Chapter 29

Images by Reflection

Infants are always fascinated by the antics of the Baby in the Mirror. Now if you want to know something about mirror images that most people don’t understand, try this. First bring this page closer and closer to your eyes, until you can no longer focus on it without straining. Then go in the bathroom and see how close you can get your face to the surface of the mirror before you can no longer easily focus on the image of your own eyes. You will find that the shortest comfortable eye-mirror distance is much less than the shortest comfortable eye-paper distance. This demonstrates that the image of your face in the mirror acts as if it had depth and
existed in the space behind the mirror. If the image was like a flat picture in a book, then you wouldn’t be able to focus on it from such a short distance.

In this chapter we will study the images formed by flat and curved mirrors on a qualitative, conceptual basis. Although this type of image is not as commonly encountered in everyday life as images formed by lenses, images formed by reflection are simpler to understand, so we discuss them first. In chapter 30 we will turn to a more mathematical treatment of images made by reflection. Surprisingly, the same equations can also be applied to lenses, which are the topic of chapter 31.

29.1 A virtual image

We can understand a mirror image using a ray diagram. Figure a shows several light rays, 1, that originated by diffuse reflection at the person’s nose. They bounce off the mirror, producing new rays, 2. To anyone whose eye is in the right position to get one of these rays, they appear to have come from a behind the mirror, 3, where they would have originated from a single point. This point is where the tip of the image-person’s nose appears to be. A similar analysis applies to every other point on the person’s face, so it looks as though there was an entire face behind the mirror. The customary way of describing the situation requires some explanation:

**Customary description in physics:** There is an image of the face behind the mirror.

**Translation:** The pattern of rays coming from the mirror is exactly the same as it would be if there were a face behind the mirror. Nothing is really behind the mirror.

This is referred to as a virtual image, because the rays do not actually cross at the point behind the mirror. They only appear to have originated there.

**self-check A**

Imagine that the person in figure a moves his face down quite a bit — a couple of feet in real life, or a few inches on this scale drawing. The mirror stays where it is. Draw a new ray diagram. Will there still be an image? If so, where is it visible from? Answer, p. 1041

The geometry of specular reflection tells us that rays 1 and 2 are at equal angles to the normal (the imaginary perpendicular line piercing the mirror at the point of reflection). This means that ray 2’s imaginary continuation, 3, forms the same angle with the mirror as ray 1. Since each ray of type 3 forms the same angles with
the mirror as its partner of type 1, we see that the distance of the image from the mirror is the same as that of the actual face from the mirror, and it lies directly across from it. The image therefore appears to be the same size as the actual face.

**An eye exam**

Example 1

Figure b shows a typical setup in an optometrist’s examination room. The patient’s vision is supposed to be tested at a distance of 6 meters (20 feet in the U.S.), but this distance is larger than the amount of space available in the room. Therefore a mirror is used to create an image of the eye chart behind the wall.

**The Praxinoscope**

Example 2

Figure c shows an old-fashioned device called a praxinoscope, which displays an animated picture when spun. The removable strip of paper with the pictures printed on it has twice the radius of the inner circle made of flat mirrors, so each picture’s virtual image is at the center. As the wheel spins, each picture’s image is replaced by the next.
Discussion question
A The figure shows an object that is off to one side of a mirror. Draw a ray diagram. Is an image formed? If so, where is it, and from which directions would it be visible?
29.2 Curved mirrors

An image in a flat mirror is a pretechnological example: even animals can look at their reflections in a calm pond. We now pass to our first nontrivial example of the manipulation of an image by technology: an image in a curved mirror. Before we dive in, let’s consider why this is an important example. If it was just a question of memorizing a bunch of facts about curved mirrors, then you would rightly rebel against an effort to spoil the beauty of your liberally educated brain by force-feeding you technological trivia. The reason this is an important example is not that curved mirrors are so important in and of themselves, but that the results we derive for curved bowl-shaped mirrors turn out to be true for a large class of other optical devices, including mirrors that bulge outward rather than inward, and lenses as well. A microscope or a telescope is simply a combination of lenses or mirrors or both. What you’re really learning about here is the basic building block of all optical devices from movie projectors to octopus eyes.

Because the mirror in figure d is curved, it bends the rays back closer together than a flat mirror would: we describe it as converging. Note that the term refers to what it does to the light rays, not to the physical shape of the mirror’s surface. (The surface itself would be described as concave. The term is not all that hard to remember, because the hollowed-out interior of the mirror is like a cave.) It is surprising but true that all the rays like 3 really do converge on a point, forming a good image. We will not prove this fact, but it is true for any mirror whose curvature is gentle enough and that is symmetric with respect to rotation about the perpendicular line passing through its center (not asymmetric like a potato chip). The old-fashioned method of making mirrors and lenses is by grinding them in grit by hand, and this automatically tends to produce an almost perfect spherical surface.

Bending a ray like 2 inward implies bending its imaginary continuation 3 outward, in the same way that raising one end of a seesaw causes the other end to go down. The image therefore forms deeper behind the mirror. This doesn’t just show that there is extra distance between the image-nose and the mirror; it also implies that the image itself is bigger from front to back. It has been magnified in the front-to-back direction.

It is easy to prove that the same magnification also applies to the image’s other dimensions. Consider a point like E in figure e. The trick is that out of all the rays diffusely reflected by E, we pick the one that happens to head for the mirror’s center, C. The equal-angle property of specular reflection plus a little straightforward geometry easily leads us to the conclusion that triangles ABC and CDE are the same shape, with ABC being simply a scaled-up version of CDE. The magnification of depth equals the ratio BC/CD, and the up-
down magnification is $AB/DE$. A repetition of the same proof shows that the magnification in the third dimension (out of the page) is also the same. This means that the image-head is simply a larger version of the real one, without any distortion. The scaling factor is called the magnification, $M$. The image in the figure is magnified by a factor $M = 1.9$.

Note that we did not explicitly specify whether the mirror was a sphere, a paraboloid, or some other shape. However, we assumed that a focused image would be formed, which would not necessarily be true, for instance, for a mirror that was asymmetric or very deeply curved.

### 29.3 A real image

If we start by placing an object very close to the mirror, $g/1$, and then move it farther and farther away, the image at first behaves as we would expect from our everyday experience with flat mirrors, receding deeper and deeper behind the mirror. At a certain point, however, a dramatic change occurs. When the object is more than a certain distance from the mirror, $g/2$, the image appears upside-down and in *front* of the mirror.

Here’s what’s happened. The mirror bends light rays inward, but
when the object is very close to it, as in g/1, the rays coming from a
given point on the object are too strongly diverging (spreading) for
the mirror to bring them back together. On reflection, the rays are
still diverging, just not as strongly diverging. But when the object
is sufficiently far away, g/2, the mirror is only intercepting the rays
that came out in a narrow cone, and it is able to bend these enough
so that they will reconverge.

Note that the rays shown in the figure, which both originated at
the same point on the object, reunite when they cross. The point
where they cross is the image of the point on the original object.
This type of image is called a real image, in contradistinction to the
virtual images we’ve studied before.

Definition: A real image is one where rays actually cross. A virtual
image is a point from which rays only appear to have come.

The use of the word “real” is perhaps unfortunate. It sounds as
though we are saying the image was an actual material object, which
of course it is not.

The distinction between a real image and a virtual image is an im-
portant one, because a real image can be projected onto a screen
or photographic film. If a piece of paper is inserted in figure g/2
at the location of the image, the image will be visible on the paper
(provided the object is bright and the room is dark). Your eye uses
a lens to make a real image on the retina.

self-check B
Sketch another copy of the face in figure g/1, even farther from the
mirror, and draw a ray diagram. What has happened to the location of
the image?

Answer, p. 1041

29.4 Images of images

If you are wearing glasses right now, then the light rays from the
page are being manipulated first by your glasses and then by the lens
of your eye. You might think that it would be extremely difficult
to analyze this, but in fact it is quite easy. In any series of optical
elements (mirrors or lenses or both), each element works on the rays
furnished by the previous element in exactly the same manner as if
the image formed by the previous element was an actual object.

Figure h shows an example involving only mirrors. The Newtonian
telescope, invented by Isaac Newton, consists of a large curved mir-
ror, plus a second, flat mirror that brings the light out of the tube.
(In very large telescopes, there may be enough room to put a camera
or even a person inside the tube, in which case the second mirror is
not needed.) The tube of the telescope is not vital; it is mainly a
A Newtonian telescope being used for visual rather than photographic observing. In real life, an eyepiece lens is normally used for additional magnification, but this simpler setup will also work.

The large curved mirror by itself would form an image I, but the small flat mirror creates an image of the image, I'. The relationship between I and I' is exactly the same as it would be if I was an actual object rather than an image: I and I' are at equal distances from the plane of the mirror, and the line between them is perpendicular to the plane of the mirror.

One surprising wrinkle is that whereas a flat mirror used by itself forms a virtual image of an object that is real, here the mirror is forming a real image of virtual image I. This shows how pointless it would be to try to memorize lists of facts about what kinds of images are formed by various optical elements under various circumstances. You are better off simply drawing a ray diagram.

The angular size of the flower depends on its distance from the eye.

Although the main point here was to give an example of an image of an image, figure i also shows an interesting case where we need to make the distinction between magnification and angular magnification. If you are looking at the moon through this telescope, then the images I and I' are much smaller than the actual moon. Otherwise, for example, image I would not fit inside the telescope! However, these images are very close to your eye compared to the actual moon. The small size of the image has been more than compensated for by the shorter distance. The important thing here is the amount of angle within your field of view that the image covers, and it is this angle that has been increased. The factor by which it is increased is called the angular magnification, $M_a$. 
k / The person uses a mirror to get a view of both sides of the ladybug. Although the flat mirror has $M = 1$, it doesn't give an angular magnification of 1. The image is farther from the eye than the object, so the angular magnification $M_a = \alpha_i/\alpha_o$ is less than one.

Discussion questions

A  Locate the images of you that will be formed if you stand between two parallel mirrors.
B Locate the images formed by two perpendicular mirrors, as in the figure. What happens if the mirrors are not perfectly perpendicular?
C Locate the images formed by the periscope.
Summary

Selected vocabulary

real image . . . . a place where an object appears to be, because the rays diffusely reflected from any given point on the object have been bent so that they come back together and then spread out again from the new point.

virtual image . . like a real image, but the rays don’t actually cross again; they only appear to have come from the point on the image.

converging . . . . describes an optical device that brings light rays closer to the optical axis.

diverging . . . . bends light rays farther from the optical axis.

magnification . . the factor by which an image’s linear size is increased (or decreased).

angular magnification . . the factor by which an image’s apparent angular size is increased (or decreased).

concave . . . . describes a surface that is hollowed out like a cave.

convex . . . . describes a surface that bulges outward.

Notation

\[ M \] . . . . . . . . the magnification of an image.

\[ M_a \] . . . . . . . . the angular magnification of an image.

Summary

A large class of optical devices, including lenses and flat and curved mirrors, operates by bending light rays to form an image. A real image is one for which the rays actually cross at each point of the image. A virtual image, such as the one formed behind a flat mirror, is one for which the rays only appear to have crossed at a point on the image. A real image can be projected onto a screen; a virtual one cannot.

Mirrors and lenses will generally make an image that is either smaller than or larger than the original object. The scaling factor is called the magnification. In many situations, the angular magnification is more important than the actual magnification.
Problems

Key
✓ A computerized answer check is available online.
∫ A problem that requires calculus.
★ A difficult problem.

1 A man is walking at 1.0 m/s directly towards a flat mirror. At what speed is his separation from his image decreasing? ✓

2 If a mirror on a wall is only big enough for you to see yourself from your head down to your waist, can you see your entire body by backing up? Test this experimentally and come up with an explanation for your observations, including a ray diagram.

Note that when you do the experiment, it’s easy to confuse yourself if the mirror is even a tiny bit off of vertical. One way to check yourself is to artificially lower the top of the mirror by putting a piece of tape or a post-it note where it blocks your view of the top of your head. You can then check whether you are able to see more of yourself both above and below by backing up.

3 In this chapter we’ve only done examples of mirrors with hollowed-out shapes (called concave mirrors). Now draw a ray diagram for a curved mirror that has a bulging outward shape (called a convex mirror). (a) How does the image’s distance from the mirror compare with the actual object’s distance from the mirror? From this comparison, determine whether the magnification is greater than or less than one. (b) Is the image real, or virtual? Could this mirror ever make the other type of image?

4 As discussed in question 3, there are two types of curved mirrors, concave and convex. Make a list of all the possible combinations of types of images (virtual or real) with types of mirrors (concave and convex). (Not all of the four combinations are physically possible.) Now for each one, use ray diagrams to determine whether increasing the distance of the object from the mirror leads to an increase or a decrease in the distance of the image from the mirror.

Draw BIG ray diagrams! Each diagram should use up about half a page of paper.

Some tips: To draw a ray diagram, you need two rays. For one of these, pick the ray that comes straight along the mirror’s axis, since its reflection is easy to draw. After you draw the two rays and locate the image for the original object position, pick a new object position that results in the same type of image, and start a new ray diagram, in a different color of pen, right on top of the first one. For the two new rays, pick the ones that just happen to hit the mirror at the same two places; this makes it much easier to get the result right without depending on extreme accuracy in your ability to draw the
5 If the user of an astronomical telescope moves her head closer to or farther away from the image she is looking at, does the magnification change? Does the angular magnification change? Explain. (For simplicity, assume that no eyepiece is being used.)

Solution, p. 1032

6 In figure g/2 in on page 856, only the image of my forehead was located by drawing rays. Either photocopy the figure or download the book and print out the relevant page. On this copy of the figure, make a new set of rays coming from my chin, and locate its image. To make it easier to judge the angles accurately, draw rays from the chin that happen to hit the mirror at the same points where the two rays from the forehead were shown hitting it. By comparing the locations of the chin’s image and the forehead’s image, verify that the image is actually upside-down, as shown in the original figure.

7 The figure shows four points where rays cross. Of these, which are image points? Explain.

8 Here’s a game my kids like to play. I sit next to a sunny window, and the sun reflects from the glass on my watch, making a disk of light on the wall or floor, which they pretend to chase as I move it around. Is the spot a disk because that’s the shape of the sun, or because it’s the shape of my watch? In other words, would a square watch make a square spot, or do we just have a circular image of the circular sun, which will be circular no matter what?

9 Suppose we have a polygonal room whose walls are mirrors, and there a pointlike light source in the room. In most such examples, every point in the room ends up being illuminated by the light source after some finite number of reflections. A difficult mathematical question, first posed in the middle of the last century, is whether it is ever possible to have an example in which the whole room is not illuminated. (Rays are assumed to be absorbed if they strike exactly at a vertex of the polygon, or if they pass exactly through the plane of a mirror.)

The problem was finally solved in 1995 by G.W. Tokarsky, who found an example of a room that was not illuminable from a certain point. Figure 9 shows a slightly simpler example found two years later by D. Castro. If a light source is placed at either of the locations shown with dots, the other dot remains unilluminated, although every other point is lit up. It is not straightforward to prove rigorously that Castro’s solution has this property. However, the plausibility of the solution can be demonstrated as follows.

Suppose the light source is placed at the right-hand dot. Locate all the images formed by single reflections. Note that they form a
regular pattern. Convince yourself that none of these images illuminates the left-hand dot. Because of the regular pattern, it becomes plausible that even if we form images of images, images of images of images, etc., none of them will ever illuminate the other dot.

There are various other versions of the problem, some of which remain unsolved. The book by Klee and Wagon gives a good introduction to the topic, although it predates Tokarsky and Castro’s work.

References:
Exercise 29: Exploring images with a curved mirror

Equipment:

- concave mirrors with deep curvature
- concave mirrors with gentle curvature
- convex mirrors

1. Obtain a curved mirror from your instructor. If it is silvered on both sides, make sure you’re working with the concave side, which bends light rays inward. Look at your own face in the mirror. Now change the distance between your face and the mirror, and see what happens. Explore the full range of possible distances between your face and the mirror.

In these observations you’ve been changing two variables at once: the distance between the object (your face) and the mirror, and the distance from the mirror to your eye. In general, scientific experiments become easier to interpret if we practice isolation of variables, i.e., only change one variable while keeping all the others constant. In parts 2 and 3 you’ll form an image of an object that’s not your face, so that you can have independent control of the object distance and the point of view.

2. With the mirror held far away from you, observe the image of something behind you, over your shoulder. Now bring your eye closer and closer to the mirror. Can you see the image with your eye very close to the mirror? See if you can explain your observation by drawing a ray diagram.
3. Now imagine the following new situation, but *don’t actually do it yet*. Suppose you lay the mirror face-up on a piece of tissue paper, put your finger a few cm above the mirror, and look at the image of your finger. As in part 2, you can bring your eye closer and closer to the mirror. Will you be able to see the image with your eye very close to the mirror? Draw a ray diagram to help you predict what you will observe.

Prediction:__________________

Now test your prediction. If your prediction was incorrect, see if you can figure out what went wrong, or ask your instructor for help.

4. For parts 4 and 5, it’s more convenient to use concave mirrors that are more gently curved; obtain one from your instructor. Lay the mirror on the tissue paper, and use it to create an image of the overhead lights on a piece of paper above it and a little off to the side. What do you have to do in order to make the image clear? Can you explain this observation using a ray diagram?

———> turn page
5. Now imagine the following experiment, but *don’t do it yet*. What will happen to the image on the paper if you cover half of the mirror with your hand?

   Prediction:__________________

   Test your prediction. If your prediction was incorrect, can you explain what happened?

6. Now imagine forming an image with a convex mirror (one that bulges outward), and that therefore bends light rays away from the central axis (i.e., is diverging). Draw a typical ray diagram.

   Is the image real, or virtual? Will there be more than one type of image?

   Prediction:__________________

   Test your prediction.
Chapter 30
Images, Quantitatively

It sounds a bit odd when a scientist refers to a theory as “beautiful,” but to those in the know it makes perfect sense. One mark of a beautiful theory is that it surprises us by being simple. The mathematical theory of lenses and curved mirrors gives us just such a surprise. We expect the subject to be complex because there are so many cases: a converging mirror forming a real image, a diverging lens that makes a virtual image, and so on for a total of six possibilities. If we want to predict the location of the images in all these situations, we might expect to need six different equations, and six more for predicting magnifications. Instead, it turns out that we can use just one equation for the location of the image and one equation for its magnification, and these two equations work in all the different cases with no changes except for plus and minus signs. This is the kind of thing the physicist Eugene Wigner referred to as “the unreasonable effectiveness of mathematics.” Sometimes
we can find a deeper reason for this kind of unexpected simplicity, but sometimes it almost seems as if God went out of Her way to make the secrets of universe susceptible to attack by the human thought-tool called math.

30.1 A real image formed by a converging mirror

Location of the image

We will now derive the equation for the location of a real image formed by a converging mirror. We assume for simplicity that the mirror is spherical, but actually this isn’t a restrictive assumption, because any shallow, symmetric curve can be approximated by a sphere. The shape of the mirror can be specified by giving the location of its center, C. A deeply curved mirror is a sphere with a small radius, so C is close to it, while a weakly curved mirror has C farther away. Given the point O where the object is, we wish to find the point I where the image will be formed.

To locate an image, we need to track a minimum of two rays coming from the same point. Since we have proved in the previous chapter that this type of image is not distorted, we can use an on-axis point, O, on the object, as in figure a/1. The results we derive will also hold for off-axis points, since otherwise the image would have to be distorted, which we know is not true. We let one of the rays be the one that is emitted along the axis; this ray is especially easy to trace, because it bounces straight back along the axis again. As our second ray, we choose one that strikes the mirror at a distance of 1 from the axis. “One what?” asks the astute reader. The answer is that it doesn’t really matter. When a mirror has shallow curvature, all the reflected rays hit the same point, so 1 could be expressed in any units you like. It could, for instance, be 1 cm, unless your mirror is smaller than 1 cm!

The only way to find out anything mathematical about the rays is to use the sole mathematical fact we possess concerning specular reflection: the incident and reflected rays form equal angles with respect to the normal, which is shown as a dashed line. Therefore the two angles shown in figure a/2 are the same, and skipping some straightforward geometry, this leads to the visually reasonable result that the two angles in figure a/3 are related as follows:

$$\theta_i + \theta_o = \text{constant}$$

(Note that $\theta_i$, $\theta_o$, which are measured from the image and the object, not from the eye like the angles we referred to in discussing angular magnification on page 858.) For example, move O farther from the mirror. The top angle in figure a/2 is increased, so the bottom angle must increase by the same amount, causing the image...
point, I, to move closer to the mirror. In terms of the angles shown in figure a/3, the more distant object has resulted in a smaller angle $\theta_0$, while the closer image corresponds to a larger $\theta_i$. One angle increases by the same amount that the other decreases, so their sum remains constant. These changes are summarized in figure a/4.

The sum $\theta_i + \theta_o$ is a constant. What does this constant represent? Geometrically, we interpret it as double the angle made by the dashed radius line. Optically, it is a measure of the strength of the mirror, i.e., how strongly the mirror focuses light, and so we call it the focal angle, $\theta_f$.

$\theta_i + \theta_o = \theta_f$.

Suppose, for example, that we wish to use a quick and dirty optical test to determine how strong a particular mirror is. We can lay it on the floor as shown in figure c, and use it to make an image of a lamp mounted on the ceiling overhead, which we assume is very far away compared to the radius of curvature of the mirror, so that the mirror intercepts only a very narrow cone of rays from the lamp. This cone is so narrow that its rays are nearly parallel, and $\theta_o$ is nearly zero. The real image can be observed on a piece of paper. By moving the paper nearer and farther, we can bring the image into focus, at which point we know the paper is located at the image point. Since $\theta_o \approx 0$, we have $\theta_i \approx \theta_f$, and we can then determine this mirror’s focal angle either by measuring $\theta_i$ directly with a protractor, or indirectly via trigonometry. A strong mirror will bring the rays together to form an image close to the mirror, and these rays will form a blunt-angled cone with a large $\theta_i$ and $\theta_f$.

**An alternative optical test**

- Figure c shows an alternative optical test. Rather than placing the object at infinity as in figure b, we adjust it so that the image is right on top of the object. Points O and I coincide, and the rays are reflected right back on top of themselves. If we measure the angle $\theta$ shown in figure c, how can we find the focal angle?

- The object and image angles are the same; the angle labeled $\theta$ in the figure equals both of them. We therefore have $\theta_i + \theta_o = \theta = \theta_f$. Comparing figures b and c, it is indeed plausible that the angles are related by a factor of two.

At this point, we could consider our work to be done. Typically, we know the strength of the mirror, and we want to find the image location for a given object location. Given the mirror’s focal angle and the object location, we can determine $\theta_o$ by trigonometry, subtract to find $\theta_i = \theta_f - \theta_o$, and then do more trig to find the image location.

There is, however, a shortcut that can save us from doing so much work. Figure a/3 shows two right triangles whose legs of length 1 coincide and whose acute angles are $\theta_o$ and $\theta_i$. These can be related...
The object and image distances

e / Mirror 1 is weaker than mirror 2. It has a shallower curvature, a longer focal length, and a smaller focal angle. It reflects rays at angles not much different than those that would be produced with a flat mirror.

by trigonometry to the object and image distances shown in figure d:

\[ \tan \theta_o = 1/d_o \quad \tan \theta_i = 1/d_i \]

Ever since chapter 2, we’ve been assuming small angles. For small angles, we can use the small-angle approximation \( \tan x \approx x \) (for \( x \) in radians), giving simply

\[ \theta_o = 1/d_o \quad \theta_i = 1/d_i. \]

We likewise define a distance called the focal length, \( f \) according to

\[ \theta_f = 1/f. \]

In figure b, \( f \) is the distance from the mirror to the place where the rays cross. We can now reexpress the equation relating the object and image positions as

\[ \frac{1}{f} = \frac{1}{d_i} + \frac{1}{d_o}. \]

Figure e summarizes the interpretation of the focal length and focal angle.

Which form is better, \( \theta_f = \theta_i + \theta_o \) or \( 1/f = 1/d_i + 1/d_o \)? The angular form has in its favor its simplicity and its straightforward visual interpretation, but there are two reasons why we might prefer the second version. First, the numerical values of the angles depend on what we mean by “one unit” for the distance shown as 1 in figure a/1. Second, it is usually easier to measure distances rather than angles, so the distance form is more convenient for number crunching. Neither form is superior overall, and we will often need to use both to solve any given problem.

A searchlight example 2

Suppose we need to create a parallel beam of light, as in a searchlight. Where should we place the lightbulb? A parallel beam has zero angle between its rays, so \( \theta_i = 0 \). To place the lightbulb correctly, however, we need to know a distance, not an angle: the distance \( d_o \) between the bulb and the mirror. The problem involves a mixture of distances and angles, so we need to get everything in terms of one or the other in order to solve it. Since

1There is a standard piece of terminology which is that the “focal point” is the point lying on the optical axis at a distance from the mirror equal to the focal length. This term isn’t particularly helpful, because it names a location where nothing normally happens. In particular, it is not normally the place where the rays come to a focus! — that would be the image point. In other words, we don’t normally have \( d_i = f \), unless perhaps \( d_o = \infty \). A recent online discussion among some physics teachers (https://carnot.physics.buffalo.edu/archives, Feb. 2006) showed that many disliked the terminology, felt it was misleading, or didn’t know it and would have misinterpreted it if they had come across it. That is, it appears to be what grammarians call a “skunked term” — a word that bothers half the population when it’s used incorrectly, and the other half when it’s used correctly.

2I would like to thank Fouad Ajami for pointing out the pedagogical advantages of using both equations side by side.
the goal is to find a distance, let’s figure out the image distance corresponding to the given angle $\theta_i = 0$. These are related by $d_i = 1/\theta_i$, so we have $d_i = \infty$. (Yes, dividing by zero gives infinity. Don’t be afraid of infinity. Infinity is a useful problem-solving device.) Solving the distance equation for $d_o$, we have

$$d_o = \left(\frac{1}{f} - \frac{1}{d_i}\right)^{-1}$$

$$= \left(\frac{1}{f} - 0\right)^{-1}$$

$$= f$$

The bulb has to be placed at a distance from the mirror equal to its focal point.

### Diopters example 3

An equation like $d_i = 1/\theta_i$ really doesn’t make sense in terms of units. Angles are unitless, since radians aren’t really units, so the right-hand side is unitless. We can’t have a left-hand side with units of distance if the right-hand side of the same equation is unitless. This is an artifact of my cavalier statement that the conical bundles of rays spread out to a distance of 1 from the axis where they strike the mirror, without specifying the units used to measure this 1. In real life, optometrists define the thing we’re calling $\theta_i = 1/d_i$ as the “dioptic strength” of a lens or mirror, and measure it in units of inverse meters ($\text{m}^{-1}$), also known as diopters (1 D=1 m$^{-1}$).

### Magnification

We have already discussed in the previous chapter how to find the magnification of a virtual image made by a curved mirror. The result is the same for a real image, and we omit the proof, which is very similar. In our new notation, the result is $M = d_i/d_o$. A numerical example is given in section 30.2.

### 30.2 Other cases with curved mirrors

The equation $d_i = \left(\frac{1}{f} - \frac{1}{d_o}\right)^{-1}$ can easily produce a negative result, but we have been thinking of $d_i$ as a distance, and distances can’t be negative. A similar problem occurs with $\theta_i = \theta_f - \theta_o$ for $\theta_o > \theta_f$. What’s going on here?

The interpretation of the angular equation is straightforward. As we bring the object closer and closer to the image, $\theta_o$ gets bigger and bigger, and eventually we reach a point where $\theta_o = \theta_f$ and $\theta_i = 0$. This large object angle represents a bundle of rays forming a cone that is very broad, so broad that the mirror can no longer bend them back so that they reconverge on the axis. The image angle $\theta_i = 0$ represents an outgoing bundle of rays that are parallel. The outgoing rays never cross, so this is not a real image, unless we want to be charitable and say that the rays cross at infinity. If we go on bringing the object even closer, we get a virtual image.
To analyze the distance equation, let’s look at a graph of \( d_i \) as a function of \( d_o \). The branch on the upper right corresponds to the case of a real image. Strictly speaking, this is the only part of the graph that we’ve proven corresponds to reality, since we never did any geometry for other cases, such as virtual images. As discussed in the previous section, making \( d_o \) bigger causes \( d_i \) to become smaller, and vice-versa.

Letting \( d_o \) be less than \( f \) is equivalent to \( \theta_o > \theta_f \): a virtual image is produced on the far side of the mirror. This is the first example of Wigner’s “unreasonable effectiveness of mathematics” that we have encountered in optics. Even though our proof depended on the assumption that the image was real, the equation we derived turns out to be applicable to virtual images, provided that we either interpret the positive and negative signs in a certain way, or else modify the equation to have different positive and negative signs.

**self-check A**

Interpret the three places where, in physically realistic parts of the graph, the graph approaches one of the dashed lines. [This will come more naturally if you have learned the concept of limits in a math class.]

Answer, p. 1041

---

*A flat mirror*  
We can even apply the equation to a flat mirror. As a sphere gets
bigger and bigger, its surface is more and more gently curved. The planet Earth is so large, for example, that we cannot even perceive the curvature of its surface. To represent a flat mirror, we let the mirror’s radius of curvature, and its focal length, become infinite. Dividing by infinity gives zero, so we have

\[ \frac{1}{d_o} = -\frac{1}{d_i}, \]

or

\[ d_o = -d_i. \]

If we interpret the minus sign as indicating a virtual image on the far side of the mirror from the object, this makes sense.

It turns out that for any of the six possible combinations of real or virtual images formed by converging or diverging lenses or mirrors, we can apply equations of the form

\[ \theta_f = \theta_i + \theta_o \]

and

\[ \frac{1}{f} = \frac{1}{d_i} + \frac{1}{d_o}, \]

with only a modification of plus or minus signs. There are two possible approaches here. The approach we have been using so far is the more popular approach in American textbooks: leave the equation the same, but attach interpretations to the resulting negative or positive values of the variables. The trouble with this approach is that one is then forced to memorize tables of sign conventions, e.g., that the value of \( d_i \) should be negative when the image is a virtual image formed by a converging mirror. Positive and negative signs also have to be memorized for focal lengths. Ugh! It’s highly unlikely that any student has ever retained these lengthy tables in his or her mind for more than five minutes after handing in the final exam in a physics course. Of course one can always look such things up when they are needed, but the effect is to turn the whole thing into an exercise in blindly plugging numbers into formulas.

As you have gathered by now, there is another method which I think is better, and which I’ll use throughout the rest of this book. In this method, all distances and angles are \textit{positive by definition}, and we put in positive and negative signs in the \textit{equations} depending on the situation. (I thought I was the first to invent this method, but I’ve been told that this is known as the European sign convention, and that it’s fairly common in Europe.) Rather than memorizing these signs, we start with the generic equations

\[ \theta_f = \pm \theta_i \pm \theta_o, \]

\[ \frac{1}{f} = \pm \frac{1}{d_i} \pm \frac{1}{d_o}, \]
and then determine the signs by a two-step method that depends on ray diagrams. There are really only two signs to determine, not four; the signs in the two equations match up in the way you’d expect. The method is as follows:

1. Use ray diagrams to decide whether \( \theta_o \) and \( \theta_i \) vary in the same way or in opposite ways. (In other words, decide whether making \( \theta_o \) greater results in a greater value of \( \theta_i \) or a smaller one.) Based on this, decide whether the two signs in the angle equation are the same or opposite. If the signs are opposite, go on to step 2 to determine which is positive and which is negative.

2. If the signs are opposite, we need to decide which is the positive one and which is the negative. Since the focal angle is never negative, the smaller angle must be the one with a minus sign.

In step 1, many students have trouble drawing the ray diagram correctly. For simplicity, you should always do your diagram for a point on the object that is on the axis of the mirror, and let one of your rays be the one that is emitted along the axis and reflected straight back on itself, as in the figures in section 30.1. As shown in figure a/4 in section 30.1, there are four angles involved: two at the mirror, one at the object (\( \theta_o \)), and one at the image (\( \theta_i \)). Make sure to draw in the normal to the mirror so that you can see the two angles at the mirror. These two angles are equal, so as you change the object position, they fan out or fan in, like opening or closing a book. Once you’ve drawn this effect, you should easily be able to tell whether \( \theta_o \) and \( \theta_i \) change in the same way or in opposite ways.

Although focal lengths are always positive in the method used in this book, you should be aware that diverging mirrors and lenses are assigned negative focal lengths in the other method, so if you see a lens labeled \( f = -30 \text{ cm} \), you’ll know what it means.

An anti-shoplifting mirror

- Convenience stores often install a diverging mirror so that the clerk has a view of the whole store and can catch shoplifters. Use a ray diagram to show that the image is reduced, bringing more into the clerk’s field of view. If the focal length of the mirror is 3.0 m, and the mirror is 7.0 m from the farthest wall, how deep is the image of the store?

- As shown in ray diagram g/1, \( d_i \) is less than \( d_o \). The magnification, \( M = d_i / d_o \), will be less than one, i.e., the image is actually reduced rather than magnified.

Apply the method outlined above for determining the plus and minus signs. Step 1: The object is the point on the opposite wall. As an experiment, g/2, move the object closer. I did these drawings using illustration software, but if you were doing them by hand, you’d want to make the scale much larger for greater accuracy. Also, although I split figure g into two separate drawings.
in order to make them easier to understand, you’re less likely to make a mistake if you do them on top of each other.

The two angles at the mirror fan out from the normal. Increasing $\theta_o$ has clearly made $\theta_i$ larger as well. (All four angles got bigger.) There must be a cancellation of the effects of changing the two terms on the right in the same way, and the only way to get such a cancellation is if the two terms in the angle equation have opposite signs:

$$\theta_f = +\theta_i - \theta_o$$

or

$$\theta_f = -\theta_i + \theta_o.$$

Step 2: Now which is the positive term and which is negative? Since the image angle is bigger than the object angle, the angle equation must be

$$\theta_f = \theta_i - \theta_o,$$

in order to give a positive result for the focal angle. The signs of the distance equation behave the same way:

$$\frac{1}{f} = \frac{1}{d_i} - \frac{1}{d_o}.$$

Solving for $d_i$, we find

$$d_i = \left(\frac{1}{f} + \frac{1}{d_o}\right)^{-1} = 2.1 \text{ m}.$$

The image of the store is reduced by a factor of $2.1/7.0 = 0.3$, i.e., it is smaller by 70%.

---

**Example 6**

A shortcut for real images

In the case of a real image, there is a shortcut for step 1, the determination of the signs. In a real image, the rays cross at both the object and the image. We can therefore time-reverse the ray diagram, so that all the rays are coming from the image and reconverging at the object. Object and image swap roles. Due to this time-reversal symmetry, the object and image cannot be treated differently in any of the equations, and they must therefore have the same signs. They are both positive, since they must add up to a positive result.

### 30.3 Aberrations

An imperfection or distortion in an image is called an aberration. An aberration can be produced by a flaw in a lens or mirror, but even with a perfect optical surface some degree of aberration is unavoidable. To see why, consider the mathematical approximation
A diverging mirror in the shape of a sphere. The image is reduced ($M < 1$). This is similar to example 5, but here the image is distorted because the mirror’s curve is not shallow.

we’ve been making, which is that the depth of the mirror’s curve is small compared to $d_o$ and $d_i$. Since only a flat mirror can satisfy this shallow-mirror condition perfectly, any curved mirror will deviate somewhat from the mathematical behavior we derived by assuming that condition. There are two main types of aberration in curved mirrors, and these also occur with lenses.

(1) An object on the axis of the lens or mirror may be imaged correctly, but off-axis objects may be out of focus or distorted. In a camera, this type of aberration would show up as a fuzziness or warping near the sides of the picture when the center was perfectly focused. An example of this is shown in figure i, and in that particular example, the aberration is not a sign that the equipment was of low quality or wasn’t right for the job but rather an inevitable result of trying to flatten a panoramic view; in the limit of a 360-degree panorama, the problem would be similar to the problem of representing the Earth’s surface on a flat map, which can’t be accomplished without distortion.

(2) The image may be sharp when the object is at certain distances and blurry when it is at other distances. The blurriness occurs because the rays do not all cross at exactly the same point. If we know in advance the distance of the objects with which the mirror or lens will be used, then we can optimize the shape of the optical surface to make in-focus images in that situation. For instance, a spherical mirror will produce a perfect image of an object that is at the center of the sphere, because each ray is reflected directly onto the radius along which it was emitted. For objects at greater distances, however, the focus will be somewhat blurry. In astronomy the objects being used are always at infinity, so a spherical mirror
is a poor choice for a telescope. A different shape (a parabola) is better specialized for astronomy.

One way of decreasing aberration is to use a small-diameter mirror or lens, or block most of the light with an opaque screen with a hole in it, so that only light that comes in close to the axis can get through. Either way, we are using a smaller portion of the lens or mirror whose curvature will be more shallow, thereby making the shallow-mirror (or thin-lens) approximation more accurate. Your eye does this by narrowing down the pupil to a smaller hole. In a camera, there is either an automatic or manual adjustment, and narrowing the opening is called “stopping down.” The disadvantage of stopping down is that light is wasted, so the image will be dimmer.
or a longer exposure must be used.

Even though the spherical mirror (solid line) is not well adapted for viewing an object at infinity, we can improve its performance greatly by stopping it down. Now the only part of the mirror being used is the central portion, where its shape is virtually indistinguishable from a parabola (dashed line).

What I would suggest you take away from this discussion for the sake of your general scientific education is simply an understanding of what an aberration is, why it occurs, and how it can be reduced, not detailed facts about specific types of aberrations.

The Hubble Space Telescope was placed into orbit with faulty optics in 1990. Its main mirror was supposed to have been nearly parabolic, since it is an astronomical telescope, meant for producing images of objects at infinity. However, contractor Perkin Elmer had delivered a faulty mirror, which produced aberrations. The large photo shows astronauts putting correcting mirrors in place in 1993. The two small photos show images produced by the telescope before and after the fix.
Summary

Selected vocabulary

focal length . . . a property of a lens or mirror, equal to the distance from the lens or mirror to the image it forms of an object that is infinitely far away

Notation

\[ f \] . . . . . . . . . . the focal length
\[ d_o \] . . . . . . . . . . the distance of the object from the mirror
\[ d_i \] . . . . . . . . . . the distance of the image from the mirror
\[ \theta_f \] . . . . . . . . . . the focal angle, defined as \(1/f\)
\[ \theta_o \] . . . . . . . . . . the object angle, defined as \(1/d_o\)
\[ \theta_i \] . . . . . . . . . . the image angle, defined as \(1/d_i\)

Other terminology and notation

\[ f > 0 \] . . . . . . . describes a converging lens or mirror; in this book, all focal lengths are positive, so there is no such implication
\[ f < 0 \] . . . . . . . describes a diverging lens or mirror; in this book, all focal lengths are positive
\[ M < 0 \] . . . . . . . indicates an inverted image; in this book, all magnifications are positive

Summary

Every lens or mirror has a property called the focal length, which is defined as the distance from the lens or mirror to the image it forms of an object that is infinitely far away. A stronger lens or mirror has a shorter focal length.

The relationship between the locations of an object and its image formed by a lens or mirror can always be expressed by equations of the form

\[ \theta_f = \pm \theta_i \pm \theta_o \]
\[ \frac{1}{f} = \pm \frac{1}{d_i} \pm \frac{1}{d_o}. \]

The choice of plus and minus signs depends on whether we are dealing with a lens or a mirror, whether the lens or mirror is converging or diverging, and whether the image is real or virtual. A method for determining the plus and minus signs is as follows:

1. Use ray diagrams to decide whether \( \theta_i \) and \( \theta_o \) vary in the same way or in opposite ways. Based on this, decide whether the two signs in the equation are the same or opposite. If the signs are opposite, go on to step 2 to determine which is positive and which is negative.

2. If the signs are opposite, we need to decide which is the positive one and which is the negative. Since the focal angle is never negative, the smaller angle must be the one with a minus sign.
Once the correct form of the equation has been determined, the magnification can be found via the equation

\[ M = \frac{d_i}{d_o}. \]

This equation expresses the idea that the entire image-world is shrunk consistently in all three dimensions.
Problems

Key
✓ A computerized answer check is available online.
∫ A problem that requires calculus.
★ A difficult problem.

1 Apply the equation \( M = \frac{d_i}{d_o} \) to the case of a flat mirror. \( \triangleright \) Solution, p. 1033

2 Use the method described in the text to derive the equation relating object distance to image distance for the case of a virtual image produced by a converging mirror. \( \triangleright \) Solution, p. 1033

3 (a) Make up a numerical example of a virtual image formed by a converging mirror with a certain focal length, and determine the magnification. (You will need the result of problem 2.) Make sure to choose values of \( d_o \) and \( f \) that would actually produce a virtual image, not a real one. Now change the location of the object a little bit and redetermine the magnification, showing that it changes. At my local department store, the cosmetics department sells hand mirrors advertised as giving a magnification of 5 times. How would you interpret this?

(b) Suppose a Newtonian telescope is being used for astronomical observing. Assume for simplicity that no eyepiece is used, and assume a value for the focal length of the mirror that would be reasonable for an amateur instrument that is to fit in a closet. Is the angular magnification different for objects at different distances? For example, you could consider two planets, one of which is twice as far as the other. \( \triangleright \) Solution, p. 1033

4 (a) Find a case where the magnification of a curved mirror is infinite. Is the angular magnification infinite from any realistic viewing position? (b) Explain why an arbitrarily large magnification can’t be achieved by having a sufficiently small value of \( d_o \). \( \triangleright \) Solution, p. 1033

5 The figure shows a device for constructing a realistic optical illusion. Two mirrors of equal focal length are put against each other with their silvered surfaces facing inward. A small object placed in the bottom of the cavity will have its image projected in the air above. The way it works is that the top mirror produces a virtual image, and the bottom mirror then creates a real image of the virtual image. (a) Show that if the image is to be positioned as shown, at the mouth of the cavity, then the focal length of the mirrors is related to the dimension \( h \) via the equation

\[
\frac{1}{f} = \frac{1}{h} + \frac{1}{h + \left( \frac{1}{h} - \frac{1}{f} \right)^{-1}}.
\]

(b) Restate the equation in terms of a single variable \( x = h/f \), and
show that there are two solutions for \( x \). Which solution is physically consistent with the assumptions of the calculation? \(*\)

6 A concave surface that reflects sound waves can act just like a converging mirror. Suppose that, standing near such a surface, you are able to find a point where you can place your head so that your own whispers are focused back on your head, so that they sound loud to you. Given your distance to the surface, what is the surface’s focal length? \(\sqrt{7}\)

7 Find the focal length of the mirror in problem 5 in chapter 28. \(\sqrt{8}\)

8 Rank the focal lengths of the mirrors in the figure, from shortest to longest. Explain.

9 (a) A converging mirror is being used to create a virtual image. What is the range of possible magnifications? (b) Do the same for the other types of images that can be formed by curved mirrors (both converging and diverging).

10 (a) A converging mirror with a focal length of 20 cm is used to create an image, using an object at a distance of 10 cm. Is the image real, or is it virtual? (b) How about \( f = 20 \text{ cm} \) and \( d_o = 30 \text{ cm} \)? (c) What if it was a diverging mirror with \( f = 20 \text{ cm} \) and \( d_o = 10 \text{ cm} \)? (d) A diverging mirror with \( f = 20 \text{ cm} \) and \( d_o = 30 \text{ cm} \).

11 A diverging mirror of focal length \( f \) is fixed, and faces down. An object is dropped from the surface of the mirror, and falls away from it with acceleration \( g \). The goal of the problem is to find the maximum velocity of the image.

(a) Describe the motion of the image verbally, and explain why we should expect there to be a maximum velocity.

(b) Use arguments based on units to determine the form of the solution, up to an unknown unitless multiplicative constant.

(c) Complete the solution by determining the unitless constant.

12 A mechanical linkage is a device that changes one type of motion into another. The most familiar example occurs in a gasoline car’s engine, where a connecting rod changes the linear motion of the piston into circular motion of the crankshaft. The top panel of the figure shows a mechanical linkage invented by Peaucellier in 1864, and independently by Lipkin around the same time. It consists of six rods joined by hinges, the four short ones forming a rhombus. Point \( O \) is fixed in space, but the apparatus is free to rotate about \( O \). Motion at \( P \) is transformed into a different motion at \( P' \) (or vice versa).

Geometrically, the linkage is a mechanical implementation of the ancient problem of inversion in a circle. Considering the case in
which the rhombus is folded flat, let the $k$ be the distance from $O$ to the point where $P$ and $P'$ coincide. Form the circle of radius $k$ with its center at $O$. As $P$ and $P'$ move in and out, points on the inside of the circle are always mapped to points on its outside, such that $rr' = k^2$. That is, the linkage is a type of analog computer that exactly solves the problem of finding the inverse of a number $r$. Inversion in a circle has many remarkable geometrical properties, discussed in H.S.M. Coxeter, *Introduction to Geometry*, Wiley, 1961.

If a pen is inserted through a hole at $P$, and $P'$ is traced over a geometrical figure, the Peaucellier linkage can be used to draw a kind of image of the figure.

A related problem is the construction of pictures, like the one in the bottom panel of the figure, called anamorphs. The drawing of the column on the paper is highly distorted, but when the reflecting cylinder is placed in the correct spot on top of the page, an undistorted image is produced inside the cylinder. (Wide-format movie technologies such as Cinemascope are based on similar principles.)

Show that the Peaucellier linkage does not convert correctly between an image and its anamorph, and design a modified version of the linkage that does. Some knowledge of analytic geometry will be helpful.
Exercise 30: Object and image distances

Equipment:
- optical benches
- converging mirrors
- illuminated objects

1. Set up the optical bench with the mirror at zero on the centimeter scale. Set up the illuminated object on the bench as well.

2. Each group will locate the image for their own value of the object distance, by finding where a piece of paper has to be placed in order to see the image on it. (The instructor will do one point as well.) Note that you will have to tilt the mirror a little so that the paper on which you project the image doesn’t block the light from the illuminated object.

   Is the image real or virtual? How do you know? Is it inverted, or uninverted?

   Draw a ray diagram.

3. Measure the image distance and write your result in the table on the board. Do the same for the magnification.

4. What do you notice about the trend of the data on the board? Draw a second ray diagram with a different object distance, and show why this makes sense. Some tips for doing this correctly: (1) For simplicity, use the point on the object that is on the mirror’s axis. (2) You need to trace two rays to locate the image. To save work, don’t just do two rays at random angles. You can either use the on-axis ray as one ray, or do two rays that come off at the same angle, one above and one below the axis. (3) Where each ray hits the mirror, draw the normal line, and make sure the ray is at equal angles on both sides of the normal.

5. We will find the mirror’s focal length from the instructor’s data-point. Then, using this focal length, calculate a theoretical prediction of the image distance, and write it on the board next to the experimentally determined image distance.
Three stages in the evolution of the eye. The flatworm has two eye pits. The nautilus’s eyes are pinhole cameras. The human eye incorporates a lens.

Chapter 31

Refraction

Economists normally consider free markets to be the natural way of judging the monetary value of something, but social scientists also use questionnaires to gauge the relative value of privileges, disadvantages, or possessions that cannot be bought or sold. They ask people to imagine that they could trade one thing for another and ask which they would choose. One interesting result is that the average light-skinned person in the U.S. would rather lose an arm than suffer the racist treatment routinely endured by African-Americans. Even more impressive is the value of sight. Many prospective parents can imagine without too much fear having a deaf child, but would have a far more difficult time coping with raising a blind one.

So great is the value attached to sight that some have imbued it with mystical aspects. Joan of Arc saw visions, and my college has a “vision statement.” Christian fundamentalists who perceive a conflict between evolution and their religion have claimed that the eye is such a perfect device that it could never have arisen through a process as helter-skelter as evolution, or that it could not have evolved because half of an eye would be useless. In fact, the structure of an eye is fundamentally dictated by physics, and it has arisen
separately by evolution somewhere between eight and 40 times, depending on which biologist you ask. We humans have a version of the eye that can be traced back to the evolution of a light-sensitive “eye spot” on the head of an ancient invertebrate. A sunken pit then developed so that the eye would only receive light from one direction, allowing the organism to tell where the light was coming from. (Modern flatworms have this type of eye.) The top of the pit then became partially covered, leaving a hole, for even greater directionality (as in the nautilus). At some point the cavity became filled with jelly, and this jelly finally became a lens, resulting in the general type of eye that we share with the bony fishes and other vertebrates. Far from being a perfect device, the vertebrate eye is marred by a serious design flaw due to the lack of planning or intelligent design in evolution: the nerve cells of the retina and the blood vessels that serve them are all in front of the light-sensitive cells, blocking part of the light. Squids and other molluscs, whose eyes evolved on a separate branch of the evolutionary tree, have a more sensible arrangement, with the light-sensitive cells out in front.

31.1 Refraction

Refraction

The fundamental physical phenomenon at work in the eye is that when light crosses a boundary between two media (such as air and the eye’s jelly), part of its energy is reflected, but part passes into the new medium. In the ray model of light, we describe the original ray as splitting into a reflected ray and a transmitted one (the one that gets through the boundary). Of course the reflected ray goes in a direction that is different from that of the original one, according to the rules of reflection we have already studied. More surprisingly — and this is the crucial point for making your eye focus light — the transmitted ray is bent somewhat as well. This bending phenomenon is called refraction. The origin of the word is the same as that of the word “fracture,” i.e., the ray is bent or “broken.” (Keep in mind, however, that light rays are not physical objects that can really be “broken.”) Refraction occurs with all waves, not just light waves.

The actual anatomy of the eye, b, is quite complex, but in essence it is very much like every other optical device based on refraction. The rays are bent when they pass through the front surface of the eye, c. Rays that enter farther from the central axis are bent more, with the result that an image is formed on the retina. There is only one slightly novel aspect of the situation. In most human-built optical devices, such as a movie projector, the light is bent as it passes into a lens, bent again as it reemerges, and then reaches a focus beyond the lens. In the eye, however, the “screen” is inside the eye, so the rays are only refracted once, on entering the jelly, and never emerge
A common misconception is that the “lens” of the eye is what does the focusing. All the transparent parts of the eye are made of fairly similar stuff, so the dramatic change in medium is when a ray crosses from the air into the eye (at the outside surface of the cornea). This is where nearly all the refraction takes place. The lens medium differs only slightly in its optical properties from the rest of the eye, so very little refraction occurs as light enters and exits the lens. The lens, whose shape is adjusted by muscles attached to it, is only meant for fine-tuning the focus to form images of near or far objects.

**Refractive properties of media**

What are the rules governing refraction? The first thing to observe is that just as with reflection, the new, bent part of the ray lies in the same plane as the normal (perpendicular) and the incident ray, d.

If you try shooting a beam of light at the boundary between two substances, say water and air, you’ll find that regardless of the angle at which you send in the beam, the part of the beam in the water is always closer to the normal line, e. It doesn’t matter if the ray is entering the water or leaving, so refraction is symmetric with respect to time-reversal, f.

If, instead of water and air, you try another combination of substances, say plastic and gasoline, again you’ll find that the ray’s angle with respect to the normal is consistently smaller in one and larger in the other. Also, we find that if substance A has rays closer to normal than in B, and B has rays closer to normal than in C, then A has rays closer to normal than C. This means that we can rank-order all materials according to their refractive properties. Isaac Newton did so, including in his list many amusing substances, such as “Danzig vitriol” and “a pseudo-topazius, being a natural, pellucid, brittle, hairy stone, of a yellow color.” Several general rules can be inferred from such a list:

- Vacuum lies at one end of the list. In refraction across the interface between vacuum and any other medium, the other medium has rays closer to the normal.

- Among gases, the ray gets closer to the normal if you increase the density of the gas by pressurizing it more.

- The refractive properties of liquid mixtures and solutions vary in a smooth and systematic manner as the proportions of the mixture are changed.

- Denser substances usually, but not always, have rays closer to the normal.
The second and third rules provide us with a method for measuring the density of an unknown sample of gas, or the concentration of a solution. The latter technique is very commonly used, and the CRC Handbook of Physics and Chemistry, for instance, contains extensive tables of the refractive properties of sugar solutions, cat urine, and so on.

**Snell’s law**

The numerical rule governing refraction was discovered by Snell, who must have collected experimental data something like what is shown on this graph and then attempted by trial and error to find the right equation. The equation he came up with was

$$\frac{\sin \theta_1}{\sin \theta_2} = \text{constant}. $$

The value of the constant would depend on the combination of media used. For instance, any one of the data points in the graph would have sufficed to show that the constant was 1.3 for an air-water interface (taking air to be substance 1 and water to be substance 2).

Snell further found that if media A and B gave a constant $K_{AB}$ and media B and C gave a constant $K_{BC}$, then refraction at an interface between A and C would be described by a constant equal to the product, $K_{AC} = K_{AB}K_{BC}$. This is exactly what one would expect if the constant depended on the ratio of some number characterizing one medium to the number characteristic of the second medium. This number is called the *index of refraction* of the medium, written as $n$ in equations. Since measuring the angles would only allow him to determine the *ratio* of the indices of refraction of two media, Snell had to pick some medium and define it as having $n = 1$. He chose to define vacuum as having $n = 1$. (The index of refraction of air at normal atmospheric pressure is 1.0003, so for most purposes it is a good approximation to assume that air has $n = 1$.) He also had to decide which way to define the ratio, and he chose to define it so that media with their rays closer to the normal would have larger indices of refraction. This had the advantage that denser media would typically have higher indices of refraction, and for this reason the index of refraction is also referred to as the optical density. Written in terms of indices of refraction, Snell’s equation becomes

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

but rewriting it in the form

$$\frac{\sin \theta_1}{\sin \theta_2} = \text{constant}. $$
\[ n_1 \sin \theta_1 = n_2 \sin \theta_2 \]

[relationship between angles of rays at the interface between media with indices of refraction \( n_1 \) and \( n_2 \); angles are defined with respect to the normal]

makes us less likely to get the 1’s and 2’s mixed up, so this the way most people remember Snell’s law. A few indices of refraction are given in the back of the book.

**Self-check A**

1. What would the graph look like for two substances with the same index of refraction?

2. Based on the graph, when does refraction at an air-water interface change the direction of a ray most strongly?  
   \[ \text{Answer, p. 1041} \]

**Finding an angle using Snell’s law example 1**

\[ \alpha \]

A submarine shines its searchlight up toward the surface of the water. What is the angle \( \alpha \) shown in the figure?

\[ \alpha \]

The tricky part is that Snell’s law refers to the angles with respect to the normal. Forgetting this is a very common mistake. The beam is at an angle of 30° with respect to the normal in the water. Let’s refer to the air as medium 1 and the water as 2. Solving Snell’s law for \( \theta_1 \), we find

\[ \theta_1 = \sin^{-1} \left( \frac{n_2}{n_1} \sin \theta_2 \right) . \]

As mentioned above, air has an index of refraction very close to 1, and water’s is about 1.3, so we find \( \theta_1 = 40° \). The angle \( \alpha \) is therefore 50°.

**The index of refraction is related to the speed of light.**

What neither Snell nor Newton knew was that there is a very simple interpretation of the index of refraction. This may come as a relief to the reader who is taken aback by the complex reasoning involving proportionalities that led to its definition. Later experiments showed that the index of refraction of a medium was inversely proportional to the speed of light in that medium. Since \( c \) is defined as the speed of light in vacuum, and \( n = 1 \) is defined as the index of refraction of vacuum, we have

\[ n = \frac{c}{v} . \]

\[ [n = \text{medium’s index of refraction}, v = \text{speed of light in that medium, } c = \text{speed of light in a vacuum}] \]

Many textbooks start with this as the definition of the index of refraction, although that approach makes the quantity’s name somewhat of a mystery, and leaves students wondering why \( c/v \) was used rather than \( v/c \). It should also be noted that measuring angles of refraction is a far more practical method for determining \( n \) than direct measurement of the speed of light in the substance of interest.
A mechanical model of Snell’s law

Why should refraction be related to the speed of light? The mechanical model shown in the figure may help to make this more plausible. Suppose medium 2 is thick, sticky mud, which slows down the car. The car’s right wheel hits the mud first, causing the right side of the car to slow down. This will cause the car to turn to the right until it moves far enough forward for the left wheel to cross into the mud. After that, the two sides of the car will once again be moving at the same speed, and the car will go straight.

Of course, light isn’t a car. Why should a beam of light have anything resembling a “left wheel” and “right wheel?” After all, the mechanical model would predict that a motorcycle would go straight, and a motorcycle seems like a better approximation to a ray of light than a car. The whole thing is just a model, not a description of physical reality.

A derivation of Snell’s law

However intuitively appealing the mechanical model may be, light is a wave, and we should be using wave models to describe refraction. In fact Snell’s law can be derived quite simply from wave concepts. Figure j shows the refraction of a water wave. The water in the upper left part of the tank is shallower, so the speed of the waves is slower there, and their wavelengths is shorter. The reflected part of the wave is also very faintly visible.
In the close-up view on the right, the dashed lines are normals to the interface. The two marked angles on the right side are both equal to $\theta_1$, and the two on the left to $\theta_2$.

Trigonometry gives

$$\sin \theta_1 = \frac{\lambda_1}{h} \quad \text{and} \quad \sin \theta_2 = \frac{\lambda_2}{h}.$$ 

Eliminating $h$ by dividing the equations, we find

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{\lambda_1}{\lambda_2}.$$ 

The frequencies of the two waves must be equal or else they would get out of step, so by $v = f\lambda$ we know that their wavelengths are proportional to their velocities. Combining $\lambda \propto v$ with $v \propto 1/n$ gives $\lambda \propto 1/n$, so we find

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

which is one form of Snell’s law.

**Ocean waves near and far from shore**

Ocean waves are formed by winds, typically on the open sea, and the wavefronts are perpendicular to the direction of the wind that formed them. At the beach, however, you have undoubtedly observed that waves tend come in with their wavefronts very nearly (but not exactly) parallel to the shoreline. This is because the speed of water waves in shallow water depends on depth: the shallower the water, the slower the wave. Although the change from the fast-wave region to the slow-wave region is gradual rather than abrupt, there is still refraction, and the wave motion is nearly perpendicular to the normal in the slow region.

**Color and refraction**

In general, the speed of light in a medium depends both on the medium and on the wavelength of the light. Another way of saying it is that a medium’s index of refraction varies with wavelength. This is why a prism can be used to split up a beam of white light into a rainbow. Each wavelength of light is refracted through a different angle.

**How much light is reflected, and how much is transmitted?**

In chapter 20 we developed an equation for the percentage of the wave energy that is transmitted and the percentage reflected at a boundary between media. This was only done in the case of waves in one dimension, however, and rather than discuss the full three-dimensional generalization it will be more useful to go into some qualitative observations about what happens. First, reflection happens...
only at the interface between two media, and two media with the same index of refraction act as if they were a single medium. Thus, at the interface between media with the same index of refraction, there is no reflection, and the ray keeps going straight. Continuing this line of thought, it is not surprising that we observe very little reflection at an interface between media with similar indices of refraction.

The next thing to note is that it is possible to have situations where no possible angle for the refracted ray can satisfy Snell’s law. Solving Snell’s law for $\theta_2$, we find

$$\theta_2 = \sin^{-1} \left( \frac{n_1}{n_2} \sin \theta_1 \right),$$

and if $n_1$ is greater than $n_2$, then there will be large values of $\theta_1$ for which the quantity $(n_1/n_2) \sin \theta$ is greater than one, meaning that your calculator will flash an error message at you when you try to take the inverse sine. What can happen physically in such a situation? The answer is that all the light is reflected, so there is no refracted ray. This phenomenon is known as total internal reflection, and is used in the fiber-optic cables that nowadays carry almost all long-distance telephone calls. The electrical signals from your phone travel to a switching center, where they are converted from electricity into light. From there, the light is sent across the country in a thin transparent fiber. The light is aimed straight into the end of the fiber, and as long as the fiber never goes through any turns that are too sharp, the light will always encounter the edge of the fiber at an angle sufficiently oblique to give total internal reflection. If the fiber-optic cable is thick enough, one can see an image at one end of whatever the other end is pointed at.

Alternatively, a bundle of cables can be used, since a single thick cable is too hard to bend. This technique for seeing around corners is useful for making surgery less traumatic. Instead of cutting a person wide open, a surgeon can make a small “keyhole” incision and insert a bundle of fiber-optic cable (known as an endoscope) into the body.

Since rays at sufficiently large angles with respect to the normal may be completely reflected, it is not surprising that the relative amount of reflection changes depending on the angle of incidence, and is greatest for large angles of incidence.
Discussion questions

A  What index of refraction should a fish have in order to be invisible to other fish?

B  Does a surgeon using an endoscope need a source of light inside the body cavity? If so, how could this be done without inserting a light bulb through the incision?

C  A denser sample of a gas has a higher index of refraction than a less dense sample (i.e., a sample under lower pressure), but why would it not make sense for the index of refraction of a gas to be proportional to density?

D  The earth’s atmosphere gets thinner and thinner as you go higher in altitude. If a ray of light comes from a star that is below the zenith, what will happen to it as it comes into the earth’s atmosphere?

E  Does total internal reflection occur when light in a denser medium encounters a less dense medium, or the other way around? Or can it occur in either case?

31.2 Lenses

Figures n/1 and n/2 show examples of lenses forming images. There is essentially nothing for you to learn about imaging with lenses that is truly new. You already know how to construct and use ray diagrams, and you know about real and virtual images. The concept of the focal length of a lens is the same as for a curved mirror. The equations for locating images and determining magnifications are of the same form. It’s really just a question of flexing your mental muscles on a few examples. The following self-checks and discussion questions will get you started.

self-check B

(1) In figures n/1 and n/2, classify the images as real or virtual.
(2) Glass has an index of refraction that is greater than that of air. Consider the topmost ray in figure n/1. Explain why the ray makes a slight left turn upon entering the lens, and another left turn when it exits.

(3) If the flame in figure n/2 was moved closer to the lens, what would happen to the location of the image? Answer, p. 1041

Discussion questions
A In figures n/1 and n/2, the front and back surfaces are parallel to each other at the center of the lens. What will happen to a ray that enters near the center, but not necessarily along the axis of the lens? Draw a BIG ray diagram, and show a ray that comes from off axis.

In discussion questions B-F, don’t draw ultra-detailed ray diagrams as in A.

B Suppose you wanted to change the setup in figure n/1 so that the location of the actual flame in the figure would instead be occupied by an image of a flame. Where would you have to move the candle to achieve this? What about in n/2?

C There are three qualitatively different types of image formation that can occur with lenses, of which figures n/1 and n/2 exhaust only two. Figure out what the third possibility is. Which of the three possibilities can result in a magnification greater than one? Cf. problem 4, p. 863.

D Classify the examples shown in figure o according to the types of images delineated in discussion question C.

E In figures n/1 and n/2, the only rays drawn were those that happened to enter the lenses. Discuss this in relation to figure o.

F In the right-hand side of figure o, the image viewed through the lens is in focus, but the side of the rose that sticks out from behind the lens is not. Why?

31.3 The lensmaker’s equation

The focal length of a spherical mirror is simply $r/2$, but we cannot expect the focal length of a lens to be given by pure geometry, since it also depends on the index of refraction of the lens. Suppose we have a lens whose front and back surfaces are both spherical. (This is no great loss of generality, since any surface with a sufficiently shallow curvature can be approximated with a sphere.) Then if the lens is immersed in a medium with an index of refraction of 1, its focal length is given approximately by

$$f = \left( n - 1 \right) \left[ \frac{1}{r_1} \pm \frac{1}{r_2} \right]^{-1},$$

where $n$ is the index of refraction and $r_1$ and $r_2$ are the radii of curvature of the two surfaces of the lens. This is known as the lensmaker’s equation. In my opinion it is not particularly worthy of memorization. The positive sign is used when both surfaces are
dispersion of white light by a prism. white light is a mixture of all the wavelengths of the visible spectrum. waves of different wavelengths undergo different amounts of refraction.

31.4 Dispersion

For most materials, we observe that the index of refraction depends slightly on wavelength, being highest at the blue end of the visible spectrum and lowest at the red. For example, white light disperses into a rainbow when it passes through a prism, q. Even when the waves involved aren’t light waves, and even when refraction isn’t of interest, the dependence of wave speed on wavelength is referred to as dispersion. Dispersion inside spherical raindrops is responsible for the creation of rainbows in the sky, and in an optical instrument such as the eye or a camera it is responsible for a type of aberration called chromatic aberration (section 30.3 and problem 2). As we’ll see in section 35.2, dispersion causes a wave that is not a pure sine wave to have its shape distorted as it travels, and also causes the speed at which energy and information are transported by the wave to be different from what one might expect from a naive calculation.
31.5  **The principle of least time for refraction**

We have seen previously how the rules governing straight-line motion of light and reflection of light can be derived from the principle of least time. What about refraction? In the figure, it is indeed plausible that the bending of the ray serves to minimize the time required to get from a point A to point B. If the ray followed the unbent path shown with a dashed line, it would have to travel a longer distance in the medium in which its speed is slower. By bending the correct amount, it can reduce the distance it has to cover in the slower medium without going too far out of its way. It is true that Snell’s law gives exactly the set of angles that minimizes the time required for light to get from one point to another. The proof of this fact is left as an exercise (problem 9, p. 903).

31.6  **Case study: the eye of the jumping spider**

Figure 3 shows an exceptionally cute jumping spider. The jumping spider does not build a web. It stalks its prey like a cat, so it needs excellent eyesight. In some ways, its visual system is more sophisticated and more functional than that of a human, illustrating how evolution does not progress systematically toward “higher” forms of life.
One way in which the spider outdoes us is that it has eight eyes to our two. (Each eye is simple, not compound like that of a fly.) The reason this works well has to do with the trade-off between magnification and field of view. The elongated principal eyes at the front of the head have a large value of $d_i$, resulting in a large magnification $M = d_i/d_o$. This high magnification is used for sophisticated visual tasks like distinguishing prey from a potential mate. (The pretty stripes on the legs in the photo are probably evolved to aid in making this distinction, which is a crucial one on a Saturday night.) As always with a high magnification, this results in a reduction in the field of view: making the image bigger means reducing the amount of the potential image that can actually fit on the retina. The animal has tunnel vision in these forward eyes. To allow it to glimpse prey from other angles, it has the additional eyes on the sides of its head. These are not elongated, and the smaller $d_i$ gives a smaller magnification but a larger field of view. When the spider sees something moving in these eyes, it turns its body so that it can take a look with the front eyes. The tiniest pair of eyes are too small to be useful. These vestigial organs, like the maladaptive human appendix, are an example of the tendency of evolution to produce unfortunate accidents due to the lack of intelligent design. The use of multiple eyes for these multiple purposes is far superior to the two-eye arrangement found in humans, octopuses, etc., especially because of its compactness. If the spider had only two spherical eyes, they would have to have the same front-to-back dimension in order to produce the same acuity, but then the eyes would take up nearly all of the front of the head.

Another beautiful feature of these eyes is that they will never need bifocals. A human eye uses muscles to adjust for seeing near and far, varying $f$ in order to achieve a fixed $d_i$ for differing values of $d_o$. On older models of H. sap., this poorly engineered feature is usually one of the first things to break down. The spider’s front eyes have muscles, like a human’s, that rotate the tube, but none that vary $f$, which is fixed. However, the retina consists of four separate layers at slightly different values of $d_i$. The figure only shows the detailed cellular structure of the rearmost layer, which is the most acute. Depending on $d_o$, the image may lie closest to any one of the four layers, and the spider can then use that layer to get a well-focused view. The layering is also believed to help eliminate problems caused by the variation of the index of refraction with wavelength (cf. problem 2, p. 902).

Although the spider’s eye is different in many ways from a human’s or an octopus’s, it shares the same fundamental construction, being essentially a lens that forms a real image on a screen inside a darkened chamber. From this perspective, the main difference is simply the scale, which is miniaturized by about a factor of $10^2$ in the linear dimensions. How far down can this scaling go? Does an
amoeba or a white blood cell lack an eye merely because it doesn’t have a nervous system that could make sense of the signals? In fact there is an optical limit on the miniaturization of any eye or camera. The spider’s eye is already so small that on the scale of the bottom panel in figure s, one wavelength of visible light would be easily distinguishable — about the length of the comma in this sentence. Chapter 32 is about optical effects that occur when the wave nature of light is important, and problem 14 on p. 928 specifically addresses the effect on this spider’s vision.