63s^1$, is ionized relatively easily, giving up its $3s$ electron to form a Na^+ ion. The energy required to ionize the atom to form Na^+ is 5.1 eV. Chlorine, which has the electronic configuration $1s^22s^22p^5$, is one electron short of the filled-shell structure of argon. If we compare the energy of a system of a free electron and a Cl atom with one in which the electron joins the atom to make the Cl^- ion, we find that the energy of the ion is lower. When the electron makes a transition from the $E = 0$ state to the negative energy state associated with the available shell in the atom, energy is released. This amount of energy is called the **electron affinity** of the atom. For chlorine, the electron affinity is 3.6 eV. Therefore, the energy required to form Na^+ and Cl^- from isolated atoms is $5.1 - 3.6 = 1.5$ eV. It costs 5.1 eV to remove

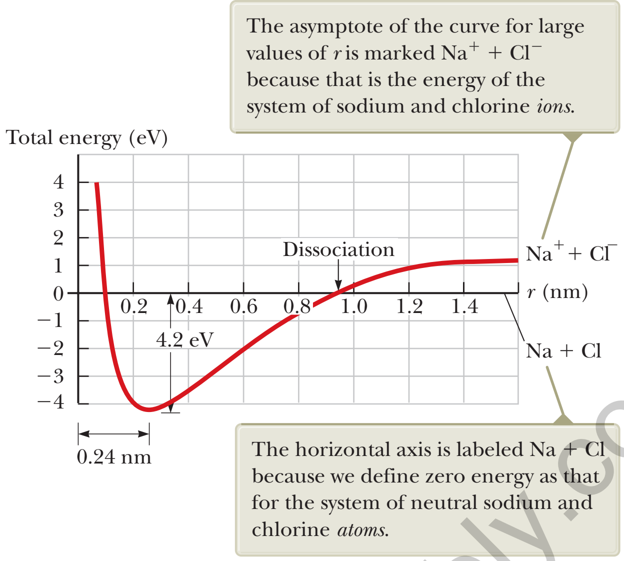


Figure 43.2 Total energy versus internuclear separation distance for Na^+ and Cl^- ions.

Pitfall Prevention 43.1

Ionic and Covalent Bonds In practice, these descriptions of ionic and covalent bonds represent extreme ends of a spectrum of bonds involving electron transfer. In a real bond, the electron may not be *completely* transferred as in an ionic bond or *equally* shared as in a covalent bond. Therefore, real bonds lie somewhere between these extremes.

the electron from the Na atom, but 3.6 eV of it is gained back when that electron is allowed to join with the Cl atom.

Now imagine that these two charged ions interact with one another to form a NaCl "molecule."[1] The total energy of the NaCl molecule versus internuclear separation distance is graphed in Figure 43.2. At very large separation distances, the energy of the system of ions is 1.5 eV as calculated above. The total energy has a minimum value of −4.2 eV at the equilibrium separation distance, which is approximately 0.24 nm. Hence, the energy required to break the Na^+−Cl^- bond and form neutral sodium and chlorine atoms, called the **dissociation energy,** is 4.2 eV. The energy of the molecule is lower than that of the system of two neutral atoms. Consequently, it is **energetically favorable** for the molecule to form: if a lower energy state of a system exists, the system tends to emit energy to achieve this lower energy state. The system of neutral sodium and chlorine atoms can reduce its total energy by transferring energy out of the system (by electromagnetic radiation, for example) and forming the NaCl molecule.

Covalent Bonding

A **covalent bond** between two atoms is one in which electrons supplied by either one or both atoms are shared by the two atoms. Many diatomic molecules—such as H_2, F_2, and CO—owe their stability to covalent bonds. The bond between two hydrogen atoms can be described by using atomic wave functions. The ground-state wave function for a hydrogen atom (Chapter 42) is

$$\psi_{1s}(r) = \frac{1}{\sqrt{\pi a_0^{\,3}}}\, e^{-r/a_0}$$

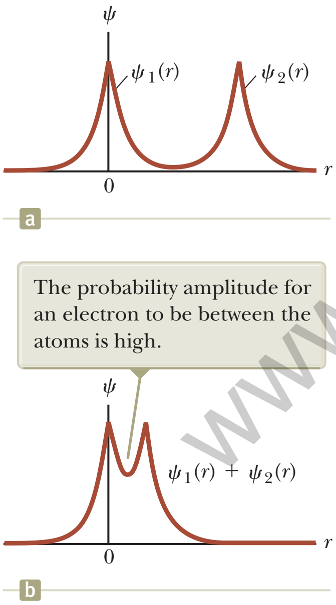


Figure 43.3 Ground-state wave functions $\psi_1(r)$ and $\psi_2(r)$ for two atoms making a covalent bond. (a) The atoms are far apart, and their wave functions overlap minimally. (b) The atoms are close together, forming a composite wave function $\psi_1(r) + \psi_2(r)$ for the system.

This wave function is graphed in Figure 43.3a for two hydrogen atoms that are far apart. There is very little overlap of the wave functions $\psi_1(r)$ for atom 1, located at $r = 0$, and $\psi_2(r)$ for atom 2, located some distance away. Suppose now the two atoms are brought close together. As that happens, their wave functions overlap and form the compound wave function $\psi_1(r) + \psi_2(r)$ shown in Figure 43.3b. Notice that the probability amplitude is larger between the atoms than it is on either side of the combination of atoms. As a result, the probability is higher that the electrons associated with the atoms will be located between the atoms than on the outer regions

[1]NaCl does not tend to form as an isolated molecule at room temperature. In the solid state, NaCl forms a crystalline array of ions as described in Section 43.3. In the liquid state or in solution with water, the Na^+ and Cl^- ions dissociate and are free to move relative to each other.

of the system. Consequently, the average position of negative charge in the system is halfway between the atoms. This scenario can be modeled as if there were a fixed negative charge between the atoms, exerting attractive Coulomb forces on both nuclei. Therefore, there is an overall attractive force between the atoms, resulting in a covalent bond.

Because of the exclusion principle, the two electrons in the ground state of H_2 must have antiparallel spins. Also because of the exclusion principle, if a third H atom is brought near the H_2 molecule, the third electron would have to occupy a higher energy level, which is not an energetically favorable situation. For this reason, the H_3 molecule is not stable and does not form.

Van der Waals Bonding

Ionic and covalent bonds occur between atoms to form molecules or ionic solids, so they can be described as bonds *within* molecules. Two additional types of bonds, van der Waals bonds and hydrogen bonds, can occur *between* molecules.

You might think that two neutral molecules would not interact by means of the electric force because they each have zero net charge. They are attracted to each other, however, by weak electrostatic forces called **van der Waals forces.** Likewise, atoms that do not form ionic or covalent bonds are attracted to each other by van der Waals forces. Noble gas atoms, for example, because of their filled shell structure, do not generally form molecules or bond to each other to form a liquid. Because of van der Waals forces, however, at sufficiently low temperatures at which thermal excitations are negligible, noble gases first condense to liquids and then solidify. (The exception is helium, which does not solidify at atmospheric pressure.)

The van der Waals force results from the following situation. While being electrically neutral, a molecule has a charge distribution with positive and negative centers at different positions in the molecule. As a result, the molecule may act as an electric dipole. (See Section 23.4.) Because of the dipole electric fields, two molecules can interact such that there is an attractive force between them.

There are three types of van der Waals forces. The first type, called the *dipole–dipole force,* is an interaction between two molecules each having a permanent electric dipole moment. For example, polar molecules such as HCl have permanent electric dipole moments and attract other polar molecules.

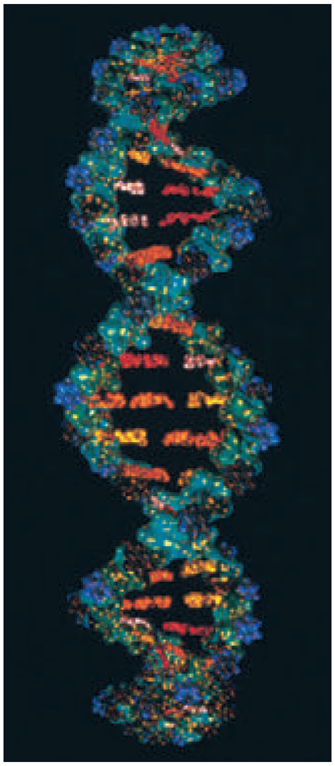
The second type, the *dipole–induced dipole force,* results when a polar molecule having a permanent electric dipole moment induces a dipole moment in a nonpolar molecule. In this case, the electric field of the polar molecule creates the dipole moment in the nonpolar molecule, which then results in an attractive force between the molecules.

The third type is called the *dispersion force,* an attractive force that occurs between two nonpolar molecules. In this case, although the average dipole moment of a nonpolar molecule is zero, the average of the square of the dipole moment is nonzero because of charge fluctuations. Two nonpolar molecules near each other tend to have dipole moments that are correlated in time so as to produce an attractive van der Waals force.

Hydrogen Bonding

Because hydrogen has only one electron, it is expected to form a covalent bond with only one other atom within a molecule. A hydrogen atom in a given molecule can also form a second type of bond between molecules called a **hydrogen bond.** Let's use the water molecule H_2O as an example. In the two covalent bonds in this molecule, the electrons from the hydrogen atoms are more likely to be found near the oxygen atom than near the hydrogen atoms, leaving essentially bare protons at the positions of the hydrogen atoms. This unshielded positive charge can be attracted to the negative end of another polar molecule. Because the proton is unshielded by electrons, the negative end of the other molecule can come very close to the proton to form a bond strong enough to form a solid crystalline structure, such as that of ordinary ice. The bonds within a water molecule are covalent, but the bonds between water molecules in ice are hydrogen bonds.

The hydrogen bond is relatively weak compared with other chemical bonds and can be broken with an input energy of approximately 0.1 eV. Because of this weakness, ice melts at the low temperature of 0°C. Even though this bond is weak, however, hydrogen bonding is a critical mechanism responsible for the linking of biological molecules and polymers. For example, in the case of the DNA (deoxyribonucleic acid) molecule, which has a double-helix structure (Fig. 43.4), hydrogen bonds form by the sharing of a proton between two atoms and create linkages between the turns of the helix.



Douglas Struthers/Stone/Getty Images

Figure 43.4 DNA molecules are held together by hydrogen bonds.

Quick Quiz 43.1 For each of the following atoms or molecules, identify the most likely type of bonding that occurs between the atoms or between the molecules. Choose from the following list: ionic, covalent, van der Waals, hydrogen. **(a)** atoms of krypton **(b)** potassium and chlorine atoms **(c)** hydrogen fluoride (HF) molecules **(d)** chlorine and oxygen atoms in a hypochlorite ion (ClO^-)

43.2 Energy States and Spectra of Molecules

Consider an individual molecule in the gaseous phase of a substance. The energy E of the molecule can be divided into four categories: (1) electronic energy, due to the interactions between the molecule's electrons and nuclei; (2) translational energy, due to the motion of the molecule's center of mass through space; (3) rotational energy, due to the rotation of the molecule about its center of mass; and (4) vibrational energy, due to the vibration of the molecule's constituent atoms:

Total energy of a molecule ▶

$$E = E_{\text{el}} + E_{\text{trans}} + E_{\text{rot}} + E_{\text{vib}}$$

We explored the roles of translational, rotational, and vibrational energy of molecules in determining the molar specific heats of gases in Sections 21.2 and 21.3. The translational energy is important in kinetic theory, but it is unrelated to internal structure of the molecule, so this molecular energy is unimportant in interpreting molecular spectra. The electronic energy of a molecule is very complex because it involves the interaction of many charged particles, but various techniques have been developed to approximate its values. Although the electronic energies can be studied, significant information about a molecule can be determined by analyzing its quantized rotational and vibrational energy states. Transitions between these states give spectral lines in the microwave and infrared regions of the electromagnetic spectrum, respectively.

Rotational Motion of Molecules

Let's consider the rotation of a molecule around its center of mass, confining our discussion to the diatomic molecule (Fig. 43.5a) but noting that the same ideas can be extended to polyatomic molecules. A diatomic molecule aligned along a y axis has only two rotational degrees of freedom, corresponding to rotations about the x and z axes passing through the molecule's center of mass. We discussed the rotation of such a molecule and its contribution to the specific heat of a gas in Section 21.3. If ω is the angular frequency of rotation about one of these axes, the rotational kinetic energy of the molecule about that axis can be expressed with Equation 10.24:

$$E_{\text{rot}} = \tfrac{1}{2}I\omega^2 \qquad \textbf{(43.2)}$$

In this equation, I is the moment of inertia of the molecule about its center of mass, given by

Moment of inertia for a diatomic molecule ▶

$$I = \left(\frac{m_1 m_2}{m_1 + m_2}\right)r^2 = \mu r^2 \qquad \textbf{(43.3)}$$

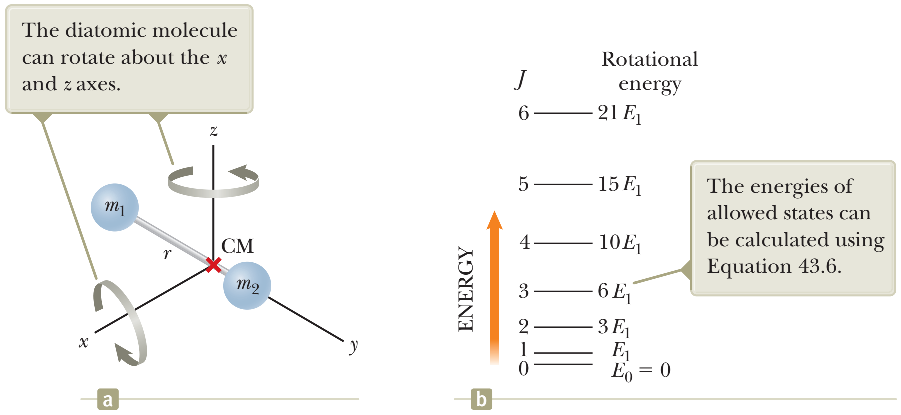


Figure 43.5 Rotation of a diatomic molecule around its center of mass. (a) A diatomic molecule oriented along the y axis. (b) Allowed rotational energies of a diatomic molecule expressed as multiples of $E_1 = \hbar^2/I$.

where m_1 and m_2 are the masses of the atoms that form the molecule, r is the atomic separation, and μ is the **reduced mass** of the molecule (see Example 41.5 and Problem 40 in Chapter 41):

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \qquad \textbf{(43.4)}$$

◀ **Reduced mass of a diatomic molecule**

The magnitude of the molecule's angular momentum about its center of mass is given by Equation 11.14, $L = I\omega$, which classically can have any value. Quantum mechanics, however, restricts the molecule to certain quantized rotational frequencies such that the angular momentum of the molecule has the values[2]

$$L = \sqrt{J(J+1)}\,\hbar \qquad J = 0, 1, 2, \ldots \qquad \textbf{(43.5)}$$

◀ **Allowed values of rotational angular momentum**

where J is an integer called the **rotational quantum number.** Combining Equations 43.5 and 43.2, we obtain an expression for the allowed values of the rotational kinetic energy of the molecule:

$$E_{\text{rot}} = \tfrac{1}{2}I\omega^2 = \frac{1}{2I}(I\omega)^2 = \frac{L^2}{2I} = \frac{(\sqrt{J(J+1)}\,\hbar)^2}{2I}$$

$$E_{\text{rot}} = E_J = \frac{\hbar^2}{2I}J(J+1) \qquad J = 0, 1, 2, \ldots \qquad \textbf{(43.6)}$$

◀ **Allowed values of rotational energy**

The allowed rotational energies of a diatomic molecule are plotted in Figure 43.5b. As the quantum number J goes up, the states become farther apart as displayed earlier for rotational energy levels in Figure 21.7.

For most molecules, transitions between adjacent rotational energy levels result in radiation that lies in the microwave range of frequencies ($f \sim 10^{11}$ Hz). When a molecule absorbs a microwave photon, the molecule jumps from a lower rotational energy level to a higher one. The allowed rotational transitions of linear molecules are regulated by the selection rule $\Delta J = \pm 1$. Given this selection rule, all absorption lines in the spectrum of a linear molecule correspond to energy separations equal to $E_J - E_{J-1}$, where $J = 1, 2, 3, \ldots$. From Equation 43.6, we see that the energies of the absorbed photons are given by

$$E_{\text{photon}} = \Delta E_{\text{rot}} = E_J - E_{J-1} = \frac{\hbar^2}{2I}[J(J+1) - (J-1)J]$$

$$E_{\text{photon}} = \frac{\hbar^2}{I}J = \frac{h^2}{4\pi^2 I}J \qquad J = 1, 2, 3, \ldots \qquad \textbf{(43.7)}$$

◀ **Energy of a photon absorbed in a transition between adjacent rotational levels**

[2]Equation 43.5 is similar to Equation 42.27 for orbital angular momentum in an atom. The relationship between the magnitude of the angular momentum of a system and the associated quantum number is the same as it is in these equations for *any* system that exhibits rotation as long as the potential energy function for the system is spherically symmetric.

where J is the rotational quantum number of the higher energy state. Because $E_{photon} = hf$, where f is the frequency of the absorbed photon, we see that the allowed frequency for the transition $J = 0$ to $J = 1$ is $f_1 = h/4\pi^2 I$. The frequency corresponding to the $J = 1$ to $J = 2$ transition is $2f_1$, and so on. These predictions are in excellent agreement with the observed frequencies.

Quick Quiz 43.2 A gas of identical diatomic molecules absorbs electromagnetic radiation over a wide range of frequencies. Molecule 1 is in the $J = 0$ rotation state and makes a transition to the $J = 1$ state. Molecule 2 is in the $J = 2$ state and makes a transition to the $J = 3$ state. Is the ratio of the frequency of the photon that excited molecule 2 to that of the photon that excited molecule 1 equal to **(a)** 1, **(b)** 2, **(c)** 3, **(d)** 4, or **(e)** impossible to determine?

Example 43.1 Rotation of the CO Molecule

The $J = 0$ to $J = 1$ rotational transition of the CO molecule occurs at a frequency of 1.15×10^{11} Hz.

(A) Use this information to calculate the moment of inertia of the molecule.

SOLUTION

Conceptualize Imagine that the two atoms in Figure 43.5a are carbon and oxygen. The center of mass of the molecule is not midway between the atoms because of the difference in masses of the C and O atoms.

Categorize The statement of the problem tells us to categorize this example as one involving a quantum-mechanical treatment and to restrict our investigation to the rotational motion of a diatomic molecule.

Analyze Use Equation 43.7 to find the energy of a photon that excites the molecule from the $J = 0$ to the $J = 1$ rotational level:

$$E_{photon} = \frac{h^2}{4\pi^2 I}(1) = \frac{h^2}{4\pi^2 I}$$

Equate this energy to $E = hf$ for the absorbed photon and solve for I:

$$\frac{h^2}{4\pi^2 I} = hf \rightarrow I = \frac{h}{4\pi^2 f}$$

Substitute the frequency given in the problem statement:

$$I = \frac{6.626 \times 10^{-34}\ \mathrm{J \cdot s}}{4\pi^2(1.15 \times 10^{11}\ \mathrm{s^{-1}})} = \boxed{1.46 \times 10^{-46}\ \mathrm{kg \cdot m^2}}$$

(B) Calculate the bond length of the molecule.

SOLUTION

Find the reduced mass μ of the CO molecule:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} = \frac{(12\ \mathrm{u})(16\ \mathrm{u})}{12\ \mathrm{u} + 16\ \mathrm{u}} = 6.86\ \mathrm{u}$$

$$= (6.86\ \mathrm{u})\left(\frac{1.66 \times 10^{-27}\ \mathrm{kg}}{1\ \mathrm{u}}\right) = 1.14 \times 10^{-26}\ \mathrm{kg}$$

Solve Equation 43.3 for r and substitute for the reduced mass and the moment of inertia from part (A):

$$r = \sqrt{\frac{I}{\mu}} = \sqrt{\frac{1.46 \times 10^{-46}\ \mathrm{kg \cdot m^2}}{1.14 \times 10^{-26}\ \mathrm{kg}}}$$

$$= 1.13 \times 10^{-10}\ \mathrm{m} = \boxed{0.113\ \mathrm{nm}}$$

Finalize The moment of inertia of the molecule and the separation distance between the atoms are both very small, as expected for a microscopic system.

WHAT IF? What if another photon of frequency 1.15×10^{11} Hz is incident on the CO molecule while that molecule is in the $J = 1$ state? What happens?

Answer Because the rotational quantum states are not equally spaced in energy, the $J = 1$ to $J = 2$ transition does not have the same energy as the $J = 0$ to $J = 1$ transition. Therefore, the molecule will *not* be excited to the $J = 2$ state. Two

43.1 continued

possibilities exist. The photon could pass by the molecule with no interaction, or the photon could induce a stimulated emission, similar to that for atoms and discussed in Section 42.9. In this case, the molecule makes a transition back to the $J = 0$ state and the original photon and a second identical photon leave the scene of the interaction.

Vibrational Motion of Molecules

If we consider a molecule to be a flexible structure in which the atoms are bonded together by "effective springs" as shown in Figure 43.6a, we can apply the particle in simple harmonic motion analysis model to the molecule as long as the atoms in the molecule are not too far from their equilibrium positions. Recall from Section 15.3 that the potential energy function for a simple harmonic oscillator is parabolic, varying as the square of the position of the particle relative to the equilibrium position. (See Eq. 15.20 and Fig. 15.9b.) Figure 43.6b shows a plot of potential energy versus atomic separation for a diatomic molecule, where r_0 is the equilibrium atomic separation. For separations close to r_0, the shape of the potential energy curve closely resembles the parabolic shape of the potential energy function in the particle in simple harmonic motion model.

According to classical mechanics, the frequency of vibration for the system shown in Figure 43.6a is given by Equation 15.14:

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \qquad \textbf{(43.8)}$$

where k is the effective spring constant and μ is the reduced mass given by Equation 43.4. In Section 21.3, we studied the contribution of a molecule's vibration to the specific heats of gases.

Quantum mechanics predicts that a molecule vibrates in quantized states as described in Section 41.7. The vibrational motion and quantized vibrational energy can be altered if the molecule acquires energy of the proper value to cause a transition between quantized vibrational states. As discussed in Section 41.7, the allowed vibrational energies are

$$E_{\text{vib}} = (v + \tfrac{1}{2})hf \qquad v = 0, 1, 2, \ldots \qquad \textbf{(43.9)}$$

where v is an integer called the **vibrational quantum number.** (We used n in Section 41.7 for a general harmonic oscillator, but v is often used for the quantum number when discussing molecular vibrations.) If the system is in the lowest vibrational state, for which $v = 0$, its ground-state energy is $\frac{1}{2}hf$. In the first excited vibrational state, $v = 1$ and the energy is $\frac{3}{2}hf$, and so on.

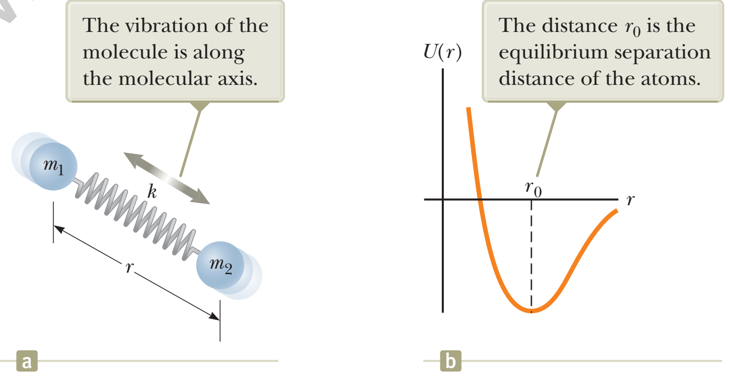


Figure 43.6 (a) Effective-spring model of a diatomic molecule. (b) Plot of the potential energy of a diatomic molecule versus atomic separation distance. Compare with Figure 15.11a.

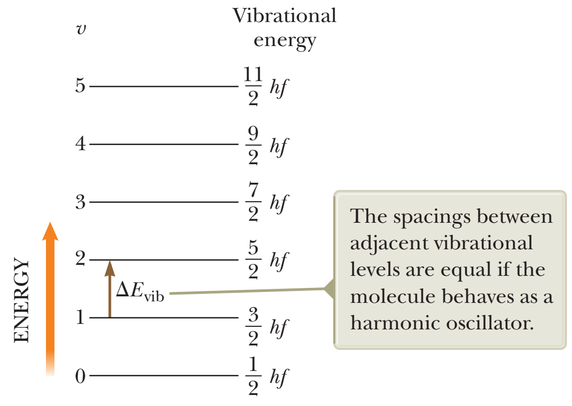


Figure 43.7 Allowed vibrational energies of a diatomic molecule, where f is the frequency of vibration of the molecule, given by Equation 43.8.

Substituting Equation 43.8 into Equation 43.9 gives the following expression for the allowed vibrational energies:

Allowed values of vibrational energy

$$E_{\text{vib}} = \left(v + \tfrac{1}{2}\right)\frac{h}{2\pi}\sqrt{\frac{k}{\mu}} \qquad v = 0, 1, 2, \ldots \tag{43.10}$$

The selection rule for the allowed vibrational transitions is $\Delta v = \pm 1$. Transitions between vibrational levels are caused by absorption of photons in the infrared region of the spectrum. The energy of an absorbed photon is equal to the energy difference between any two successive vibrational levels. Therefore, the photon energy is given by

$$E_{\text{photon}} = \Delta E_{\text{vib}} = \frac{h}{2\pi}\sqrt{\frac{k}{\mu}} \tag{43.11}$$

The vibrational energies of a diatomic molecule are plotted in Figure 43.7. At ordinary temperatures, most molecules have vibrational energies corresponding to the $v = 0$ state because the spacing between vibrational states is much greater than $k_B T$, where k_B is Boltzmann's constant and T is the temperature.

Quick Quiz 43.3 A gas of identical diatomic molecules absorbs electromagnetic radiation over a wide range of frequencies. Molecule 1, initially in the $v = 0$ vibrational state, makes a transition to the $v = 1$ state. Molecule 2, initially in the $v = 2$ state, makes a transition to the $v = 3$ state. What is the ratio of the frequency of the photon that excited molecule 2 to that of the photon that excited molecule 1? **(a)** 1 **(b)** 2 **(c)** 3 **(d)** 4 **(e)** impossible to determine

Example 43.2 Vibration of the CO Molecule AM

The frequency of the photon that causes the $v = 0$ to $v = 1$ transition in the CO molecule is 6.42×10^{13} Hz. We ignore any changes in the rotational energy for this example.

(A) Calculate the force constant k for this molecule.

SOLUTION

Conceptualize Imagine that the two atoms in Figure 43.6a are carbon and oxygen. As the molecule vibrates, a given point on the imaginary spring is at rest. This point is not midway between the atoms because of the difference in masses of the C and O atoms.

Categorize The statement of the problem tells us to categorize this example as one involving a quantum-mechanical treatment and to restrict our investigation to the vibrational motion of a diatomic molecule. The molecule is analyzed with portions of the *particle in simple harmonic motion* analysis model.

▶ 43.2 continued

Analyze Set Equation 43.11 equal to the photon energy hf and solve for the force constant:

$$\frac{h}{2\pi}\sqrt{\frac{k}{\mu}} = hf \;\rightarrow\; k = 4\pi^2\mu f^2$$

Substitute the frequency given in the problem statement and the reduced mass from Example 43.1:

$$k = 4\pi^2(1.14 \times 10^{-26}\text{ kg})(6.42 \times 10^{13}\text{ s}^{-1})^2 = \boxed{1.85 \times 10^3\text{ N/m}}$$

(B) What is the classical amplitude A of vibration for this molecule in the $v = 0$ vibrational state?

SOLUTION

Equate the maximum elastic potential energy $\frac{1}{2}kA^2$ in the molecule (Eq. 15.21) to the vibrational energy given by Equation 43.10 with $v = 0$ and solve for A:

$$\tfrac{1}{2}kA^2 = \frac{h}{4\pi}\sqrt{\frac{k}{\mu}} \;\rightarrow\; A = \sqrt{\frac{h}{2\pi}}\left(\frac{1}{\mu k}\right)^{1/4}$$

Substitute the value for k from part (A) and the value for μ:

$$A = \sqrt{\frac{6.626 \times 10^{-34}\text{ J}\cdot\text{s}}{2\pi}}\left[\frac{1}{(1.14 \times 10^{-26}\text{ kg})(1.85 \times 10^3\text{ N/m})}\right]^{1/4}$$

$$= 4.79 \times 10^{-12}\text{ m} = \boxed{0.004\,79\text{ nm}}$$

Finalize Comparing this result with the bond length of 0.113 nm we calculated in Example 43.1 shows that the classical amplitude of vibration is approximately 4% of the bond length.

Molecular Spectra

In general, a molecule vibrates and rotates simultaneously. To a first approximation, these motions are independent of each other, so the total energy of the molecule for these motions is the sum of Equations 43.6 and 43.9:

$$E = (v + \tfrac{1}{2})hf + \frac{\hbar^2}{2I}J(J+1) \qquad \textbf{(43.12)}$$

The energy levels of any molecule can be calculated from this expression, and each level is indexed by the two quantum numbers v and J. From these calculations, an energy-level diagram like the one shown in Figure 43.8a (page 1350) can be constructed. For each allowed value of the vibrational quantum number v, there is a complete set of rotational levels corresponding to $J = 0, 1, 2, \ldots$. The energy separation between successive rotational levels is much smaller than the separation between successive vibrational levels. As noted earlier, most molecules at ordinary temperatures are in the $v = 0$ vibrational state; these molecules can be in various rotational states as Figure 43.8a shows.

When a molecule absorbs a photon with the appropriate energy, the vibrational quantum number v increases by one unit while the rotational quantum number J either increases or decreases by one unit as can be seen in Figure 43.8. Therefore, the molecular absorption spectrum in Figure 43.8b consists of two groups of lines: one group to the right of center and satisfying the selection rules $\Delta J = +1$ and $\Delta v = +1$, and the other group to the left of center and satisfying the selection rules $\Delta J = -1$ and $\Delta v = +1$.

The energies of the absorbed photons can be calculated from Equation 43.12:

$$E_{\text{photon}} = \Delta E = hf + \frac{\hbar^2}{I}(J+1) \qquad J = 0, 1, 2, \ldots \qquad (\Delta J = +1) \qquad \textbf{(43.13)}$$

$$E_{\text{photon}} = \Delta E = hf - \frac{\hbar^2}{I}J \qquad J = 1, 2, 3, \ldots \qquad (\Delta J = -1) \qquad \textbf{(43.14)}$$

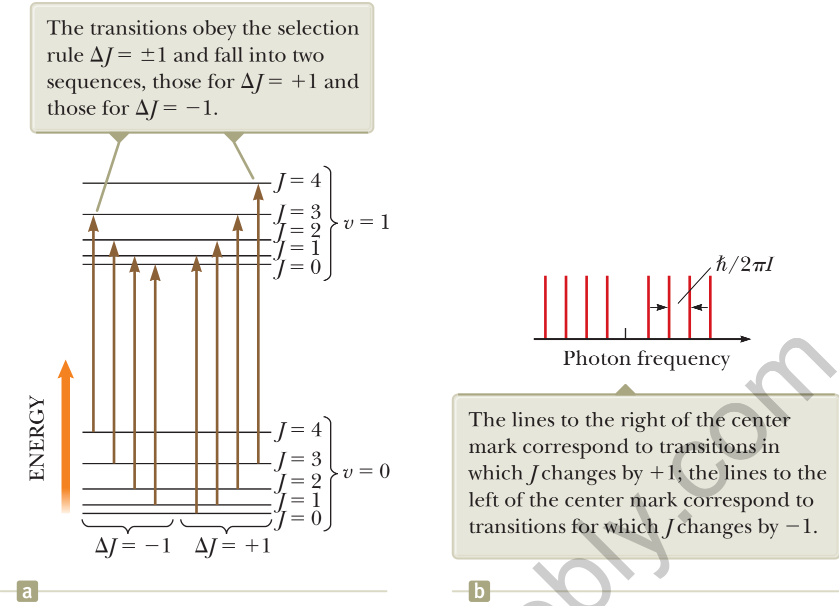


Figure 43.8 (a) Absorptive transitions between the $v = 0$ and $v = 1$ vibrational states of a diatomic molecule. Compare the energy levels in this figure with those in Figure 21.7. (b) Expected lines in the absorption spectrum of a molecule. These same lines appear in the emission spectrum.

where J is the rotational quantum number of the *initial* state. Equation 43.13 generates the series of equally spaced lines *higher* than the frequency f, whereas Equation 43.14 generates the series *lower* than this frequency. Adjacent lines are separated in frequency by the fundamental unit $\hbar/2\pi I$. Figure 43.8b shows the expected frequencies in the absorption spectrum of the molecule; these same frequencies appear in the emission spectrum.

The experimental absorption spectrum of the HCl molecule shown in Figure 43.9 follows this pattern very well and reinforces our model. One peculiarity is apparent, however: each line is split into a doublet. This doubling occurs because two chlorine isotopes (Cl-35 and Cl-37; see Section 44.1) were present in the sample used to obtain this spectrum. Because the isotopes have different masses, the two HCl molecules have different values of I.

The intensity of the spectral lines in Figure 43.9 follows an interesting pattern, rising first as one moves away from the central gap (located at about 8.65×10^{13} Hz, corresponding to the forbidden $J = 0$ to $J = 0$ transition) and then falling. This intensity is determined by a product of two functions of J. The first function corresponds to the number of available states for a given value of J. This function is $2J + 1$, corresponding to the number of values of m_J, the molecular rotation analog to m_ℓ for atomic states. For example, the $J = 2$ state has five substates with five values of m_J ($m_J = -2, -1, 0, 1, 2$), whereas the $J = 1$ state has only three substates ($m_J = -1, 0, 1$). Therefore, on average and without regard for the second function described below, five-thirds as many molecules make the transition from the $J = 2$ state as from the $J = 1$ state.

The second function determining the envelope of the intensity of the spectral lines is the Boltzmann factor, introduced in Section 21.5. The number of molecules in an excited rotational state is given by

$$n = n_0 e^{-\hbar^2 J(J+1)/(2Ik_B T)}$$

where n_0 is the number of molecules in the $J = 0$ state.

Multiplying these factors together indicates that the intensity of spectral lines should be described by a function of J as follows:

Intensity variation in the vibration–rotation spectrum of a molecule

$$I \propto (2J + 1)e^{-\hbar^2 J(J+1)/(2Ik_B T)} \qquad \textbf{(43.15)}$$

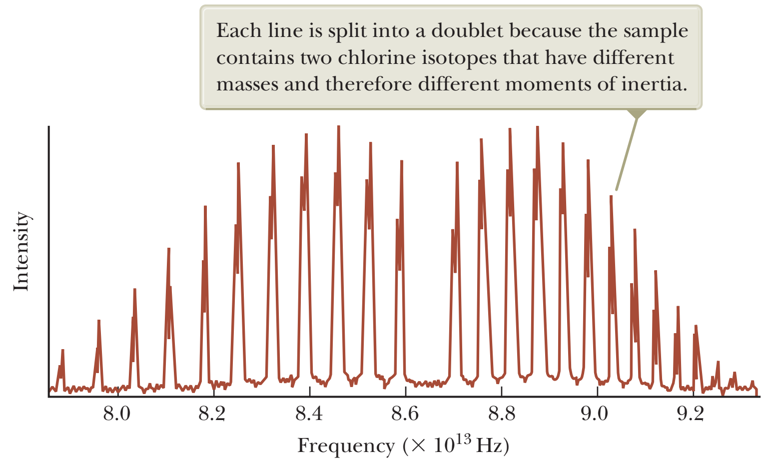


Figure 43.9 Experimental absorption spectrum of the HCl molecule.

The factor $(2J + 1)$ increases with J while the exponential second factor decreases. The product of the two factors gives a behavior that closely describes the envelope of the spectral lines in Figure 43.9.

The excitation of rotational and vibrational energy levels is an important consideration in current models of global warming. Most of the absorption lines for CO_2 are in the infrared portion of the spectrum. Therefore, visible light from the Sun is not absorbed by atmospheric CO_2 but instead strikes the Earth's surface, warming it. In turn, the surface of the Earth, being at a much lower temperature than the Sun, emits thermal radiation that peaks in the infrared portion of the electromagnetic spectrum (Section 40.1). This infrared radiation is absorbed by the CO_2 molecules in the air instead of radiating out into space. Atmospheric CO_2 acts like a one-way valve for energy from the Sun and is responsible, along with some other atmospheric molecules, for raising the temperature of the Earth's surface above its value in the absence of an atmosphere. This phenomenon is commonly called the "greenhouse effect." The burning of fossil fuels in today's industrialized society adds more CO_2 to the atmosphere. This addition of CO_2 increases the absorption of infrared radiation, raising the Earth's temperature further. In turn, this increase in temperature causes substantial climatic changes.

As seen in Figure 43.10, the amount of carbon dioxide in the atmosphere has been steadily increasing since the middle of the 20th century. This graph shows hard data that indicate that the atmosphere is undergoing a distinct change, although not all scientists agree on the interpretation of what that change means in terms of global temperatures.

The Intergovernmental Panel on Climate Change (IPCC) is a scientific body that assesses the available information related to global warming and associated effects

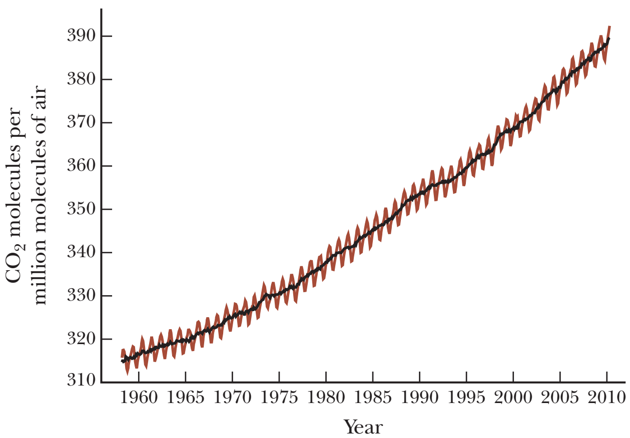


Figure 43.10 The concentration of atmospheric carbon dioxide in parts per million (ppm) of dry air as a function of time. These data were recorded at the Mauna Loa Observatory in Hawaii. The yearly variations (red-brown curve) coincide with growing seasons because vegetation absorbs carbon dioxide from the air. The steady increase in the average concentration (black curve) is of concern to scientists.

related to climate change. It was originally established in 1988 by two United Nations organizations, the World Meteorological Organization and the United Nations Environment Programme. The IPCC has published four assessment reports on climate change, the most recent in 2007, and a fifth report is scheduled to be released in 2014. The 2007 report concludes that there is a probability of greater than 90% that the increased global temperature measured by scientists is due to the placement of greenhouse gases such as carbon dioxide in the atmosphere by humans. The report also predicts a global temperature increase between 1°C and 6°C in the 21st century, a sea level rise from 18 cm to 59 cm, and very high probabilities of weather extremes, including heat waves, droughts, cyclones, and heavy rainfall.

In addition to its scientific aspects, global warming is a social issue with many facets. These facets encompass international politics and economics, because global warming is a worldwide problem. Changing our policies requires real costs to solve the problem. Global warming also has technological aspects, and new methods of manufacturing, transportation, and energy supply must be designed to slow down or reverse the increase in temperature.

Conceptual Example 43.3 **Comparing Figures 43.8 and 43.9**

In Figure 43.8a, the transitions indicated correspond to spectral lines that are equally spaced as shown in Figure 43.8b. The actual spectrum in Figure 43.9, however, shows lines that move closer together as the frequency increases. Why does the spacing of the actual spectral lines differ from the diagram in Figure 43.8?

SOLUTION

In Figure 43.8, we modeled the rotating diatomic molecule as a rigid object (Chapter 10). In reality, however, as the molecule rotates faster and faster, the effective spring in Figure 43.6a stretches and provides the increased force associated with the larger centripetal acceleration of each atom. As the molecule stretches along its length, its moment of inertia I increases. Therefore, the rotational part of the energy expression in Equation 43.12 has an extra dependence on J in the moment of inertia I. Because the increasing moment of inertia is in the denominator, as J increases, the energies do not increase as rapidly with J as indicated in Equation 43.12. With each higher energy level being lower than indicated by Equation 43.12, the energy associated with a transition to that level is smaller, as is the frequency of the absorbed photon, destroying the even spacing of the spectral lines and giving the spacing that decreases with increasing frequency seen in Figure 43.9.

43.3 Bonding in Solids

A crystalline solid consists of a large number of atoms arranged in a regular array, forming a periodic structure. The ions in the NaCl crystal are ionically bonded, as already noted, and the carbon atoms in diamond form covalent bonds with one another. The metallic bond described at the end of this section is responsible for the cohesion of copper, silver, sodium, and other solid metals.

Ionic Solids

Many crystals are formed by ionic bonding, in which the dominant interaction between ions is the Coulomb force. Consider a portion of the NaCl crystal shown in Figure 43.11a. The red spheres are sodium ions, and the blue spheres are chlorine ions. As shown in Figure 43.11b, each Na^+ ion has six nearest-neighbor Cl^- ions. Similarly, in Figure 43.11c, we see that each Cl^- ion has six nearest-neighbor Na^+ ions. Each Na^+ ion is attracted to its six Cl^- neighbors. The corresponding potential energy is $-6k_e e^2/r$, where k_e is the Coulomb constant and r is the separation distance between each Na^+ and Cl^-. In addition, there are 12 next-nearest-neighbor

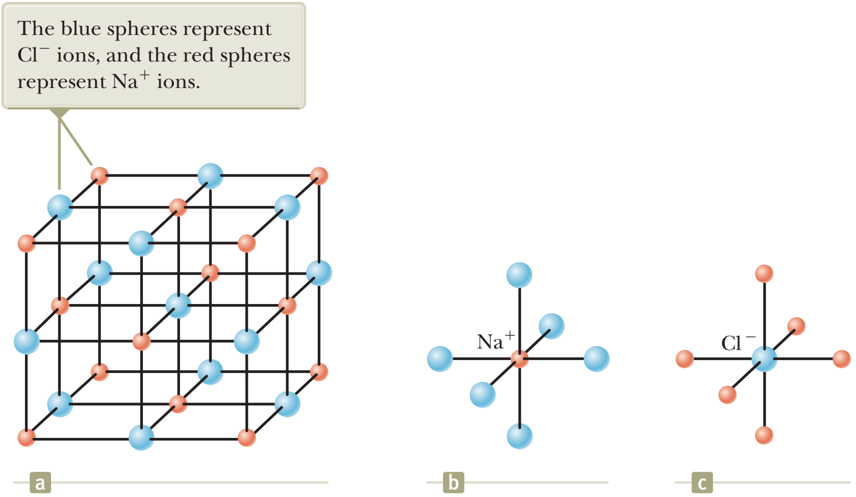


Figure 43.11 (a) Crystalline structure of NaCl. (b) Each positive sodium ion is surrounded by six negative chlorine ions. (c) Each chlorine ion is surrounded by six sodium ions.

Na^+ ions at a distance of $\sqrt{2}r$ from the Na^+ ion, and these 12 positive ions exert weaker repulsive forces on the central Na^+. Furthermore, beyond these 12 Na^+ ions are more Cl^- ions that exert an attractive force, and so on. The net effect of all these interactions is a resultant negative electric potential energy

$$U_{\text{attractive}} = -\alpha k_e \frac{e^2}{r} \tag{43.16}$$

where α is a dimensionless number known as the **Madelung constant.** The value of α depends only on the particular crystalline structure of the solid. For example, $\alpha = 1.747\,6$ for the NaCl structure. When the constituent ions of a crystal are brought close together, a repulsive force exists because of electrostatic forces and the exclusion principle as discussed in Section 43.1. The potential energy term B/r^m in Equation 43.1 accounts for this repulsive force. We do not include neighbors other than nearest neighbors here because the repulsive forces occur only for ions that are very close together. (Electron shells must overlap for exclusion-principle effects to become important.) Therefore, we can express the total potential energy of the crystal as

$$U_{\text{total}} = -\alpha k_e \frac{e^2}{r} + \frac{B}{r^m} \tag{43.17}$$

where m in this expression is some small integer.

A plot of total potential energy versus ion separation distance is shown in Figure 43.12. The potential energy has its minimum value U_0 at the equilibrium separation, when $r = r_0$. It is left as a problem (Problem 59) to show that

$$U_0 = -\alpha k_e \frac{e^2}{r_0}\left(1 - \frac{1}{m}\right) \tag{43.18}$$

This minimum energy U_0 is called the **ionic cohesive energy** of the solid, and its absolute value represents the energy required to separate the solid into a collection of isolated positive and negative ions. Its value for NaCl is -7.84 eV per ion pair.

To calculate the **atomic cohesive energy,** which is the binding energy relative to the energy of the neutral atoms, 5.14 eV must be added to the ionic cohesive energy value to account for the transition from Na^+ to Na and 3.62 eV must be subtracted to account for the conversion of Cl^- to Cl. Therefore, the atomic cohesive energy of NaCl is

$$-7.84 \text{ eV} + 5.14 \text{ eV} - 3.62 \text{ eV} = -6.32 \text{ eV}$$

In other words, 6.32 eV of energy per ion pair is needed to separate the solid into isolated neutral atoms of Na and Cl.

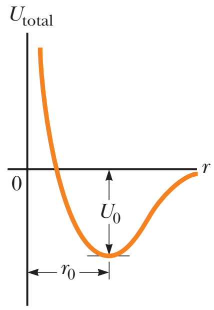


Figure 43.12 Total potential energy versus ion separation distance for an ionic solid, where U_0 is the ionic cohesive energy and r_0 is the equilibrium separation distance between ions.

Ionic crystals form relatively stable, hard crystals. They are poor electrical conductors because they contain no free electrons; each electron in the solid is bound tightly to one of the ions, so it is not sufficiently mobile to carry current. Ionic crystals have high melting points; for example, the melting point of NaCl is 801°C. Ionic crystals are transparent to visible radiation because the shells formed by the electrons in ionic solids are so tightly bound that visible radiation does not possess sufficient energy to promote electrons to the next allowed shell. Infrared radiation is absorbed strongly because the vibrations of the ions have natural resonant frequencies in the low-energy infrared region.

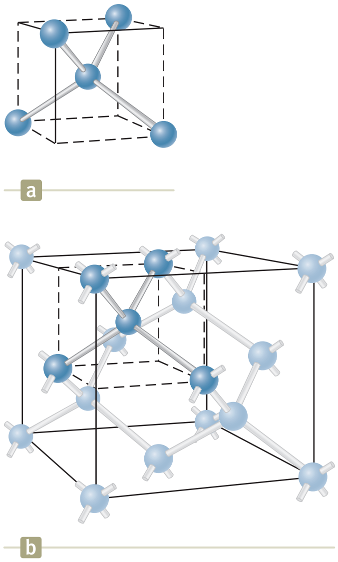


Figure 43.13 (a) Each carbon atom in a diamond crystal is covalently bonded to four other carbon atoms so that a tetrahedral structure is formed. (b) The crystal structure of diamond, showing the tetrahedral bond arrangement.

Covalent Solids

Solid carbon, in the form of diamond, is a crystal whose atoms are covalently bonded. Because atomic carbon has the electronic configuration $1s^2 2s^2 2p^2$, it is four electrons short of filling its $n = 2$ shell, which can accommodate eight electrons. Because of this electron structure, two carbon atoms have a strong attraction for each other, with a cohesive energy of 7.37 eV. In the diamond structure, each carbon atom is covalently bonded to four other carbon atoms located at four corners of a cube as shown in Figure 43.13a.

The crystalline structure of diamond is shown in Figure 43.13b. Notice that each carbon atom forms covalent bonds with four nearest-neighbor atoms. The basic structure of diamond is called tetrahedral (each carbon atom is at the center of a regular tetrahedron), and the angle between the bonds is 109.5°. Other crystals such as silicon and germanium have the same structure.

Carbon is interesting in that it can form several different types of structures. In addition to the diamond structure, it forms graphite, with completely different properties. In this form, the carbon atoms form flat layers with hexagonal arrays of atoms. A very weak interaction between the layers allows the layers to be removed easily under friction, as occurs in the graphite used in pencil lead.

Carbon atoms can also form a large hollow structure; in this case, the compound is called **buckminsterfullerene** after the famous architect R. Buckminster Fuller, who invented the geodesic dome. The unique shape of this molecule (Fig. 43.14) provides a "cage" to hold other atoms or molecules. Related structures, called "buckytubes" because of their long, narrow cylindrical arrangements of carbon atoms, may provide the basis for extremely strong, yet lightweight, materials.

A current area of active research is in the properties and applications of **graphene.** Graphene consists of a monolayer of carbon atoms, with the atoms arranged in hexagons so that the monolayer looks like chicken wire. Graphite flakes that are shed from a pencil while writing contain small fragments of graphene. Pioneers in graphene research include Andre Geim (b. 1958) and Konstantin Novoselov (b. 1974) of the University of Manchester, who received the Nobel Prize in Physics in 2010 for their experiments. Graphene has interesting electronic, thermal, and optical properties that are currently under investigation. Its mechanical properties include a breaking strength 200 times that of steel. Potential applications under

A cylinder of nearly pure crystalline silicon (Si), approximately 25 cm long. Such crystals are cut into wafers and processed to make various semiconductor devices.

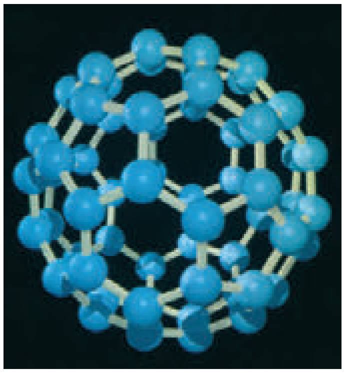


Figure 43.14 Computer rendering of a "buckyball," short for the molecule buckminsterfullerene. These nearly spherical molecular structures that look like soccer balls were named for the inventor of the geodesic dome. This form of carbon, C_{60}, was discovered by astrophysicists investigating the carbon gas that exists between stars. Scientists are actively studying the properties and potential uses of buckminsterfullerene and related molecules.

Table 43.1 Atomic Cohesive Energies of Some Covalent Solids

Solid	Cohesive Energy (eV per ion pair)
C (diamond)	7.37
Si	4.63
Ge	3.85
InAs	5.70
SiC	6.15
ZnS	6.32
CuCl	9.24

study include graphene nanoribbons, quantum dots, transistors, optical modulators, and integrated circuits.

The atomic cohesive energies of some covalent solids are given in Table 43.1. The large energies account for the hardness of covalent solids. Diamond is particularly hard and has an extremely high melting point (about 4 000 K). Covalently bonded solids usually have high bond energies and high melting points, and are good electrical insulators.

Metallic Solids

Metallic bonds are generally weaker than ionic or covalent bonds. The outer electrons in the atoms of a metal are relatively free to move throughout the material, and the number of such mobile electrons in a metal is large. The metallic structure can be viewed as a "sea" or a "gas" of nearly free electrons surrounding a lattice of positive ions (Fig. 43.15). The bonding mechanism in a metal is the attractive force between the entire collection of positive ions and the electron gas. Metals have a cohesive energy in the range of 1 to 3 eV per atom, which is less than the cohesive energies of ionic or covalent solids.

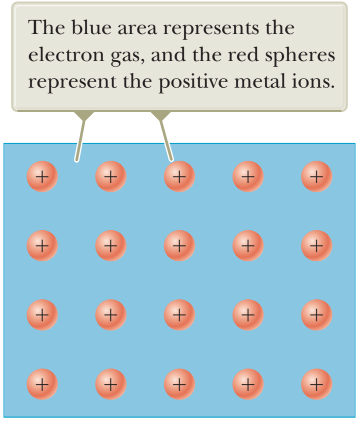


Figure 43.15 Highly schematic diagram of a metal.

Light interacts strongly with the free electrons in metals. Hence, visible light is absorbed and re-emitted quite close to the surface of a metal, which accounts for the shiny nature of metal surfaces. In addition to the high electrical conductivity of metals produced by the free electrons, the nondirectional nature of the metallic bond allows many different types of metal atoms to be dissolved in a host metal in varying amounts. The resulting *solid solutions,* or *alloys* (steel, bronze, brass, etc.), may be designed to have particular properties, such as tensile strength, ductility, electrical and thermal conductivity, and resistance to corrosion.

Because the bonding in metals is between all the electrons and all the positive ions, metals tend to bend when stressed. This bending is in contrast to nonmetallic solids, which tend to fracture when stressed. Fracturing results because bonding in nonmetallic solids is primarily with nearest-neighbor ions or atoms. When the distortion causes sufficient stress between some set of nearest neighbors, fracture occurs.

43.4 Free-Electron Theory of Metals

In Section 27.3, we described a classical free-electron theory of electrical conduction in metals, a structural model that led to Ohm's law. According to this theory, a metal is modeled as a classical gas of conduction electrons moving through a fixed lattice of ions. Although this theory predicts the correct functional form of Ohm's law, it does not predict the correct values of electrical and thermal conductivities.

A quantum-based free-electron theory of metals remedies the shortcomings of the classical model by taking into account the wave nature of the electrons. In this model, based on the quantum particle under boundary conditions analysis model, the outer-shell electrons are free to move through the metal but are trapped within a three-dimensional box formed by the metal surfaces. Therefore, each electron is represented as a particle in a box. As discussed in Section 41.2, particles in a box are restricted to quantized energy levels.

Statistical physics can be applied to a collection of particles in an effort to relate microscopic properties to macroscopic properties as we saw with kinetic theory of gases in Chapter 21. In the case of electrons, it is necessary to use *quantum statistics,* with the requirement that each state of the system can be occupied by only two electrons (one with spin up and the other with spin down) as a consequence of the exclusion principle. The probability that a particular state having energy E is occupied by one of the electrons in a solid is

Fermi–Dirac distribution function

$$f(E) = \frac{1}{e^{(E-E_F)/k_B T} + 1} \tag{43.19}$$

where $f(E)$ is called the **Fermi–Dirac distribution function** and E_F is called the **Fermi energy.** A plot of $f(E)$ versus E at $T = 0$ K is shown in Figure 43.16a. Notice that $f(E) = 1$ for $E < E_F$ and $f(E) = 0$ for $E > E_F$. That is, at 0 K, all states having energies less than the Fermi energy are occupied and all states having energies greater than the Fermi energy are vacant. A plot of $f(E)$ versus E at some temperature $T > 0$ K is shown in Figure 43.16b. This curve shows that as T increases, the distribution rounds off slightly. Because of thermal excitation, states near and below E_F lose population and states near and above E_F gain population. The Fermi energy E_F also depends on temperature, but the dependence is weak in metals.

Let's now follow up on our discussion of the particle in a box in Chapter 41 to generalize the results to a three-dimensional box. Recall that if a particle of mass m is confined to move in a one-dimensional box of length L, the allowed states have quantized energy levels given by Equation 41.14:

$$E_n = \left(\frac{h^2}{8mL^2}\right)n^2 = \left(\frac{\hbar^2\pi^2}{2mL^2}\right)n^2 \quad n = 1, 2, 3, \ldots$$

Now imagine a piece of metal in the shape of a solid cube of sides L and volume L^3 and focus on one electron that is free to move anywhere in this volume. Therefore, the electron is modeled as a particle in a three-dimensional box. In this model, we require that $\psi(x, y, z) = 0$ at the boundaries of the metal. It can be shown (see Problem 37) that the energy for such an electron is

$$E = \frac{\hbar^2\pi^2}{2m_e L^2}(n_x^2 + n_y^2 + n_z^2) \tag{43.20}$$

where m_e is the mass of the electron and n_x, n_y, and n_z are quantum numbers. As we expect, the energies are quantized, and each allowed value of the energy is characterized by this set of three quantum numbers (one for each degree of freedom) and the spin quantum number m_s. For example, the ground state, corresponding to

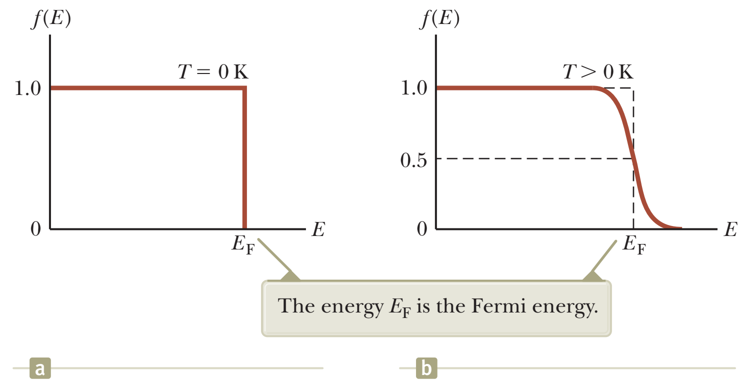


Figure 43.16 Plot of the Fermi–Dirac distribution function $f(E)$ versus energy at (a) $T = 0$ K and (b) $T > 0$ K.

$n_x = n_y = n_z = 1$, has an energy equal to $3\hbar^2\pi^2/2m_eL^2$ and can be occupied by two electrons, corresponding to spin up and spin down.

Because of the macroscopic size L of the box, the energy levels for the electrons are very close together. As a result, we can treat the quantum numbers as continuous variables. Under this assumption, the number of allowed states per unit volume that have energies between E and $E + dE$ is

$$g(E)\,dE = \frac{8\sqrt{2}\,\pi m_e^{3/2}}{h^3}E^{1/2}\,dE \qquad \textbf{(43.21)}$$

(See Example 43.5.) The function $g(E)$ is called the **density-of-states function.**

If a metal is in thermal equilibrium, the number of electrons per unit volume $N(E)\,dE$ that have energy between E and $E + dE$ is equal to the product of the number of allowed states per unit volume and the probability that a state is occupied; that is, $N(E)\,dE = g(E)f(E)\,dE$:

$$N(E)\,dE = \left(\frac{8\sqrt{2}\,\pi m_e^{3/2}}{h^3}E^{1/2}\right)\left(\frac{1}{e^{(E-E_F)/k_BT}+1}\right)dE \qquad \textbf{(43.22)}$$

Plots of $N(E)$ versus E for two temperatures are given in Figure 43.17.

If n_e is the total number of electrons per unit volume, we require that

$$n_e = \int_0^\infty N(E)\,dE = \frac{8\sqrt{2}\,\pi m_e^{3/2}}{h^3}\int_0^\infty \frac{E^{1/2}\,dE}{e^{(E-E_F)/k_BT}+1} \qquad \textbf{(43.23)}$$

We can use this condition to calculate the Fermi energy. At $T = 0$ K, the Fermi–Dirac distribution function $f(E) = 1$ for $E < E_F$ and $f(E) = 0$ for $E > E_F$. Therefore, at $T = 0$ K, Equation 43.23 becomes

$$n_e = \frac{8\sqrt{2}\,\pi m_e^{3/2}}{h^3}\int_0^{E_F} E^{1/2}\,dE = \tfrac{2}{3}\frac{8\sqrt{2}\,\pi m_e^{3/2}}{h^3}E_F^{3/2} \qquad \textbf{(43.24)}$$

Solving for the Fermi energy at 0 K gives

$$E_F(0) = \frac{h^2}{2m_e}\left(\frac{3n_e}{8\pi}\right)^{2/3} \qquad \textbf{(43.25)}$$

◀ **Fermi energy at $T = 0$ K**

The Fermi energies for metals are in the range of a few electron volts. Representative values for various metals are given in Table 43.2. It is left as a problem (Problem 39) to show that the average energy of a free electron in a metal at 0 K is

$$E_{avg} = \tfrac{3}{5}E_F \qquad \textbf{(43.26)}$$

In summary, we can consider a metal to be a system comprising a very large number of energy levels available to the free electrons. These electrons fill the levels in accordance with the Pauli exclusion principle, beginning with $E = 0$ and ending with E_F. At $T = 0$ K, all levels below the Fermi energy are filled and all levels above the Fermi energy are empty. At 300 K, a small fraction of the free electrons are excited above the Fermi energy.

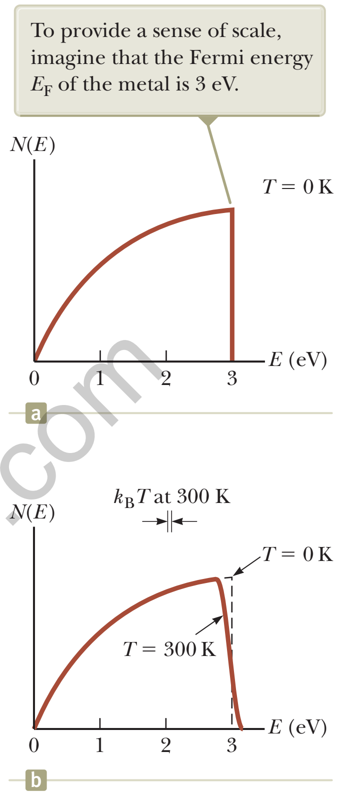


Figure 43.17 Plot of the electron distribution function versus energy in a metal at (a) $T = 0$ K and (b) $T = 300$ K.

Table 43.2 Calculated Values of the Fermi Energy for Metals at 300 K Based on the Free-Electron Theory

Metal	Electron Concentration (m^{-3})	Fermi Energy (eV)
Li	4.70×10^{28}	4.72
Na	2.65×10^{28}	3.23
K	1.40×10^{28}	2.12
Cu	8.46×10^{28}	7.05
Ag	5.85×10^{28}	5.48
Au	5.90×10^{28}	5.53

Example 43.4 The Fermi Energy of Gold

Each atom of gold (Au) contributes one free electron to the metal. Compute the Fermi energy for gold.

SOLUTION

Conceptualize Imagine electrons filling available levels at $T = 0$ K in gold until the solid is neutral. The highest energy filled is the Fermi energy.

Categorize We evaluate the result using a result from this section, so we categorize this example as a substitution problem.

Substitute the concentration of free electrons in gold from Table 43.2 into Equation 43.25 to calculate the Fermi energy at 0 K:

$$E_F(0) = \frac{(6.626 \times 10^{-34}\ \text{J}\cdot\text{s})^2}{2(9.11 \times 10^{-31}\ \text{kg})}\left[\frac{3(5.90 \times 10^{28}\ \text{m}^{-3})}{8\pi}\right]^{2/3}$$

$$= 8.85 \times 10^{-19}\ \text{J} = \boxed{5.53\ \text{eV}}$$

Example 43.5 Deriving Equation 43.21

Based on the allowed states of a particle in a three-dimensional box, derive Equation 43.21.

SOLUTION

Conceptualize Imagine a particle confined to a three-dimensional box, subject to boundary conditions in three dimensions. Imagine also a three-dimensional *quantum number space* whose axes represent n_x, n_y, and n_z. The allowed states in this space can be represented as dots located at integral values of the three quantum numbers as in Figure 43.18. This space is not traditional space in which a location is specified by coordinates x, y, and z; rather, it is a space in which allowed states can be specified by coordinates representing the quantum numbers. The number of allowed states having energies between E and $E + dE$ corresponds to the number of dots in the spherical shell of radius n and thickness dn.

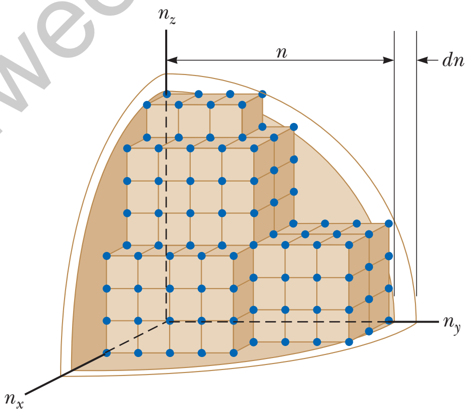


Figure 43.18 The dots representing the allowed states are located at integer values of n_x, n_y, and n_z and are therefore at the corners of cubes with sides of "length" 1.

Categorize We categorize this problem as that of a quantum system in which the energies of the particle are quantized. Furthermore, we can base the solution to the problem on our understanding of the particle in a one-dimensional box.

Analyze As noted previously, the allowed states of the particle in a three-dimensional box are described by three quantum numbers n_x, n_y, and n_z. For a macroscopic sample of metal, the number of allowed values of these quantum numbers is tremendous, so on a macroscopic scale, the allowed states in the number space can be modeled as continuous.

Defining $E_0 = \hbar^2\pi^2/2m_eL^2$ and $n = (E/E_0)^{1/2}$, rewrite Equation 43.20:

$$(1)\quad n_x^2 + n_y^2 + n_z^2 = \frac{2m_eL^2}{\hbar^2\pi^2}E = \frac{E}{E_0} = n^2$$

In the quantum number space, Equation (1) is the equation of a sphere of radius n. Therefore, the number of allowed states having energies between E and $E + dE$ is equal to the number of points in a spherical shell of radius n and thickness dn.

Find the "volume" of this shell, which represents the total number of states $G(E)\,dE$:

$$(2)\quad G(E)\,dE = \tfrac{1}{8}(4\pi n^2\,dn) = \tfrac{1}{2}\pi n^2\,dn$$

We have taken one-eighth of the total volume because we are restricted to the octant of a three-dimensional space in which all three quantum numbers are positive.

Replace n in Equation (2) with its equivalent in terms of E using the relation $n^2 = E/E_0$ from Equation (1):

$$G(E)\,dE = \tfrac{1}{2}\pi\left(\frac{E}{E_0}\right)d\left[\left(\frac{E}{E_0}\right)^{1/2}\right] = \tfrac{1}{2}\pi\frac{E}{(E_0)^{3/2}}\,d[(E)^{1/2}]$$

43.5 continued

Evaluate the differential:

$$G(E)\,dE = \tfrac{1}{2}\pi\left[\frac{E}{(E_0)^{3/2}}\right](\tfrac{1}{2}E^{-1/2}\,dE) = \tfrac{1}{4}\pi E_0^{-3/2}E^{1/2}\,dE$$

Substitute for E_0 from its definition above:

$$G(E)\,dE = \tfrac{1}{4}\pi\left(\frac{\hbar^2\pi^2}{2m_eL^2}\right)^{-3/2}E^{1/2}\,dE$$

$$= \frac{\sqrt{2}}{2}\,\frac{m_e^{3/2}L^3}{\hbar^3\pi^2}E^{1/2}\,dE$$

Letting $g(E)$ represent the number of states per unit volume, where L^3 is the volume V of the cubical box in normal space, find $g(E) = G(E)/V$:

$$g(E)\,dE = \frac{G(E)}{V}\,dE = \frac{\sqrt{2}}{2}\,\frac{m_e^{3/2}}{\hbar^3\pi^2}E^{1/2}\,dE$$

Substitute $\hbar = h/2\pi$:

$$g(E)\,dE = \frac{4\sqrt{2}\,\pi m_e^{3/2}}{h^3}E^{1/2}\,dE$$

Multiply by 2 for the two possible spin states in each particle-in-a-box state:

$$g(E)\,dE = \frac{8\sqrt{2}\,\pi m_e^{3/2}}{h^3}E^{1/2}\,dE$$

Finalize This result is Equation 43.21, which is what we set out to derive.

43.5 Band Theory of Solids

In Section 43.4, the electrons in a metal were modeled as particles free to move around inside a three-dimensional box and we ignored the influence of the parent atoms. In this section, we make the model more sophisticated by incorporating the contribution of the parent atoms that form the crystal.

Recall from Section 41.1 that the probability density $|\psi|^2$ for a system is physically significant, but the probability amplitude ψ is not. Let's consider as an example an atom that has a single s electron outside of a closed shell. Both of the following wave functions are valid for such an atom with atomic number Z:

$$\psi_s^{\,+}(r) = +Af(r)e^{-Zr/na_0} \qquad \psi_s^{\,-}(r) = -Af(r)e^{-Zr/na_0}$$

where A is the normalization constant and $f(r)$ is a function[3] of r that varies with the value of n. Choosing either of these wave functions leads to the same value of $|\psi|^2$, so both choices are equivalent. A difference arises, however, when two atoms are combined.

If two identical atoms are very far apart, they do not interact and their electronic energy levels can be considered to be those of isolated atoms. Suppose the two atoms are sodium, each having a lone 3s electron that is in a well-defined quantum state. As the two sodium atoms are brought closer together, their wave functions begin to overlap as we discussed for covalent bonding in Section 43.1. The properties of the combined system differ depending on whether the two atoms are combined with wave functions $\psi_s^{\,+}(r)$ as in Figure 43.19a or whether they are combined with one having wave function $\psi_s^{\,+}(r)$ and the other $\psi_s^{\,-}(r)$ as in Figure 43.19b. The choice of two atoms with wave function $\psi_s^{\,-}(r)$ is physically equivalent to that with two positive wave functions, so we do not consider it separately. When two wave functions $\psi_s^{\,+}(r)$ are combined, the result is a composite wave function in which the probability amplitudes add between the atoms. If $\psi_s^{\,+}(r)$ combines with $\psi_s^{\,-}(r)$,

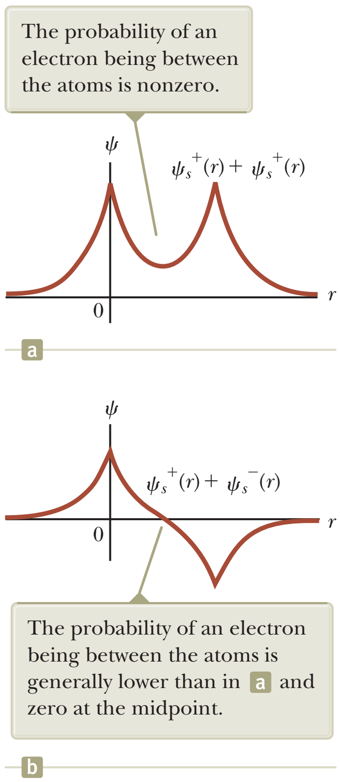


Figure 43.19 The wave functions of two atoms combine to form a composite wave function for the two-atom system when the atoms are close together. (a) Two atoms with wave functions $\psi_s^{\,+}(r)$ combine. (b) Two atoms with wave functions $\psi_s^{\,+}(r)$ and $\psi_s^{\,-}(r)$ combine.

[3]The functions $f(r)$ are called *Laguerre polynomials*. They can be found in the quantum treatment of the hydrogen atom in modern physics textbooks.

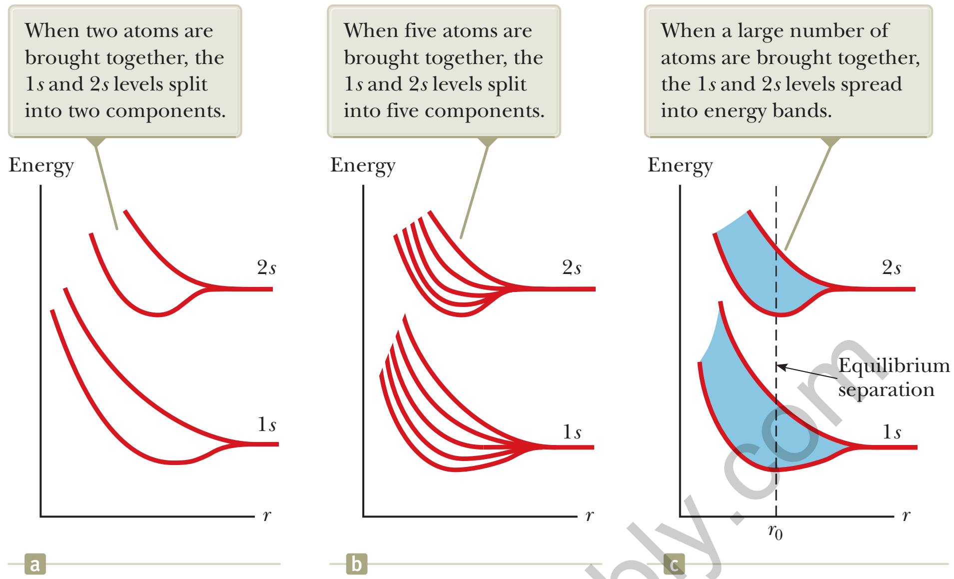


Figure 43.20 Energies of the 1s and 2s levels in sodium as a function of the separation distance r between atoms.

however, the wave functions between the nuclei subtract. Therefore, the composite probability amplitudes for the two possibilities are different. These two possible combinations of wave functions represent two possible states of the two-atom system. We interpret these curves as representing the probability amplitude of finding an electron. The positive–positive curve shows some probability of finding the electron at the midpoint between the atoms. The positive–negative function shows no such probability. A state with a high probability of an electron *between* two positive nuclei must have a different energy than a state with a high probability of the electron being elsewhere! Therefore, the states are *split* into two energy levels due to the two ways of combining the wave functions. The energy difference is relatively small, so the two states are close together on an energy scale.

Figure 43.20a shows this splitting effect as a function of separation distance. For large separations r, the electron clouds do not overlap and there is no splitting. As the atoms are brought closer so that r decreases, the electron clouds overlap and we need to consider the system of two atoms.

When a large number of atoms are brought together to form a solid, a similar phenomenon occurs. The individual wave functions can be brought together in various combinations of $\psi_s^+(r)$ and $\psi_s^-(r)$, each possible combination corresponding to a different energy. As the atoms are brought close together, the various isolated-atom energy levels split into multiple energy levels for the composite system. This splitting in levels for five atoms in close proximity is shown in Figure 43.20b. In this case, there are five energy levels corresponding to five different combinations of isolated-atom wave functions.

As the number of atoms grows, the number of combinations of wave functions grows, as does the number of possible energies. If we extend this argument to the large number of atoms found in solids (on the order of 10^{23} atoms per cubic centimeter), we obtain a huge number of levels of varying energy so closely spaced that they may be regarded as a continuous **band** of energy levels as shown in Figure 43.20c. In the case of sodium, it is customary to refer to the continuous distributions of allowed energy levels as s bands because the bands originate from the s levels of the individual sodium atoms.

Each energy level in the atom can spread into a band when the atoms are combined into a solid. Figure 43.21 shows the allowed energy bands of sodium at a fixed separation distance between the atoms. Notice that energy gaps, corresponding to *forbidden energies,* occur between the allowed bands. In addition, some bands exhibit sufficient spreading in energy that there is an overlap between bands arising from different quantum states (3s and 3p).

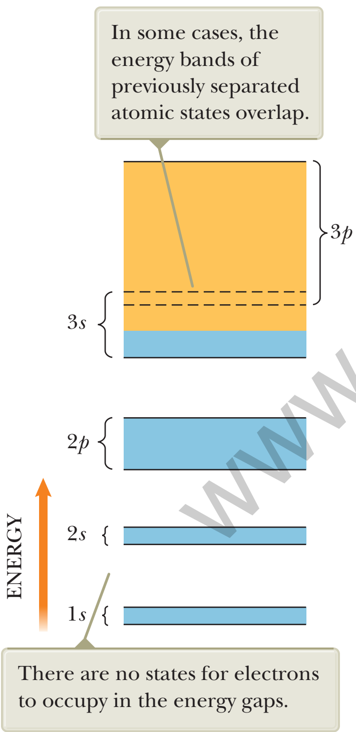


Figure 43.21 Energy bands of a sodium crystal. Blue represents energy bands occupied by the sodium electrons when the atom is in its ground state. Gold represents energy bands that are empty.

As indicated by the blue-shaded areas in Figure 43.21, the 1*s*, 2*s*, and 2*p* bands of sodium are each full of electrons because the 1*s*, 2*s*, and 2*p* states of each atom are full. An energy level in which the orbital angular momentum is ℓ can hold $2(2\ell + 1)$ electrons. The factor 2 arises from the two possible electron spin orientations, and the factor $2\ell + 1$ corresponds to the number of possible orientations of the orbital angular momentum. The capacity of each band for a system of N atoms is $2(2\ell + 1)N$ electrons. Therefore, the 1*s* and 2*s* bands each contain $2N$ electrons ($\ell = 0$), and the 2*p* band contains $6N$ electrons ($\ell = 1$). Because sodium has only one 3*s* electron and there are a total of N atoms in the solid, the 3*s* band contains only N electrons and is partially full as indicated by the blue coloring in Figure 43.21. The 3*p* band, which is the higher region of the overlapping bands, is completely empty (all gold in the figure).

Band theory allows us to build simple models to understand the behavior of conductors, insulators, and semiconductors as well as that of semiconductor devices, as we shall discuss in the following sections.

43.6 Electrical Conduction in Metals, Insulators, and Semiconductors

Good electrical conductors contain a high density of free charge carriers, and the density of free charge carriers in insulators is nearly zero. Semiconductors, first introduced in Section 23.2, are a class of technologically important materials in which charge-carrier densities are intermediate between those of insulators and those of conductors. In this section, we discuss the mechanisms of conduction in these three classes of materials in terms of a model based on energy bands.

Metals

If a material is to be a good electrical conductor, the charge carriers in the material must be free to move in response to an applied electric field. Let's consider the electrons in a metal as the charge carriers. The motion of the electrons in response to an electric field represents an increase in energy of the system (the metal lattice and the free electrons) corresponding to the additional kinetic energy of the moving electrons. The system is described by the nonisolated system model for energy. Equation 8.2 becomes $W = \Delta K$, where the work is done on the electrons by the electric field. Therefore, when an electric field is applied to a conductor, electrons must move upward to an available higher energy state on an energy-level diagram to represent the additional kinetic energy.

Figure 43.22 shows a half-filled band in a metal at $T = 0$ K, where the blue region represents levels filled with electrons. Because electrons obey Fermi–Dirac statistics, all levels below the Fermi energy are filled with electrons and all levels above the Fermi energy are empty. The Fermi energy lies in the band at the highest filled state. At temperatures slightly greater than 0 K, some electrons are thermally excited to levels above E_F, but overall there is little change from the 0 K case. If a potential difference is applied to the metal, however, electrons having energies near the Fermi energy require only a small amount of additional energy from the applied electric field to reach nearby empty energy states above the Fermi energy. Therefore, electrons in a metal experiencing only a weak applied electric field are free to move because many empty levels are available close to the occupied energy levels. The model of metals based on band theory demonstrates that metals are excellent electrical conductors.

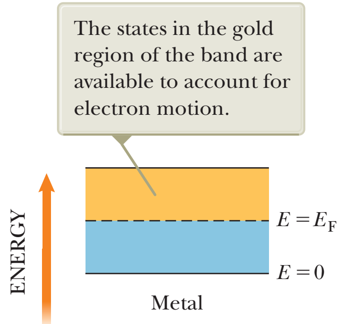


Figure 43.22 Half-filled band of a metal, an electrical conductor. At $T = 0$ K, the Fermi energy lies in the middle of the band.

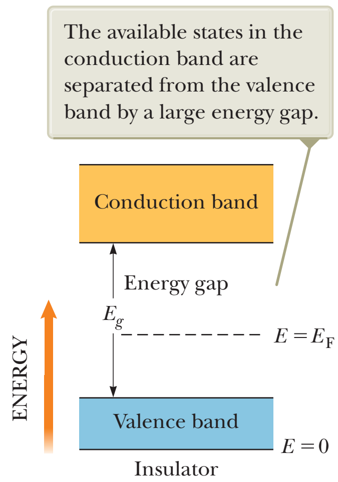


Figure 43.23 An electrical insulator at $T = 0$ K has a filled valence band and an empty conduction band. The Fermi level lies somewhere between these bands in the region known as the energy gap.

Insulators

Now consider the two outermost energy bands of a material in which the lower band is filled with electrons and the higher band is empty at 0 K (Fig. 43.23). The lower, filled band is called the **valence band,** and the upper, empty band is the **conduction band.** (The conduction band is the one that is partially filled in a metal.) It is common to refer to the energy separation between the valence and conduction bands as the **energy gap** E_g of the material. The Fermi energy lies somewhere in the energy gap[4] as shown in Figure 43.23.

Suppose a material has a relatively large energy gap of, for example, approximately 5 eV. At 300 K (room temperature), $k_B T = 0.025$ eV, which is much smaller than the energy gap. At such temperatures, the Fermi–Dirac distribution predicts that very few electrons are thermally excited into the conduction band. There are no available states that lie close in energy above the valence band and into which electrons can move upward to account for the extra kinetic energy associated with motion through the material in response to an electric field. Consequently, the electrons do not move; the material is an insulator. Although an insulator has many vacant states in its conduction band that can accept electrons, these states are separated from the filled states by a large energy gap. Only a few electrons occupy these states, so the overall electrical conductivity of insulators is very small.

Table 43.3 **Energy-Gap Values for Some Semiconductors**

Crystal	E_g (eV) 0 K	300 K
Si	1.17	1.14
Ge	0.74	0.67
InP	1.42	1.34
GaP	2.32	2.26
GaAs	1.52	1.42
CdS	2.58	2.42
CdTe	1.61	1.56
ZnO	3.44	3.2
ZnS	3.91	3.6

Semiconductors

Semiconductors have the same type of band structure as an insulator, but the energy gap is much smaller, on the order of 1 eV. Table 43.3 shows the energy gaps for some representative materials. The band structure of a semiconductor is shown in Figure 43.24. Because the Fermi level is located near the middle of the gap for a semiconductor and E_g is small, appreciable numbers of electrons are thermally excited from the valence band to the conduction band. Because of the many empty levels above the thermally filled levels in the conduction band, a small applied potential difference can easily raise the electrons in the conduction band into available energy states, resulting in a moderate current.

At $T = 0$ K, all electrons in these materials are in the valence band and no energy is available to excite them across the energy gap. Therefore, semiconductors are poor conductors at very low temperatures. Because the thermal excitation of electrons across the narrow gap is more probable at higher temperatures, the conductivity of semiconductors increases rapidly with temperature, contrasting sharply with the conductivity of metals, which decreases slowly with increasing temperature.

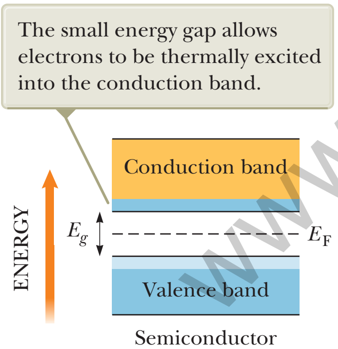


Figure 43.24 Band structure of a semiconductor at ordinary temperatures ($T \approx 300$ K). The energy gap is much smaller than in an insulator.

Charge carriers in a semiconductor can be negative, positive, or both. When an electron moves from the valence band into the conduction band, it leaves behind a vacant site, called a **hole,** in the otherwise filled valence band. This hole (electron-deficient site) acts as a charge carrier in the sense that a free electron from a nearby site can transfer into the hole. Whenever an electron does so, it creates a new hole at the site it abandoned. Therefore, the net effect can be viewed as the hole migrating through the material in the direction opposite the direction of electron movement. The hole behaves as if it were a particle with a positive charge $+e$.

A pure semiconductor crystal containing only one element or one compound is called an **intrinsic semiconductor.** In these semiconductors, there are equal numbers of conduction electrons and holes. Such combinations of charges are called **electron–hole pairs.** In the presence of an external electric field, the holes move in the direction of the field and the conduction electrons move in the direction opposite the field (Fig. 43.25). Because the electrons and holes have opposite signs, both motions correspond to a current in the same direction.

[4]We defined the Fermi energy as the energy of the highest filled state at $T = 0$, which might suggest that the Fermi energy should be at the top of the valence band in Figure 43.23. A more sophisticated general treatment of the Fermi energy, however, shows that it is located at that energy at which the probability of occupation is one-half (see Fig. 43.16b). According to this definition, the Fermi energy lies in the energy gap between the bands.

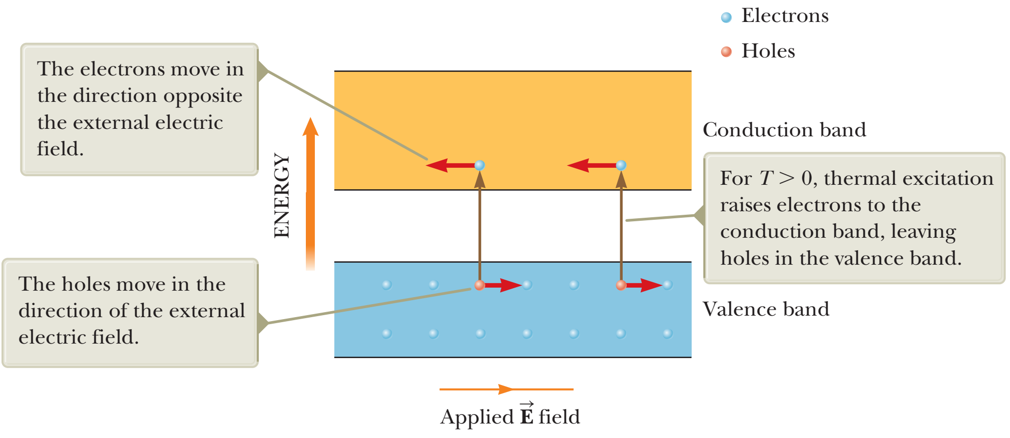


Figure 43.25 Movement of charges (holes and electrons) in an intrinsic semiconductor.

Quick Quiz 43.4 Consider the data on three materials given in the table.

Material	Conduction Band	E_g
A	Empty	1.2 eV
B	Half full	1.2 eV
C	Empty	8.0 eV

Identify each material as a conductor, an insulator, or a semiconductor.

Doped Semiconductors

When impurities are added to a semiconductor, both the band structure of the semiconductor and its resistivity are modified. The process of adding impurities, called **doping,** is important in controlling the conductivity of semiconductors. For example, when an atom containing five outer-shell electrons, such as arsenic, is added to a Group IV semiconductor, four of the electrons form covalent bonds with atoms of the semiconductor and one is left over (Fig. 43.26a). This extra electron is nearly free of its parent atom and can be modeled as having an energy level that lies in the energy gap, immediately below the conduction band (Fig. 43.26b). Such a pentavalent atom in effect donates an electron to the structure and hence is referred to as a **donor atom.** Because the spacing between the energy level of the electron of the donor atom and the bottom of the conduction band is very

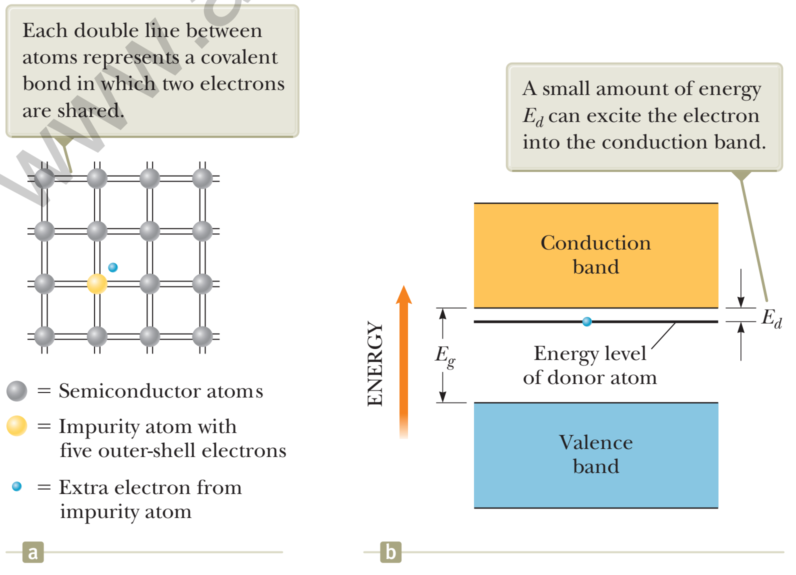


Figure 43.26 (a) Two-dimensional representation of a semiconductor consisting of Group IV atoms (gray) and an impurity atom (yellow) that has five outer-shell electrons. (b) Energy-band diagram for a semiconductor in which the nearly free electron of the impurity atom lies in the energy gap, immediately below the bottom of the conduction band.

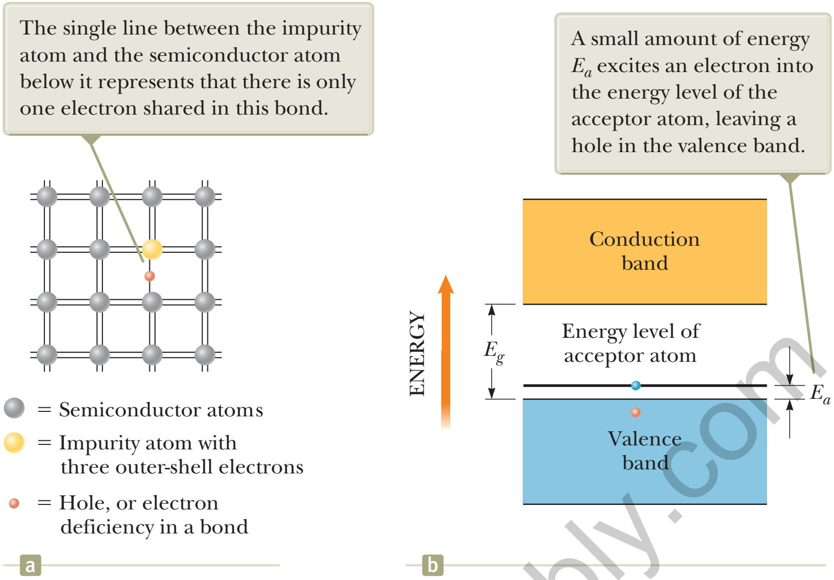


Figure 43.27 (a) Two-dimensional representation of a semiconductor consisting of Group IV atoms (gray) and an impurity atom (yellow) having three outer-shell electrons. (b) Energy-band diagram for a semiconductor in which the energy level associated with the trivalent impurity atom lies in the energy gap, immediately above the top of the valence band.

small (typically, approximately 0.05 eV), only a small amount of thermal excitation is needed to cause this electron to move into the conduction band. (Recall that the average energy of an electron at room temperature is approximately $k_BT \approx$ 0.025 eV.) Semiconductors doped with donor atoms are called ***n*-type semiconductors** because the majority of charge carriers are electrons, which are ***n***egatively charged.

If a Group IV semiconductor is doped with atoms containing three outer-shell electrons, such as indium and aluminum, the three electrons form covalent bonds with neighboring semiconductor atoms, leaving an electron deficiency—a hole—where the fourth bond would be if an impurity-atom electron were available to form it (Fig. 43.27a). This situation can be modeled by placing an energy level in the energy gap, immediately above the valence band, as in Figure 43.27b. An electron from the valence band has enough energy at room temperature to fill this impurity level, leaving behind a hole in the valence band. This hole can carry current in the presence of an electric field. Because a trivalent atom accepts an electron from the valence band, such impurities are referred to as **acceptor atoms.** A semiconductor doped with trivalent (acceptor) impurities is known as a ***p*-type semiconductor** because the majority of charge carriers are ***p***ositively charged holes.

When conduction in a semiconductor is the result of acceptor or donor impurities, the material is called an **extrinsic semiconductor.** The typical range of doping densities for extrinsic semiconductors is 10^{13} to 10^{19} cm^{-3}, whereas the electron density in a typical semiconductor is roughly 10^{21} cm^{-3}.

43.7 Semiconductor Devices

The electronics of the first half of the 20th century was based on vacuum tubes, in which electrons pass through empty space between a cathode and an anode. We have seen vacuum tube devices in Figure 29.6 (the television picture tube), Figure 29.10 (circular electron beam), Figure 29.15a (Thomson's apparatus for measuring e/m_e for the electron), and Figure 40.9 (photoelectric effect apparatus).

The transistor was invented in 1948, leading to a shift away from vacuum tubes and toward semiconductors as the basis of electronic devices. This phase of electronics has been under way for several decades. As discussed in Chapter 41, there may be a new phase of electronics in the near future using nanotechnological devices employing quantum dots and other nanoscale structures.

In this section, we discuss electronic devices based on semiconductors, which are still in wide use and will be for many years to come.

The Junction Diode

A fundamental unit of a semiconductor device is formed when a *p*-type semiconductor is joined to an *n*-type semiconductor to form a ***p–n* junction.** A **junction diode** is a device that is based on a single *p–n* junction. The role of a diode of any type is to pass current in one direction but not the other. Therefore, it acts as a one-way valve for current.

The *p–n* junction shown in Figure 43.28a consists of three distinct regions: a *p* region, an *n* region, and a small area that extends several micrometers to either side of the interface, called a *depletion region*.

The depletion region may be visualized as arising when the two halves of the junction are brought together. The mobile *n*-side donor electrons nearest the junction (deep-blue area in Fig. 43.28a) diffuse to the *p* side and fill holes located there, leaving behind immobile positive ions. While this process occurs, we can model the holes that are being filled as diffusing to the *n* side, leaving behind a region (brown area in Fig. 43.28a) of fixed negative ions.

Because the two sides of the depletion region each carry a net charge, an internal electric field on the order of 10^4 to 10^6 V/cm exists in the depletion region (see Fig. 43.28b). This field produces an electric force on any remaining mobile charge carriers that sweeps them out of the depletion region, so named because it is a region depleted of mobile charge carriers. This internal electric field creates an internal potential difference ΔV_0 that prevents further diffusion of holes and electrons across the junction and thereby ensures zero current in the junction when no potential difference is applied.

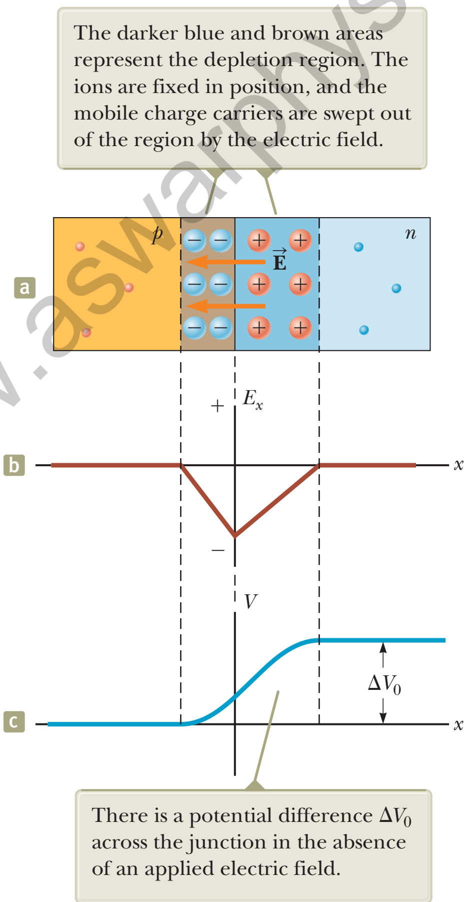


Figure 43.28 (a) Physical arrangement of a *p–n* junction. (b) Component E_x of the internal electric field versus x for the *p–n* junction. (c) Internal electric potential difference ΔV versus x for the *p–n* junction.

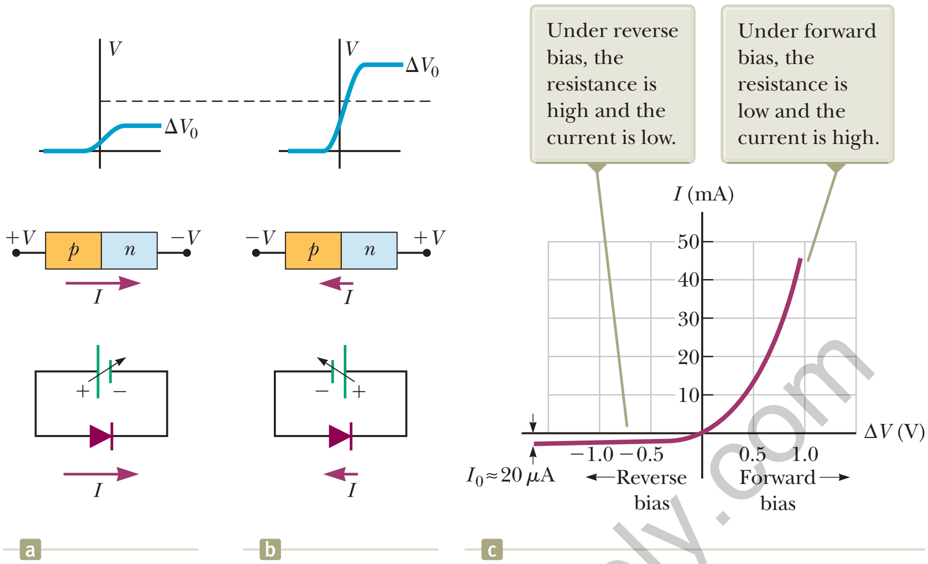


Figure 43.29 (a) A p–n junction under forward bias. The middle diagram shows the potentials applied at the ends of the junction. Below that is a circuit diagram showing a battery with an adjustable voltage. The upper diagram shows how the potential varies across the junction. The dashed line shows the potential difference across the unbiased junction. (b) When the battery is reversed and the p–n junction is under reverse bias, the current is very small. (c) The characteristic curve for a real p–n junction.

The operation of the junction as a diode is easiest to understand in terms of the potential difference graph shown in Figure 43.28c. If a voltage ΔV is applied to the junction such that the p side is connected to the positive terminal of a voltage source as shown in Figure 43.29a, the internal potential difference ΔV_0 across the junction decreases as shown at the top of the figure; the decrease results in a current that increases exponentially with increasing forward voltage, or *forward bias.* For *reverse bias* (where the n side of the junction is connected to the positive terminal of a voltage source), the internal potential difference ΔV_0 increases with increasing reverse bias as in Figure 43.29b; the increase results in a very small reverse current that quickly reaches a saturation value I_0. The current–voltage relationship for an ideal diode is

$$I = I_0\left(e^{e\,\Delta V/k_B T} - 1\right) \qquad \textbf{(43.27)}$$

where the first e is the base of the natural logarithm, the second e represents the magnitude of the electron charge, k_B is Boltzmann's constant, and T is the absolute temperature. Figure 43.29c shows an I–ΔV plot characteristic of a real p–n junction, demonstrating the diode behavior.

Light-Emitting and Light-Absorbing Diodes

Light-emitting diodes (LEDs) and semiconductor lasers are common examples of devices that depend on the behavior of semiconductors. LEDs are used in LCD television displays, household lighting, flashlights, and camera flash units. Semiconductor lasers are often used for pointers in presentations and in playback equipment for digitally recorded information.

Light emission and absorption in semiconductors is similar to light emission and absorption by gaseous atoms except that in the discussion of semiconductors we must incorporate the concept of energy bands rather than the discrete energy levels in single atoms. As shown in Figure 43.30a, an electron excited electrically into the conduction band can easily recombine with a hole (especially if the electron is injected into a p region). As this recombination takes place, a photon of energy E_g is emitted. With proper design of the semiconductor and the associated plastic envelope or mirrors, the light from a large number of these transitions serves as the source of an LED or a semiconductor laser.

Conversely, an electron in the valence band may absorb an incoming photon of light and be promoted to the conduction band, leaving a hole behind (Fig. 43.30b). This absorbed energy can be used to operate an electrical circuit.

One device that operates on this principle is the **photovoltaic solar cell,** which appears in many handheld calculators. An early large-scale application of arrays of

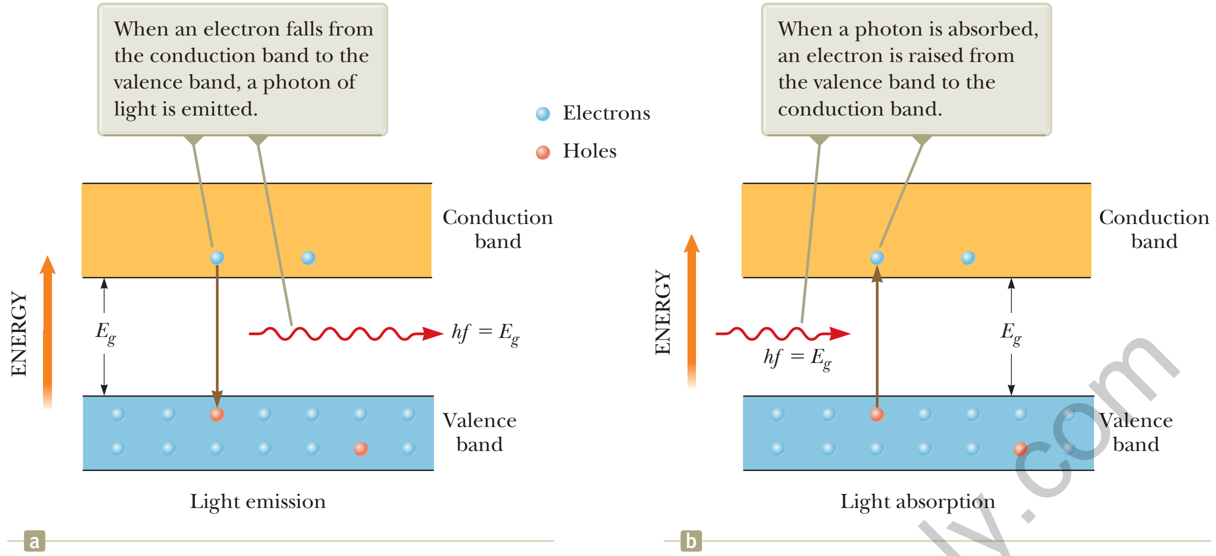


Figure 43.30 (a) Light emission from a semiconductor. (b) Light absorption by a semiconductor.

photovoltaic cells is the energy supply for orbiting spacecraft. The solar panels of the Hubble Space Telescope can be seen in the chapter-opening photograph for Chapter 38 on page 1160.

During the early years of the current century, application of photovoltaics for ground-based generation of electricity has been one of the world's fastest-growing energy technologies. At the time of this printing, the global generation of energy by means of photovoltaics is over 70 GW. A homeowner can install arrays of photovoltaic panels on the roof of his or her house and generate enough energy to operate the home as well as feed excess energy back into the electrical grid. Several photovoltaic power plants have recently been completed, including the Agua Caliente Solar Project in Arizona (200 MW completed in 2012, and 397 MW projected at completion), the Golmud Solar Park in China (200 MW), and the Charanka Solar Park in India (214 MW completed in 2012, and 500 MW projected at completion), the latter of which will be one location in the Gujarat Solar Park, a collection of several sites that is hoped to eventually supply close to 1 GW of power.

Example 43.6 Where's the Remote?

Estimate the band gap of the semiconductor in the infrared LED of a typical television remote control.

SOLUTION

Conceptualize Imagine electrons in Figure 43.30a falling from the conduction band to the valence band, emitting infrared photons in the process.

Categorize We use concepts discussed in this section, so we categorize this example as a substitution problem.

In Chapter 34, we learned that the wavelength of infrared light ranges from 700 nm to 1 mm. Let's pick a number that is easy to work with, such as 1 000 nm (which is not a bad estimate because remote controls typically operate in the range of 880 to 950 nm.)

Estimate the energy hf of the photons from the remote control:

$$E = hf = \frac{hc}{\lambda} = \frac{1\,240 \text{ eV}\cdot\text{nm}}{1\,000 \text{ nm}} = \boxed{1.2 \text{ eV}}$$

This value corresponds to an energy gap E_g of approximately 1.2 eV in the LED's semiconductor.

The Transistor

The invention of the transistor by John Bardeen (1908–1991), Walter Brattain (1902–1987), and William Shockley (1910–1989) in 1948 totally revolutionized the world of electronics. For this work, these three men shared the Nobel Prize in Physics in 1956. By 1960, the transistor had replaced the vacuum tube in many electronic applications. The advent of the transistor created a multitrillion-dollar industry that produces such popular devices as personal computers, wireless keyboards, smartphones, electronic book readers, and computer tablets.

A **junction transistor** consists of a semiconducting material in which a very narrow *n* region is sandwiched between two *p* regions or a *p* region is sandwiched between two *n* regions. In either case, the transistor is formed from two *p*–*n* junctions. These types of transistors were used widely in the early days of semiconductor electronics.

During the 1960s, the electronics industry converted many electronic applications from the junction transistor to the **field-effect transistor,** which is much easier to manufacture and just as effective. Figure 43.31a shows the structure of a very common device, the **MOSFET,** or **metal-oxide-semiconductor field-effect transistor.** You are likely using millions of MOSFET devices when you are working on your computer.

There are three metal connections (the M in MOSFET) to the transistor: the *source, drain,* and *gate*. The source and drain are connected to *n*-type semiconductor regions (the S in MOSFET) at either end of the structure. These regions are connected by a narrow channel of additional *n*-type material, the *n* channel. The source and drain regions and the *n* channel are embedded in a *p*-type substrate material, which forms a depletion region, as in the junction diode, along the bottom of the *n* channel. (Depletion regions also exist at the junctions underneath the source and drain regions, but we will ignore them because the operation of the device depends primarily on the behavior in the channel.)

The gate is separated from the *n* channel by a layer of insulating silicon dioxide (the O in MOSFET, for oxide). Therefore, it does not make electrical contact with the rest of the semiconducting material.

Imagine that a voltage source ΔV_{SD} is applied across the source and drain as shown in Figure 43.31b. In this situation, electrons flow through the upper region of the *n* channel. Electrons cannot flow through the depletion region in the lower part of the *n* channel because this region is depleted of charge carriers. Now a second voltage ΔV_{SG} is applied across the source and gate as in Figure 43.31c. The positive potential on the gate electrode results in an electric field below the gate that is directed downward in the *n* channel (the field in "field-effect"). This electric field exerts upward forces on electrons in the region below the gate, causing them to move into the *n* channel. Consequently, the depletion region becomes smaller,

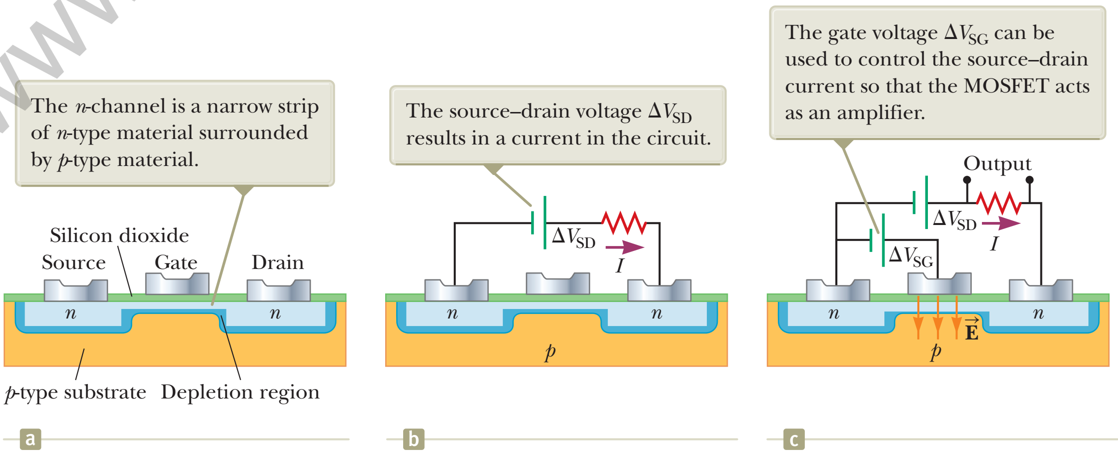


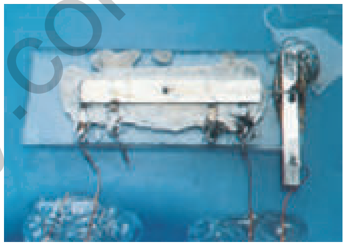
Figure 43.31 (a) The structure of a metal-oxide-semiconductor field-effect transistor (MOSFET). (b) A source–drain voltage is applied. (c) A gate voltage is applied.

widening the area through which there is current between the top of the n channel and the depletion region. As the area becomes wider, the current increases.

If a varying voltage, such as that generated from music stored in the memory of a smartphone, is applied to the gate, the area through which the source–drain current exists varies in size according to the varying gate voltage. A small variation in gate voltage results in a large variation in current and a correspondingly large voltage across the resistor in Figure 43.31c. Therefore, the MOSFET acts as a voltage amplifier. A circuit consisting of a chain of such transistors can result in a very small initial signal from a microphone being amplified enough to drive powerful speakers at an outdoor concert.

The Integrated Circuit

Invented independently by Jack Kilby (1923–2005, Nobel Prize in Physics, 2000) at Texas Instruments in late 1958 and by Robert Noyce (1927–1990) at Fairchild Camera and Instrument in early 1959, the integrated circuit has been justly called "the most remarkable technology ever to hit mankind." Kilby's first device is shown in Figure 43.32. Integrated circuits have indeed started a "second industrial revolution" and are found at the heart of computers, watches, cameras, automobiles, aircraft, robots, space vehicles, and all sorts of communication and switching networks.



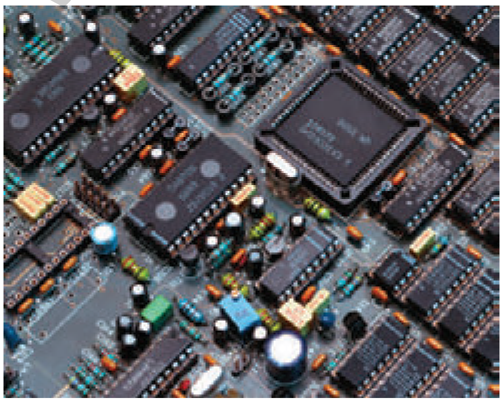
Courtesy of Texas Instruments, Inc.

Figure 43.32 Jack Kilby's first integrated circuit, tested on September 12, 1958.

In simplest terms, an **integrated circuit** is a collection of interconnected transistors, diodes, resistors, and capacitors fabricated on a single piece of silicon known as a *chip*. Contemporary electronic devices often contain many integrated circuits (Fig. 43.33). State-of-the-art chips easily contain several million components within a 1-cm^2 area, and the number of components per square inch has increased steadily since the integrated circuit was invented. The dramatic advances in chip technology can be seen by looking at microchips manufactured by Intel. The 4004 chip, introduced in 1971, contained 2 300 transistors. This number increased to 3.2 million 24 years later in 1995 with the Pentium processor. Sixteen years later, the Core i7 Sandy Bridge processor introduced in November 2011 contained 2 270 million transistors.

Integrated circuits were invented partly to solve the interconnection problem spawned by the transistor. In the era of vacuum tubes, power and size considerations of individual components set modest limits on the number of components that could be interconnected in a given circuit. With the advent of the tiny, low-power, highly reliable transistor, design limits on the number of components disappeared and were replaced by the problem of wiring together hundreds of thousands of components. The magnitude of this problem can be appreciated when we consider that second-generation computers (consisting of discrete transistors rather than integrated circuits) contained several hundred thousand components requiring more than a million joints that had to be hand-soldered and tested.

In addition to solving the interconnection problem, integrated circuits possess the advantages of miniaturization and fast response, two attributes critical for high-speed computers. Because the response time of a circuit depends on the time



Steve Allen/Brand X/Getty Images

Figure 43.33 Integrated circuits are prevalent in many electronic devices. All the flat circuit elements with black-topped surfaces in this photograph are integrated circuits.

interval required for electrical signals traveling at the speed of light to pass from one component to another, miniaturization and close packing of components result in fast response times.

43.8 Superconductivity

We learned in Section 27.5 that there is a class of metals and compounds known as **superconductors** whose electrical resistance decreases to virtually zero below a certain temperature T_c called the *critical temperature* (Table 27.3). Let's now look at these amazing materials in greater detail, using what we know about the properties of solids to help us understand the behavior of superconductors.

Let's start by examining the Meissner effect, introduced in Section 30.6 as the exclusion of magnetic flux from the interior of superconductors. The Meissner effect is illustrated in Figure 43.34 for a superconducting material in the shape of a long cylinder. Notice that the magnetic field penetrates the cylinder when its temperature is greater than T_c (Fig. 43.34a). As the temperature is lowered to below T_c, however, the field lines are spontaneously expelled from the interior of the superconductor (Fig. 43.34b). Therefore, a superconductor is more than a perfect conductor (resistivity $\rho = 0$); it is also a perfect diamagnet ($\vec{\mathbf{B}} = 0$). The property that $\vec{\mathbf{B}} = 0$ in the interior of a superconductor is as fundamental as the property of zero resistance. If the magnitude of the applied magnetic field exceeds a critical value B_c, defined as the value of B that destroys a material's superconducting properties, the field again penetrates the sample.

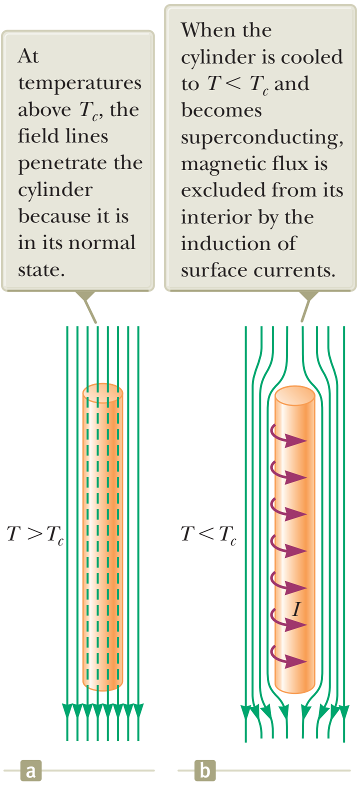


Figure 43.34 A superconductor in the form of a long cylinder in the presence of an external magnetic field.

Because a superconductor is a perfect diamagnet, it repels a permanent magnet. In fact, one can perform a demonstration of the Meissner effect by floating a small permanent magnet above a superconductor and achieving magnetic levitation as seen in Figure 30.27 in Section 30.6.

Recall from our study of electricity that a good conductor expels static electric fields by moving charges to its surface. In effect, the surface charges produce an electric field that exactly cancels the externally applied field inside the conductor. In a similar manner, a superconductor expels magnetic fields by forming surface currents. To see why that happens, consider again the superconductor shown in Figure 43.34. Let's assume the sample is initially at a temperature $T > T_c$ as illustrated in Figure 43.34a so that the magnetic field penetrates the cylinder. As the cylinder is cooled to a temperature $T < T_c$, the field is expelled as shown in Figure 43.34b. Surface currents induced on the superconductor's surface produce a magnetic field that exactly cancels the externally applied field inside the superconductor. As you would expect, the surface currents disappear when the external magnetic field is removed.

A successful theory for superconductivity in metals was published in 1957 by John Bardeen, L. N. Cooper (b. 1930), and J. R. Schrieffer (b. 1931); it is generally called BCS theory, based on the first letters of their last names. This theory led to a Nobel Prize in Physics for the three scientists in 1972. In this theory, two electrons can interact via distortions in the array of lattice ions so that there is a net attractive force between the electrons.[5] As a result, the two electrons are bound into an entity called a *Cooper pair,* which behaves like a particle with integral spin. Particles with integral spin are called *bosons.* (As noted in Pitfall Prevention 42.6, *fermions* make up another class of particles, those with half-integral spin.) An important feature of bosons is that they do not obey the Pauli exclusion principle. Consequently, at very low temperatures, it is possible for all bosons in a collection of such particles to be in the lowest quantum state. The entire collection of Cooper pairs in the metal is described by a single wave function. Above the energy level associated with this wave function is an energy gap equal to the binding energy of a Cooper pair. Under

[5]A highly simplified explanation of this attraction between electrons is as follows. The attractive Coulomb force between one electron and the surrounding positively charged lattice ions causes the ions to move inward slightly toward the electron. As a result, there is a higher concentration of positive charge in this region than elsewhere in the lattice. A second electron is attracted to the higher concentration of positive charge.

the action of an applied electric field, the Cooper pairs experience an electric force and move through the metal. A random scattering event of a Cooper pair from a lattice ion would represent resistance to the electric current. Such a collision would change the energy of the Cooper pair because some energy would be transferred to the lattice ion. There are no available energy levels below that of the Cooper pair (it is already in the lowest state), however, and none available above because of the energy gap. As a result, collisions do not occur and there is no resistance to the movement of Cooper pairs.

An important development in physics that elicited much excitement in the scientific community was the discovery of high-temperature copper oxide-based superconductors. The excitement began with a 1986 publication by J. Georg Bednorz (b. 1950) and K. Alex Müller (b. 1927), scientists at the IBM Zurich Research Laboratory in Switzerland. In their seminal paper,[6] Bednorz and Müller reported strong evidence for superconductivity at 30 K in an oxide of barium, lanthanum, and copper. They were awarded the Nobel Prize in Physics in 1987 for their remarkable discovery. Shortly thereafter, a new family of compounds was open for investigation and research activity in the field of superconductivity proceeded vigorously. In early 1987, groups at the University of Alabama at Huntsville and the University of Houston announced superconductivity at approximately 92 K in an oxide of yttrium, barium, and copper ($YBa_2Cu_3O_7$). Later that year, teams of scientists from Japan and the United States reported superconductivity at 105 K in an oxide of bismuth, strontium, calcium, and copper. Superconductivity at temperatures as high as 150 K have been reported in an oxide containing mercury. In 2006, Japanese scientists discovered superconductivity for the first time in iron-based materials, beginning with LaFePO, with a critical temperature of 4 K. The highest critical temperature that has been reported so far in the iron-based materials is 55 K, a milestone held by fluorine-doped SmFeAsO. These newly discovered materials have rejuvenated the field of high-T_c superconductivity. Today, one cannot rule out the possibility of room-temperature superconductivity, and the mechanisms responsible for the behavior of high-temperature superconductors are still under investigation. The search for novel superconducting materials continues both for scientific reasons and because practical applications become more probable and widespread as the critical temperature is raised.

Although BCS theory was very successful in explaining superconductivity in metals, there is currently no widely accepted theory for high-temperature superconductivity. It remains an area of active research.

Summary

Concepts and Principles

Two or more atoms combine to form molecules because of a net attractive force between the atoms. The mechanisms responsible for molecular bonding can be classified as follows:

- **Ionic bonds** form primarily because of the Coulomb attraction between oppositely charged ions. Sodium chloride (NaCl) is one example.
- **Covalent bonds** form when the constituent atoms of a molecule share electrons. For example, the two electrons of the H_2 molecule are equally shared between the two nuclei.
- **Van der Waals bonds** are weak electrostatic bonds between molecules or between atoms that do not form ionic or covalent bonds. These bonds are responsible for the condensation of noble gas atoms and nonpolar molecules into the liquid phase.
- **Hydrogen bonds** form between the center of positive charge in a polar molecule that includes one or more hydrogen atoms and the center of negative charge in another polar molecule.

continued

[6]J. G. Bednorz and K. A. Müller, *Z. Phys. B* **64:**189, 1986.

The allowed values of the rotational energy of a diatomic molecule are

$$E_{\text{rot}} = E_J = \frac{\hbar^2}{2I}J(J+1) \quad J = 0, 1, 2, \ldots \quad \textbf{(43.6)}$$

where I is the moment of inertia of the molecule and J is an integer called the **rotational quantum number.** The selection rule for transitions between rotational states is $\Delta J = \pm 1$.

The allowed values of the vibrational energy of a diatomic molecule are

$$E_{\text{vib}} = \left(v + \tfrac{1}{2}\right)\frac{h}{2\pi}\sqrt{\frac{k}{\mu}} \quad v = 0, 1, 2, \ldots \quad \textbf{(43.10)}$$

where v is the **vibrational quantum number,** k is the force constant of the "effective spring" bonding the molecule, and μ is the **reduced mass** of the molecule. The selection rule for allowed vibrational transitions is $\Delta v = \pm 1$, and the energy difference between any two adjacent levels is the same, regardless of which two levels are involved.

Bonding mechanisms in solids can be classified in a manner similar to the schemes for molecules. For example, the Na^+ and Cl^- ions in NaCl form **ionic bonds,** whereas the carbon atoms in diamond form **covalent bonds.** The **metallic bond** is characterized by a net attractive force between positive ion cores and the mobile free electrons of a metal.

In the **free-electron theory of metals,** the free electrons fill the quantized levels in accordance with the Pauli exclusion principle. The number of states per unit volume available to the conduction electrons having energies between E and $E + dE$ is

$$N(E)\,dE = \left(\frac{8\sqrt{2}\,\pi m_e^{3/2}}{h^3}E^{1/2}\right)\left(\frac{1}{e^{(E-E_F)/k_BT}+1}\right)dE \quad \textbf{(43.22)}$$

where E_F is the **Fermi energy.** At $T = 0$ K, all levels below E_F are filled, all levels above E_F are empty, and

$$E_F(0) = \frac{h^2}{2m_e}\left(\frac{3n_e}{8\pi}\right)^{2/3} \quad \textbf{(43.25)}$$

where n_e is the total number of conduction electrons per unit volume. Only those electrons having energies near E_F can contribute to the electrical conductivity of the metal.

In a crystalline solid, the energy levels of the system form a set of **bands.** Electrons occupy the lowest energy states, with no more than one electron per state. Energy gaps are present between the bands of allowed states.

A **semiconductor** is a material having an energy gap of approximately 1 eV and a valence band that is filled at $T = 0$ K. Because of the small energy gap, a significant number of electrons can be thermally excited from the valence band into the conduction band. The band structures and electrical properties of a Group IV semiconductor can be modified by the addition of either donor atoms containing five outer-shell electrons or acceptor atoms containing three outer-shell electrons. A semiconductor **doped** with donor impurity atoms is called an ***n*-type semiconductor,** and one doped with acceptor impurity atoms is called a ***p*-type semiconductor.**

Objective Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

1. Is each one of the following statements true or false for a superconductor below its critical temperature? (a) It can carry infinite current. (b) It must carry some nonzero current. (c) Its interior electric field must be zero. (d) Its internal magnetic field must be zero. (e) No internal energy appears when it carries electric current.

2. An infrared absorption spectrum of a molecule is shown in Figure OQ43.2. Notice that the highest peak on either side of the gap is the third peak from the gap. After this spectrum is taken, the temperature of the sample of molecules is raised to a much higher value. Compared with Figure OQ43.2, in this new spectrum is the highest absorption peak (a) at the same frequency, (b) farther from the gap, or (c) closer to the gap?

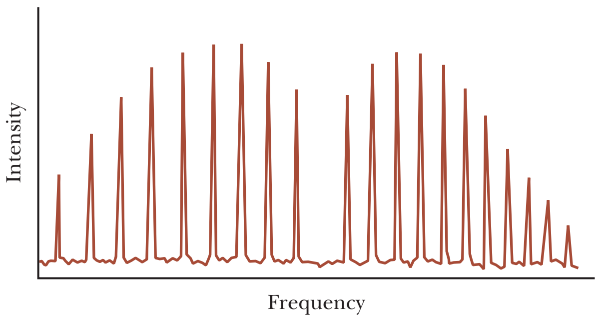


Figure OQ43.2

3. What kind of bonding likely holds the atoms together in the following solids (i), (ii), and (iii)? Choose your answers from these possibilities: (a) ionic bonding, (b) covalent bonding, and (c) metallic bonding. **(i)** The solid is opaque, shiny, flexible, and a good electric conductor. **(ii)** The crystal is transparent, brittle, and soluble in water. It is a poor conductor of electricity. **(iii)** The crystal is opaque, brittle, very hard, and a good electric insulator.

4. The Fermi energy for silver is 5.48 eV. In a piece of solid silver, free-electron energy levels are measured near 2 eV and near 6 eV. **(i)** Near which of these energies are the energy levels closer together? (a) 2 eV (b) 6 eV (c) The spacing is the same. **(ii)** Near which of these energies are more electrons occupying energy levels? (a) 2 eV (b) 6 eV (c) The number of electrons is the same.

5. As discussed in Chapter 27, the conductivity of metals decreases with increasing temperature due to electron collisions with vibrating atoms. In contrast, the conductivity of semiconductors increases with increasing temperature. What property of a semiconductor is responsible for this behavior? (a) Atomic vibrations decrease as temperature increases. (b) The number of conduction electrons and the number of holes increase steeply with increasing temperature. (c) The energy gap decreases with increasing temperature. (d) Electrons do not collide with atoms in a semiconductor.

6. **(i)** Should you expect an *n*-type doped semiconductor to have (a) higher, (b) lower, or (c) the same conductivity as an intrinsic (pure) semiconductor? **(ii)** Should you expect a *p*-type doped semiconductor to have (a) higher, (b) lower, or (c) the same conductivity as an intrinsic (pure) semiconductor?

7. Consider a typical material composed of covalently bonded diatomic molecules. Rank the following energies from the largest in magnitude to the smallest in magnitude. (a) the latent heat of fusion per molecule (b) the molecular binding energy (c) the energy of the first excited state of molecular rotation (d) the energy of the first excited state of molecular vibration

Conceptual Questions

1. denotes answer available in *Student Solutions Manual/Study Guide*

Conceptual Questions 5 and 6 in Chapter 27 can be assigned with this chapter.

1. The energies of photons of visible light range between the approximate values 1.8 eV and 3.1 eV. Explain why silicon, with an energy gap of 1.14 eV at room temperature (see Table 43.3), appears opaque, whereas diamond, with an energy gap of 5.47 eV, appears transparent.

2. Discuss the three major forms of excitation of a molecule (other than translational motion) and the relative energies associated with these three forms.

3. How can the analysis of the rotational spectrum of a molecule lead to an estimate of the size of that molecule?

4. Pentavalent atoms such as arsenic are donor atoms in a semiconductor such as silicon, whereas trivalent atoms such as indium are acceptors. Inspect the periodic table in Appendix C and determine what other elements might make good donors or acceptors.

5. When a photon is absorbed by a semiconductor, an electron–hole pair is created. Give a physical explanation of this statement using the energy-band model as the basis for your description.

6. (a) Discuss the differences in the band structures of metals, insulators, and semiconductors. (b) How does the band-structure model enable you to understand the electrical properties of these materials better?

7. (a) What essential assumptions are made in the free-electron theory of metals? (b) How does the energy-band model differ from the free-electron theory in describing the properties of metals?

8. How do the vibrational and rotational levels of heavy hydrogen (D_2) molecules compare with those of H_2 molecules?

9. Discuss models for the different types of bonds that form stable molecules.

10. Discuss the differences between crystalline solids, amorphous solids, and gases.

Problems

ENHANCED WebAssign The problems found in this chapter may be assigned online in Enhanced WebAssign

1. straightforward; **2.** intermediate; **3.** challenging

1. full solution available in the *Student Solutions Manual/Study Guide*

AMT Analysis Model tutorial available in Enhanced WebAssign

GP Guided Problem

M Master It tutorial available in Enhanced WebAssign

W Watch It video solution available in Enhanced WebAssign

Section 43.1 Molecular Bonds

1. A van der Waals dispersion force between helium atoms produces a very shallow potential well, with a depth on the order of 1 meV. At approximately what temperature would you expect helium to condense?

2. W **Review.** A K^+ ion and a Cl^- ion are separated by a distance of 5.00×10^{-10} m. Assuming the two ions act like charged particles, determine (a) the force each ion exerts on the other and (b) the potential energy of the two-ion system in electron volts.

3. Potassium chloride is an ionically bonded molecule that is sold as a salt substitute for use in a low-sodium diet. The electron affinity of chlorine is 3.6 eV. An energy input of 0.70 eV is required to form separate K^+ and Cl^- ions from separate K and Cl atoms. What is the ionization energy of K?

4. In the potassium iodide (KI) molecule, assume the K and I atoms bond ionically by the transfer of one electron from K to I. (a) The ionization energy of K is 4.34 eV, and the electron affinity of I is 3.06 eV. What energy is needed to transfer an electron from K to I, to form K^+ and I^- ions from neutral atoms? This quantity is sometimes called the activation energy E_a. (b) A model potential energy function for the KI molecule is the Lennard–Jones potential:

$$U(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + E_a$$

where r is the internuclear separation distance and ϵ and σ are adjustable parameters. The E_a term is added to ensure the correct asymptotic behavior at large r. At the equilibrium separation distance, $r = r_0 = 0.305$ nm, $U(r)$ is a minimum, and $dU/dr = 0$. In addition, $U(r_0)$ is the negative of the dissociation energy: $U(r_0) = -3.37$ eV. Find σ and ϵ. (c) Calculate the force needed to break up a KI molecule. (d) Calculate the force constant for small oscillations about $r = r_0$. *Suggestion:* Set $r = r_0 + s$, where $s/r_0 << 1$, and expand $U(r)$ in powers of s/r_0 up to second-order terms.

5. One description of the potential energy of a diatomic molecule is given by the Lennard–Jones potential,

$$U = \frac{A}{r^{12}} - \frac{B}{r^6}$$

where A and B are constants and r is the separation distance between the atoms. For the H_2 molecule, take $A = 0.124 \times 10^{-120}$ eV · m^{12} and $B = 1.488 \times 10^{-60}$ eV · m^6. Find (a) the separation distance r_0 at which the energy of the molecule is a minimum and (b) the energy E required to break up the H_2 molecule.

6. One description of the potential energy of a diatomic molecule is given by the Lennard–Jones potential,

$$U = \frac{A}{r^{12}} - \frac{B}{r^6}$$

where A and B are constants and r is the separation distance between the atoms. Find, in terms of A and B, (a) the value r_0 at which the energy is a minimum and (b) the energy E required to break up a diatomic molecule.

Section 43.2 Energy States and Spectra of Molecules

7. The CO molecule makes a transition from the $J = 1$ to the $J = 2$ rotational state when it absorbs a photon of frequency 2.30×10^{11} Hz. (a) Find the moment of inertia of this molecule from these data. (b) Compare your answer with that obtained in Example 43.1 and comment on the significance of the two results.

8. The cesium iodide (CsI) molecule has an atomic separation of 0.127 nm. (a) Determine the energy of the second excited rotational state, with $J = 2$. (b) Find the frequency of the photon absorbed in the $J = 1$ to $J = 2$ transition.

9. M An HCl molecule is excited to its second rotational energy level, corresponding to $J = 2$. If the distance between its nuclei is 0.127 5 nm, what is the angular speed of the molecule about its center of mass?

10. The photon frequency that would be absorbed by the NO molecule in a transition from vibration state $v = 0$ to $v = 1$, with no change in rotation state, is 56.3 THz. The bond between the atoms has an effective spring constant of 1 530 N/m. (a) Use this information to calculate the reduced mass of the NO molecule. (b) Compute a value for μ using Equation 43.4. (c) Compare your results to parts (a) and (b) and explain their difference, if any.

11. Assume the distance between the protons in the H_2 molecule is 0.750×10^{-10} m. (a) Find the energy of the first excited rotational state, with $J = 1$. (b) Find the wavelength of radiation emitted in the transition from $J = 1$ to $J = 0$.

12. *Why is the following situation impossible?* The effective force constant of a vibrating HCl molecule is $k = 480$ N/m. A beam of infrared radiation of wavelength 6.20×10^3 nm is directed through a gas of HCl molecules. As a result, the molecules are excited from the ground vibrational state to the first excited vibrational state.

13. AMT The effective spring constant describing the potential energy of the HI molecule is 320 N/m and that for the HF molecule is 970 N/m. Calculate the minimum amplitude of vibration for (a) the HI molecule and (b) the HF molecule.

14. The rotational spectrum of the HCl molecule contains lines with wavelengths of 0.060 4, 0.069 0, 0.080 4, 0.096 4, and 0.120 4 mm. What is the moment of inertia of the molecule?

15. The atoms of an NaCl molecule are separated by a distance $r = 0.280$ nm. Calculate (a) the reduced mass of an NaCl molecule, (b) the moment of inertia of an NaCl molecule, and (c) the wavelength of radiation emitted when an NaCl molecule undergoes a transition from the $J = 2$ state to the $J = 1$ state.

16. A diatomic molecule consists of two atoms having masses m_1 and m_2 separated by a distance r. Show that the moment of inertia about an axis through the center of mass of the molecule is given by Equation 43.3, $I = \mu r^2$.

17. The nuclei of the O_2 molecule are separated by a distance 1.20×10^{-10} m. The mass of each oxygen atom in the molecule is 2.66×10^{-26} kg. (a) Determine the rotational energies of an oxygen molecule in electron volts for the levels corresponding to $J = 0$, 1, and 2. (b) The effective force constant k between the atoms in the oxygen molecule is 1 177 N/m. Determine the vibrational energies (in electron volts) corresponding to $v = 0$, 1, and 2.

18. Figure P43.18 is a model of a benzene molecule. All atoms lie in a plane, and the carbon atoms ($m_C = 1.99 \times 10^{-26}$ kg) form a regular hexagon, as do the hydrogen atoms ($m_H = 1.67 \times 10^{-27}$ kg). The carbon atoms are 0.110 nm apart center to center, and the adjacent carbon and hydrogen atoms are 0.100 nm apart center to center. (a) Calculate the moment of inertia of the molecule about an axis perpendicular to the plane of the paper through the center point O. (b) Determine the allowed rotational energies about this axis.

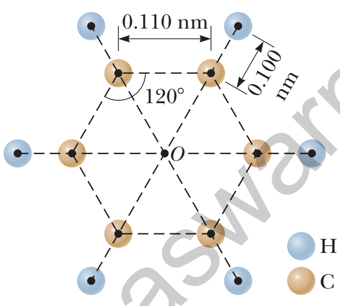


Figure P43.18

19. (a) In an HCl molecule, take the Cl atom to be the isotope ^{35}Cl. The equilibrium separation of the H and Cl atoms is 0.127 46 nm. The atomic mass of the H atom is 1.007 825 u and that of the ^{35}Cl atom is 34.968 853 u. Calculate the longest wavelength in the rotational spectrum of this molecule. (b) **What If?** Repeat the calculation in part (a), but take the Cl atom to be the isotope ^{37}Cl, which has atomic mass 36.965 903 u. The equilibrium separation distance is the same as in part (a). (c) Naturally occurring chlorine contains approximately three parts of ^{35}Cl to one part of ^{37}Cl. Because of the two different Cl masses, each line in the microwave rotational spectrum of HCl is split into a doublet as shown in Figure P43.19. Calculate the separation in wavelength between the doublet lines for the longest wavelength.

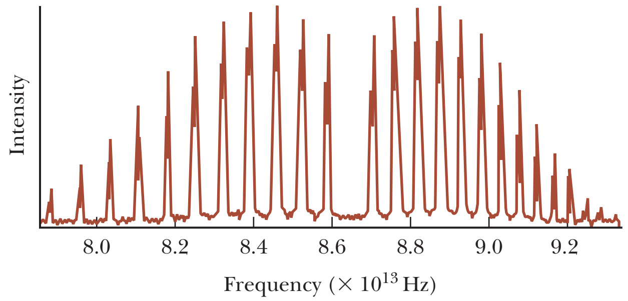


Figure P43.19 Problems 19 and 20.

20. Estimate the moment of inertia of an HCl molecule from its infrared absorption spectrum shown in Figure P43.19.

21. W An H_2 molecule is in its vibrational and rotational ground states. It absorbs a photon of wavelength 2.211 2 μm and makes a transition to the $v = 1$, $J = 1$ energy level. It then drops to the $v = 0$, $J = 2$ energy level while emitting a photon of wavelength 2.405 4 μm. Calculate (a) the moment of inertia of the H_2 molecule about an axis through its center of mass and perpendicular to the H–H bond, (b) the vibrational frequency of the H_2 molecule, and (c) the equilibrium separation distance for this molecule.

22. Photons of what frequencies can be spontaneously emitted by CO molecules in the state with $v = 1$ and $J = 0$?

23. Most of the mass of an atom is in its nucleus. Model the mass distribution in a diatomic molecule as two spheres of uniform density, each of radius 2.00×10^{-15} m and mass 1.00×10^{-26} kg, located at points along the y axis as in Figure 43.5a, and separated by 2.00×10^{-10} m. Rotation about the axis joining the nuclei in the diatomic molecule is ordinarily ignored because the first excited state would have an energy that is too high to access. To see why, calculate the ratio of the energy of the first excited state for rotation about the y axis to the energy of the first excited state for rotation about the x axis.

Section 43.3 Bonding in Solids

24. Use a magnifying glass to look at the grains of table salt that come out of a salt shaker. Compare what you see with Figure 43.11a. The distance between a sodium ion and a nearest-neighbor chlorine ion is 0.261 nm. (a) Make an order-of-magnitude estimate of the number N of atoms in a typical grain of salt. (b) **What If?** Suppose you had a number of grains of salt equal to this number N. What would be the volume of this quantity of salt?

25. Use Equation 43.18 to calculate the ionic cohesive energy for NaCl. Take $\alpha = 1.747\ 6$, $r_0 = 0.281$ nm, and $m = 8$.

26. Consider a one-dimensional chain of alternating singly-ionized positive and negative ions. Show that the potential energy associated with one of the ions and its interactions with the rest of this hypothetical crystal is

$$U(r) = -\alpha k_e \frac{e^2}{r}$$

where the Madelung constant is $\alpha = 2 \ln 2$ and r is the distance between ions. *Suggestion:* Use the series expansion for $\ln(1 + x)$.

Section 43.4 Free-Electron Theory of Metals

Section 43.5 Band Theory of Solids

27. Calculate the energy of a conduction electron in silver at 800 K, assuming the probability of finding an electron in that state is 0.950. The Fermi energy of silver is 5.48 eV at this temperature.

28. (a) State what the Fermi energy depends on according to the free-electron theory of metals and how the Fermi energy depends on that quantity. (b) Show that Equation 43.25 can be expressed as $E_F = (3.65 \times 10^{-19})n_e^{2/3}$, where E_F is in electron volts when n_e is in electrons per cubic meter. (c) According to Table 43.2, by what factor does the free-electron concentration in copper exceed that in potassium? (d) Which of these metals has the larger Fermi energy? (e) By what factor is the Fermi energy larger? (f) Explain whether this behavior is predicted by Equation 43.25.

29. When solid silver starts to melt, what is the approximate fraction of the conduction electrons that are thermally excited above the Fermi level?

30. (a) Find the typical speed of a conduction electron in copper, taking its kinetic energy as equal to the Fermi energy, 7.05 eV. (b) Suppose the copper is a current-carrying wire. How does the speed found in part (a) compare with a typical drift speed (see Section 27.1) of electrons in the wire of 0.1 mm/s?

31. The Fermi energy of copper at 300 K is 7.05 eV. (a) What is the average energy of a conduction electron in copper at 300 K? (b) At what temperature would the average translational energy of a molecule in an ideal gas be equal to the energy calculated in part (a)?

32. W Consider a cube of gold 1.00 mm on an edge. Calculate the approximate number of conduction electrons in this cube whose energies lie in the range 4.000 to 4.025 eV.

33. M Sodium is a monovalent metal having a density of 0.971 g/cm^3 and a molar mass of 23.0 g/mol. Use this information to calculate (a) the density of charge carriers and (b) the Fermi energy of sodium.

34. *Why is the following situation impossible?* A hypothetical metal has the following properties: its Fermi energy is 5.48 eV, its density is 4.90×10^3 kg/m^3, its molar mass is 100 g/mol, and it has one free electron per atom.

35. For copper at 300 K, calculate the probability that a state with an energy equal to 99.0% of the Fermi energy is occupied.

36. For a metal at temperature T, calculate the probability that a state with an energy equal to βE_F is occupied where β is a fraction between 0 and 1.

37. **Review.** An electron moves in a three-dimensional box of edge length L and volume L^3. The wave function of the particle is $\psi = A \sin(k_x x) \sin(k_y y) \sin(k_z z)$. Show that its energy is given by Equation 43.20,

$$E = \frac{\hbar^2 \pi^2}{2m_e L^2}(n_x^2 + n_y^2 + n_z^2)$$

where the quantum numbers (n_x, n_y, n_z) are integers ≥ 1. *Suggestion:* The Schrödinger equation in three dimensions may be written

$$\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right) = (U - E)\psi$$

38. M (a) Consider a system of electrons confined to a three-dimensional box. Calculate the ratio of the number of allowed energy levels at 8.50 eV to the number at 7.05 eV. (b) **What If?** Copper has a Fermi energy of 7.05 eV at 300 K. Calculate the ratio of the number of occupied levels in copper at an energy of 8.50 eV to the number at the Fermi energy. (c) How does your answer to part (b) compare with that obtained in part (a)?

39. Show that the average kinetic energy of a conduction electron in a metal at 0 K is $E_{avg} = \frac{3}{5}E_F$. *Suggestion:* In general, the average kinetic energy is

$$E_{avg} = \frac{1}{n_e}\int_0^\infty EN(E)\,dE$$

where n_e is the density of particles, $N(E)\,dE$ is given by Equation 43.22, and the integral is over all possible values of the energy.

Section 43.6 Electrical Conduction in Metals, Insulators, and Semiconductors

40. Most solar radiation has a wavelength of 1 μm or less. (a) What energy gap should the material in a solar cell have if it is to absorb this radiation? (b) Is silicon an appropriate solar cell material (see Table 43.3)? Explain your answer.

41. The energy gap for silicon at 300 K is 1.14 eV. (a) Find the lowest-frequency photon that can promote an electron from the valence band to the conduction band. (b) What is the wavelength of this photon?

42. Light from a hydrogen discharge tube is incident on a CdS crystal. (a) Which spectral lines from the Balmer series are absorbed and (b) which are transmitted?

43. A light-emitting diode (LED) made of the semiconductor GaAsP emits red light ($\lambda = 650$ nm). Determine the energy-band gap E_g for this semiconductor.

44. The longest wavelength of radiation absorbed by a certain semiconductor is 0.512 μm. Calculate the energy gap for this semiconductor.

45. You are asked to build a scientific instrument that is thermally isolated from its surroundings. The isolation container may be a calorimeter, but these design criteria could apply to other containers as well. You wish to use a laser external to the container to raise the temperature of a target inside the instrument. You decide to use a diamond window in the container. Diamond has an energy gap of 5.47 eV. What is the shortest laser wavelength you can use to warm the sample inside the instrument?

46. GP **Review.** When a phosphorus atom is substituted for a silicon atom in a crystal, four of the phosphorus valence electrons form bonds with neighboring atoms and the remaining electron is much more loosely bound. You can model the electron as free to move through the crystal lattice. The phosphorus nucleus has one more positive charge than does the silicon nucleus, however, so the extra electron provided by the phosphorus atom is attracted to this single nuclear charge $+e$. The energy levels of the extra electron are similar to those of the electron in the Bohr hydrogen atom with two important exceptions. First, the Coulomb attraction between the electron and the positive charge on the phosphorus nucleus is reduced by a factor of $1/\kappa$ from what it would be in free space (see Eq. 26.21), where κ is the dielectric constant of the crystal. As a result, the orbit radii are greatly increased over those of the hydrogen atom. Second, the influence of the periodic electric potential of the lattice causes the electron to move as if it had an effective mass m^*, which is quite different from the mass m_e of a free electron. You can use the Bohr model of hydrogen to obtain relatively accurate values for the allowed energy levels of the extra electron. We wish to find the typical energy of these donor states, which play an important role in semiconductor devices. Assume $\kappa = 11.7$ for silicon and $m^* = 0.220m_e$. (a) Find a symbolic expression for the smallest radius of the electron orbit in terms of a_0, the Bohr radius. (b) Substitute numerical values to find the numerical value of the smallest radius. (c) Find a symbolic expression for the energy levels E_n' of the electron in the Bohr orbits around the donor atom in terms of m_e, m^*, κ, and E_n, the energy of the hydrogen atom in the Bohr model. (d) Find the numerical value of the energy for the ground state of the electron.

Section 43.7 Semiconductor Devices

47. Assuming $T = 300$ K, (a) for what value of the bias voltage ΔV in Equation 43.27 does $I = 9.00I_0$? (b) **What If?** What if $I = -0.900I_0$?

48. A diode, a resistor, and a battery are connected in a series circuit. The diode is at a temperature for which $k_BT = 25.0$ meV, and the saturation value of the current is $I_0 = 1.00\ \mu$A. The resistance of the resistor is $R = 745\ \Omega$, and the battery maintains a constant potential difference of $\boldsymbol{\mathcal{E}} = 2.42$ V between its terminals. (a) Use Kirchhoff's loop rule to show that

$$\boldsymbol{\mathcal{E}} - \Delta V = I_0R\left(e^{e\Delta V/k_BT} - 1\right)$$

where ΔV is the voltage across the diode. (b) To solve this transcendental equation for the voltage ΔV, graph the left-hand side of the above equation and the right-hand side as functions of ΔV and find the value of ΔV at which the curves cross. (c) Find the current I in the circuit. (d) Find the ohmic resistance of the diode, defined as the ratio $\Delta V/I$, at the voltage in part (b). (e) Find the dynamic resistance of the diode, which is defined as the derivative $d(\Delta V)/dI$, at the voltage in part (b).

49. W You put a diode in a microelectronic circuit to protect the system in case an untrained person installs the battery backward. In the correct forward-bias situation, the current is 200 mA with a potential difference of 100 mV across the diode at room temperature (300 K). If the battery were reversed, so that the potential difference across the diode is still 100 mV but with the opposite sign, what would be the magnitude of the current in the diode?

50. M A diode is at room temperature so that $k_BT = 0.025\,0$ eV. Taking the applied voltages across the diode to be $+1.00$ V (under forward bias) and -1.00 V (under reverse bias), calculate the ratio of the forward current to the reverse current if the diode is described by Equation 43.27.

Section 43.8 Superconductivity

Problem 30 in Chapter 30 and Problems 73 through 76 in Chapter 32 can also be assigned with this section.

51. A thin rod of superconducting material 2.50 cm long is placed into a 0.540-T magnetic field with its cylindrical axis along the magnetic field lines. (a) Sketch the directions of the applied field and the induced surface current. (b) Find the magnitude of the surface current on the curved surface of the rod.

52. A direct and relatively simple demonstration of zero DC resistance can be carried out using the four-point probe method. The probe shown in Figure P43.52 consists of a disk of $YBa_2Cu_3O_7$ (a high-T_c superconductor) to which four wires are attached. Current is maintained through the sample by applying a DC voltage between points a and b, and it is measured with a DC ammeter. The current can be varied with the variable resistance R. The potential difference ΔV_{cd} between c and d is measured with a digital voltmeter. When the probe is immersed in liquid nitrogen, the sample quickly cools

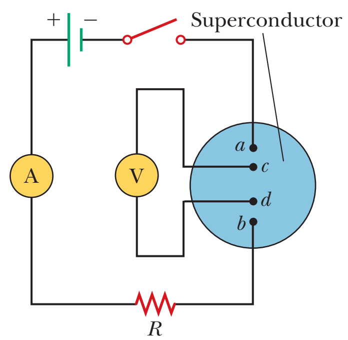


Figure P43.52

to 77 K, below the critical temperature of the material, 92 K. The current remains approximately constant, but ΔV_{cd} *drops abruptly to zero*. (a) Explain this observation on the basis of what you know about superconductors. (b) The data in the accompanying table represent actual values of ΔV_{cd} for different values of I taken on the sample at room temperature in the senior author's laboratory. A 6-V battery in series with a variable resistor R supplied the current. The values of R ranged from 10 Ω to 100 Ω. Make an I–ΔV plot of the data and determine whether the sample behaves in a linear manner. (c) From the data, obtain a value for the DC resistance of the sample at room temperature. (d) At room temperature, it was found that $\Delta V_{cd} = 2.234$ mV for $I = 100.3$ mA, but after the sample was cooled to 77 K, $\Delta V_{cd} = 0$ and $I = 98.1$ mA. What do you think might have caused the slight decrease in current?

Current Versus Potential Difference ΔV_{cd} Measured in a Bulk Ceramic Sample of $YBa_2Cu_3O_{7-\delta}$ at Room Temperature

I (mA)	ΔV_{cd} (mV)
57.8	1.356
61.5	1.441
68.3	1.602
76.8	1.802
87.5	2.053
102.2	2.398
123.7	2.904
155	3.61

53. A superconducting ring of niobium metal 2.00 cm in diameter is immersed in a uniform 0.020 0-T magnetic field directed perpendicular to the ring and carries no current. Determine the current generated in the ring when the magnetic field is suddenly decreased to zero. The inductance of the ring is 3.10×10^{-8} H.

Additional Problems

54. The effective spring constant associated with bonding in the N_2 molecule is 2 297 N/m. The nitrogen atoms each have a mass of 2.32×10^{-26} kg, and their nuclei are 0.120 nm apart. Assume the molecule is rigid. The first excited vibrational state of the molecule is above the vibrational ground state by an energy difference ΔE. Calculate the J value of the rotational state that is above the rotational ground state by the same energy difference ΔE.

55. The hydrogen molecule comes apart (dissociates) when it is excited internally by 4.48 eV. Assuming this molecule behaves like a harmonic oscillator having classical angular frequency $\omega = 8.28 \times 10^{14}$ rad/s, find the highest vibrational quantum number for a state below the 4.48-eV dissociation energy.

56. (a) Starting with Equation 43.17, show that the force exerted on an ion in an ionic solid can be written as

$$F = -\alpha k_e \frac{e^2}{r^2}\left[1 - \left(\frac{r_0}{r}\right)^{m-1}\right]$$

where α is the Madelung constant and r_0 is the equilibrium separation. (b) Imagine that an ion in the solid is displaced a small distance s from r_0. Show that the ion experiences a restoring force $F = -Ks$, where

$$K = \frac{\alpha k_e e^2}{r_0^3}(m-1)$$

(c) Use the result of part (b) to find the frequency of vibration of a Na^+ ion in NaCl. Take $m = 8$ and use the value $\alpha = 1.747\ 6$.

57. Under pressure, liquid helium can solidify as each atom bonds with four others, and each bond has an average energy of 1.74×10^{-23} J. Find the latent heat of fusion for helium in joules per gram. (The molar mass of He is 4.00 g/mol.)

58. The dissociation energy of ground-state molecular hydrogen is 4.48 eV, but it only takes 3.96 eV to dissociate it when it starts in the first excited vibrational state with $J = 0$. Using this information, determine the depth of the H_2 molecular potential-energy function.

59. Starting with Equation 43.17, show that the ionic cohesive energy of an ionically bonded solid is given by Equation 43.18.

60. The Fermi–Dirac distribution function can be written as

$$f(E) = \frac{1}{e^{(E-E_F)/k_BT} + 1} = \frac{1}{e^{(E/E_F - 1)T_F/T} + 1}$$

where T_F is the *Fermi temperature*, defined according to

$$k_B T_F \equiv E_F$$

(a) Write a spreadsheet to calculate and plot $f(E)$ versus E/E_F at a fixed temperature T. (b) Describe the curves obtained for $T = 0.1T_F$, $0.2T_F$, and $0.5T_F$.

61. A particle moves in one-dimensional motion through a field for which the potential energy of the particle–field system is

$$U(x) = \frac{A}{x^3} - \frac{B}{x}$$

where $A = 0.150$ eV · nm³ and $B = 3.68$ eV · nm. The shape of this function is shown in Figure P43.61. (a) Find the equilibrium position x_0 of the particle. (b) Determine the depth U_0 of this potential well.

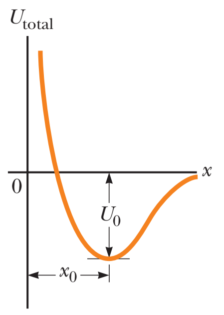


Figure P43.61 Problems 61 and 62.

(c) In moving along the x axis, what maximum force toward the negative x direction does the particle experience?

62. A particle of mass m moves in one-dimensional motion through a field for which the potential energy of the particle–field system is

$$U(x) = \frac{A}{x^3} - \frac{B}{x}$$

where A and B are constants. The general shape of this function is shown in Figure P43.61. (a) Find the equilibrium position x_0 of the particle in terms of m, A, and B. (b) Determine the depth U_0 of this potential well. (c) In moving along the x axis, what maximum force toward the negative x direction does the particle experience?

Challenge Problems

63. W As you will learn in Chapter 44, carbon-14 (^{14}C) is an unstable isotope of carbon. It has the same chemical properties and electronic structure as the much more abundant isotope carbon-12 (^{12}C), but it has different nuclear properties. Its mass is 14 u, greater than that of carbon-12 because of the two extra neutrons in the carbon-14 nucleus. Assume the CO molecular potential energy is the same for both isotopes of carbon and the examples in Section 43.2 contain accurate data and results for carbon monoxide with carbon-12 atoms. (a) What is the vibrational frequency of ^{14}CO? (b) What is the moment of inertia of ^{14}CO? (c) What wavelengths of light can be absorbed by ^{14}CO in the ($v = 0, J = 10$) state that cause it to end up in the $v = 1$ state?

64. As an alternative to Equation 43.1, another useful model for the potential energy of a diatomic molecule is the Morse potential

$$U(r) = B\left[e^{-a(r - r_0)} - 1\right]^2$$

where B, a, and r_0 are parameters used to adjust the shape of the potential and its depth. (a) What is the equilibrium separation of the nuclei? (b) What is the depth of the potential well, defined as the difference in energy between the potential's minimum value and its asymptote as r approaches infinity? (c) If μ is the reduced mass of the system of two nuclei and assuming the potential is nearly parabolic about the well minimum, what is the vibrational frequency of the diatomic molecule in its ground state? (d) What amount of energy needs to be supplied to the ground-state molecule to separate the two nuclei to infinity?

CHAPTER

44

Nuclear Structure

Ötzi the Iceman, a Copper Age man, was discovered by German tourists in the Italian Alps in 1991 when a glacier melted enough to expose his remains. Analysis of his corpse has exposed his last meal, illnesses he suffered, and places he lived. Radioactivity was used to determine that he lived in about 3300 BC. Ötzi can be seen today in the Südtiroler Archäologiemuseum (South Tyrol Museum of Archaeology) in Bolzano, Italy. *(©Vienna Report Agency/Sygma/Corbis)*

The year 1896 marks the birth of nuclear physics when French physicist Antoine-Henri Becquerel (1852–1908) discovered radioactivity in uranium compounds. This discovery prompted scientists to investigate the details of radioactivity and, ultimately, the structure of the nucleus. Pioneering work by Ernest Rutherford showed that the radiation emitted from radioactive substances is of three types—alpha, beta, and gamma rays—classified according to the nature of their electric charge and their ability to penetrate matter and ionize air. Later experiments showed that alpha rays are helium nuclei, beta rays are electrons, and gamma rays are high-energy photons.

In 1911, Rutherford, Hans Geiger, and Ernest Marsden performed the alpha-particle scattering experiments described in Section 42.2. These experiments established that the nucleus of an atom can be modeled as a point mass and point charge and that most of the atomic mass is contained in the nucleus. Subsequent studies revealed the presence of a new type of force, the short-range nuclear force, which is predominant at particle separation distances less than approximately 10^{-14} m and is zero for large distances.

In this chapter, we discuss the properties and structure of the atomic nucleus. We start by describing the basic properties of nuclei, followed by a discussion of nuclear forces and binding energy, nuclear models, and the phenomenon of radioactivity. Finally, we explore the various processes by which nuclei decay and the ways that nuclei can react with each other.

44.1 Some Properties of Nuclei

All nuclei are composed of two types of particles: protons and neutrons. The only exception is the ordinary hydrogen nucleus, which is a single proton. We describe the atomic nucleus by the number of protons and neutrons it contains, using the following quantities:

- the **atomic number** Z, which equals the number of protons in the nucleus (sometimes called the *charge number*)
- the **neutron number** N, which equals the number of neutrons in the nucleus
- the **mass number** $A = Z + N$, which equals the number of **nucleons** (neutrons plus protons) in the nucleus

A **nuclide** is a specific combination of atomic number and mass number that represents a nucleus. In representing nuclides, it is convenient to use the symbol ${}^{A}_{Z}\mathrm{X}$ to convey the numbers of protons and neutrons, where X represents the chemical symbol of the element. For example, ${}^{56}_{26}\mathrm{Fe}$ (iron) has mass number 56 and atomic number 26; therefore, it contains 26 protons and 30 neutrons. When no confusion is likely to arise, we omit the subscript Z because the chemical symbol can always be used to determine Z. Therefore, ${}^{56}_{26}\mathrm{Fe}$ is the same as ${}^{56}\mathrm{Fe}$ and can also be expressed as "iron-56" or "Fe-56."

The nuclei of all atoms of a particular element contain the same number of protons but often contain different numbers of neutrons. Nuclei related in this way are called **isotopes.** The isotopes of an element have the same Z value but different N and A values.

The natural abundance of isotopes can differ substantially. For example ${}^{11}_{6}\mathrm{C}$, ${}^{12}_{6}\mathrm{C}$, ${}^{13}_{6}\mathrm{C}$, and ${}^{14}_{6}\mathrm{C}$ are four isotopes of carbon. The natural abundance of the ${}^{12}_{6}\mathrm{C}$ isotope is approximately 98.9%, whereas that of the ${}^{13}_{6}\mathrm{C}$ isotope is only about 1.1%. Some isotopes, such as ${}^{11}_{6}\mathrm{C}$ and ${}^{14}_{6}\mathrm{C}$, do not occur naturally but can be produced by nuclear reactions in the laboratory or by cosmic rays.

Even the simplest element, hydrogen, has isotopes: ${}^{1}_{1}\mathrm{H}$, the ordinary hydrogen nucleus; ${}^{2}_{1}\mathrm{H}$, deuterium; and ${}^{3}_{1}\mathrm{H}$, tritium.

Pitfall Prevention 44.1

Mass Number Is Not Atomic Mass The mass number A should not be confused with the atomic mass. Mass number is an integer specific to an isotope and has no units; it is simply a count of the number of nucleons. Atomic mass has units and is generally not an integer because it is an average of the masses of a given element's naturally occurring isotopes.

Quick Quiz 44.1 For each part of this Quick Quiz, choose from the following answers: (a) protons (b) neutrons (c) nucleons. **(i)** The three nuclei ${}^{12}\mathrm{C}$, ${}^{13}\mathrm{N}$, and ${}^{14}\mathrm{O}$ have the same number of what type of particle? **(ii)** The three nuclei ${}^{12}\mathrm{N}$, ${}^{13}\mathrm{N}$, and ${}^{14}\mathrm{N}$ have the same number of what type of particle? **(iii)** The three nuclei ${}^{14}\mathrm{C}$, ${}^{14}\mathrm{N}$, and ${}^{14}\mathrm{O}$ have the same number of what type of particle?

Charge and Mass

The proton carries a single positive charge e, equal in magnitude to the charge $-e$ on the electron ($e = 1.6 \times 10^{-19}$ C). The neutron is electrically neutral as its name implies. Because the neutron has no charge, it was difficult to detect with early experimental apparatus and techniques. Today, neutrons are easily detected with devices such as plastic scintillators.

Nuclear masses can be measured with great precision using a mass spectrometer (see Section 29.3) and by the analysis of nuclear reactions. The proton is approximately 1 836 times as massive as the electron, and the masses of the proton and the neutron are almost equal. The **atomic mass unit** u is defined in such a way that the mass of one atom of the isotope ${}^{12}\mathrm{C}$ is exactly 12 u, where 1 u is equal to $1.660\,539 \times 10^{-27}$ kg. According to this definition, the proton and neutron each have a mass of approximately 1 u and the electron has a mass that is only a small fraction of this value. The masses of these particles and others important to the phenomena discussed in this chapter are given in Table 44.1 (page 1382).

Table 44.1 Masses of Selected Particles in Various Units

Particle	Mass kg	u	MeV/c^2
Proton	$1.672\ 62 \times 10^{-27}$	1.007 276	938.27
Neutron	$1.674\ 93 \times 10^{-27}$	1.008 665	939.57
Electron (β particle)	$9.109\ 38 \times 10^{-31}$	$5.485\ 79 \times 10^{-4}$	0.510 999
${}^1_1$H atom	$1.673\ 53 \times 10^{-27}$	1.007 825	938.783
${}^4_2$He nucleus (α particle)	$6.644\ 66 \times 10^{-27}$	4.001 506	3 727.38
${}^4_2$He atom	$6.646\ 48 \times 10^{-27}$	4.002 603	3 728.40
${}^{12}_6$C atom	$1.992\ 65 \times 10^{-27}$	12.000 000	11 177.9

You might wonder how six protons and six neutrons, each having a mass larger than 1 u, can be combined with six electrons to form a carbon-12 atom having a mass of exactly 12 u. The bound system of ^{12}C has a lower rest energy (Section 39.8) than that of six separate protons and six separate neutrons. According to Equation 39.24, $E_R = mc^2$, this lower rest energy corresponds to a smaller mass for the bound system. The difference in mass accounts for the binding energy when the particles are combined to form the nucleus. We shall discuss this point in more detail in Section 44.2.

It is often convenient to express the atomic mass unit in terms of its *rest-energy equivalent*. For one atomic mass unit,

$$E_R = mc^2 = (1.660\ 539 \times 10^{-27}\ \text{kg})(2.997\ 92 \times 10^8\ \text{m/s})^2 = 931.494\ \text{MeV}$$

where we have used the conversion $1\ \text{eV} = 1.602\ 176 \times 10^{-19}$ J.

Based on the rest-energy expression in Equation 39.24, nuclear physicists often express mass in terms of the unit MeV/c^2.

Example 44.1 The Atomic Mass Unit

Use Avogadro's number to show that $1\ \text{u} = 1.66 \times 10^{-27}$ kg.

SOLUTION

Conceptualize From the definition of the mole given in Section 19.5, we know that exactly 12 g (= 1 mol) of ^{12}C contains Avogadro's number of atoms.

Categorize We evaluate the atomic mass unit that was introduced in this section, so we categorize this example as a substitution problem.

Find the mass m of one ^{12}C atom:

$$m = \frac{0.012\ \text{kg}}{6.02 \times 10^{23}\ \text{atoms}} = 1.99 \times 10^{-26}\ \text{kg}$$

Because one atom of ^{12}C is defined to have a mass of 12.0 u, divide by 12.0 to find the mass equivalent to 1 u:

$$1\ \text{u} = \frac{1.99 \times 10^{-26}\ \text{kg}}{12.0} = 1.66 \times 10^{-27}\ \text{kg}$$

The Size and Structure of Nuclei

In Rutherford's scattering experiments, positively charged nuclei of helium atoms (alpha particles) were directed at a thin piece of metallic foil. As the alpha particles moved through the foil, they often passed near a metal nucleus. Because of the positive charge on both the incident particles and the nuclei, the particles were deflected from their straight-line paths by the Coulomb repulsive force.

Rutherford used the isolated system (energy) analysis model to find an expression for the separation distance d at which an alpha particle approaching a nucleus head-on is turned around by Coulomb repulsion. In such a head-on collision, the mechanical energy of the nucleus–alpha particle system is conserved. The initial kinetic energy of the incoming particle is transformed completely to electric potential energy of the system when the alpha particle stops momentarily at the point of closest approach (the final configuration of the system) before moving back along the same path (Fig. 44.1). Applying Equation 8.2, the conservation of energy principle, to the system gives

$$\Delta K + \Delta U = 0$$

$$\left(0 - \tfrac{1}{2}mv^2\right) + \left(k_e \frac{q_1 q_2}{d} - 0\right) = 0$$

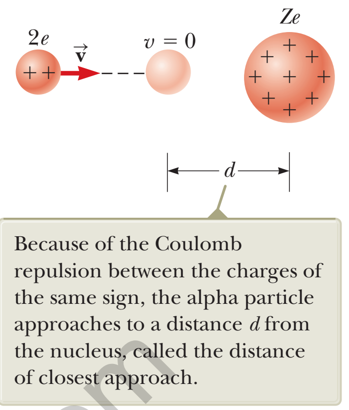


Figure 44.1 An alpha particle on a head-on collision course with a nucleus of charge Ze.

where m is the mass of the alpha particle and v is its initial speed. Solving for d gives

$$d = 2k_e \frac{q_1 q_2}{mv^2} = 2k_e \frac{(2e)(Ze)}{mv^2} = 4k_e \frac{Ze^2}{mv^2}$$

where Z is the atomic number of the target nucleus. From this expression, Rutherford found that the alpha particles approached nuclei to within 3.2×10^{-14} m when the foil was made of gold. Therefore, the radius of the gold nucleus must be less than this value. From the results of his scattering experiments, Rutherford concluded that the positive charge in an atom is concentrated in a small sphere, which he called the nucleus, whose radius is no greater than approximately 10^{-14} m.

Because such small lengths are common in nuclear physics, an often-used convenient length unit is the femtometer (fm), which is sometimes called the **fermi** and is defined as

$$1 \text{ fm} \equiv 10^{-15} \text{ m}$$

In the early 1920s, it was known that the nucleus of an atom contains Z protons and has a mass nearly equivalent to that of A protons, where on average $A \approx 2Z$ for lighter nuclei ($Z \leq 20$) and $A > 2Z$ for heavier nuclei. To account for the nuclear mass, Rutherford proposed that each nucleus must also contain $A - Z$ neutral particles that he called neutrons. In 1932, British physicist James Chadwick (1891–1974) discovered the neutron, and he was awarded the Nobel Prize in Physics in 1935 for this important work.

Since the time of Rutherford's scattering experiments, a multitude of other experiments have shown that most nuclei are approximately spherical and have an average radius given by

$$r = aA^{1/3} \quad (44.1)$$

◀ **Nuclear radius**

where a is a constant equal to 1.2×10^{-15} m and A is the mass number. Because the volume of a sphere is proportional to the cube of its radius, it follows from Equation 44.1 that the volume of a nucleus (assumed to be spherical) is directly proportional to A, the total number of nucleons. This proportionality suggests that *all nuclei have nearly the same density*. When nucleons combine to form a nucleus, they combine as though they were tightly packed spheres (Fig. 44.2). This fact has led to an analogy between the nucleus and a drop of liquid, in which the density of the drop is independent of its size. We shall discuss the liquid-drop model of the nucleus in Section 44.3.

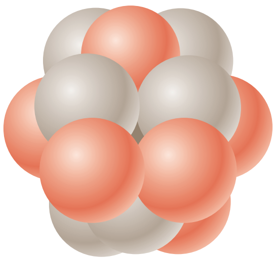


Figure 44.2 A nucleus can be modeled as a cluster of tightly packed spheres, where each sphere is a nucleon.

Example 44.2 The Volume and Density of a Nucleus

Consider a nucleus of mass number A.

(A) Find an approximate expression for the mass of the nucleus.

continued

44.2 continued

SOLUTION

Conceptualize Imagine the nucleus to be a collection of protons and neutrons as shown in Figure 44.2. The mass number A counts *both* protons and neutrons.

Categorize Let's assume A is large enough that we can imagine the nucleus to be spherical.

Analyze The mass of the proton is approximately equal to that of the neutron. Therefore, if the mass of one of these particles is m, the mass of the nucleus is approximately Am.

(B) Find an expression for the volume of this nucleus in terms of A.

SOLUTION

Assume the nucleus is spherical and use Equation 44.1:

$$(1)\quad V_{\text{nucleus}} = \tfrac{4}{3}\pi r^3 = \tfrac{4}{3}\pi a^3 A$$

(C) Find a numerical value for the density of this nucleus.

SOLUTION

Use Equation 1.1 and substitute Equation (1):

$$\rho = \frac{m_{\text{nucleus}}}{V_{\text{nucleus}}} = \frac{Am}{\frac{4}{3}\pi a^3 A} = \frac{3m}{4\pi a^3}$$

Substitute numerical values:

$$\rho = \frac{3(1.67 \times 10^{-27}\text{ kg})}{4\pi(1.2 \times 10^{-15}\text{ m})^3} = 2.3 \times 10^{17}\text{ kg/m}^3$$

Finalize The nuclear density is approximately 2.3×10^{14} times the density of water ($\rho_{\text{water}} = 1.0 \times 10^3\text{ kg/m}^3$).

WHAT IF? What if the Earth could be compressed until it had this density? How large would it be?

Answer Because this density is so large, we predict that an Earth of this density would be very small.

Use Equation 1.1 and the mass of the Earth to find the volume of the compressed Earth:

$$V = \frac{M_E}{\rho} = \frac{5.97 \times 10^{24}\text{ kg}}{2.3 \times 10^{17}\text{ kg/m}^3} = 2.6 \times 10^7\text{ m}^3$$

From this volume, find the radius:

$$V = \tfrac{4}{3}\pi r^3 \;\rightarrow\; r = \left(\frac{3V}{4\pi}\right)^{1/3} = \left[\frac{3(2.6 \times 10^7\text{ m}^3)}{4\pi}\right]^{1/3}$$

$$r = 1.8 \times 10^2\text{ m}$$

An Earth of this radius is indeed a small Earth!

Nuclear Stability

You might expect that the very large repulsive Coulomb forces between the close-packed protons in a nucleus should cause the nucleus to fly apart. Because that does not happen, there must be a counteracting attractive force. The **nuclear force** is a very short range (about 2 fm) attractive force that acts between all nuclear particles. The protons attract each other by means of the nuclear force, and, at the same time, they repel each other through the Coulomb force. The nuclear force also acts between pairs of neutrons and between neutrons and protons. The nuclear force dominates the Coulomb repulsive force within the nucleus (at short ranges), so stable nuclei can exist.

The nuclear force is independent of charge. In other words, the forces associated with the proton–proton, proton–neutron, and neutron–neutron interactions