clearly made θ_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

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

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

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

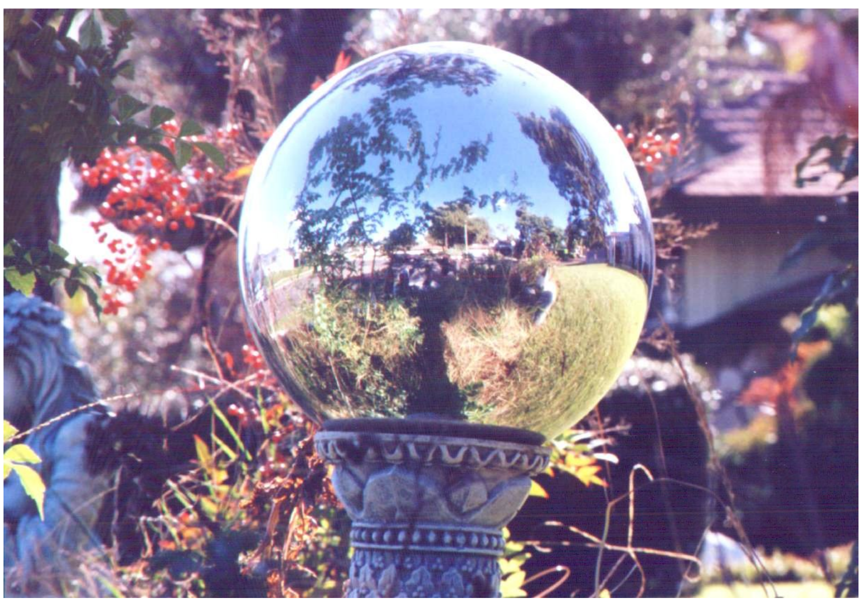
in order to give a positive result for the focal angle. The signs of the distance equation behave the same way:

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

Solving for d_i, we find

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

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



h / 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.

A shortcut for real images *example 6*

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.

3.3 ⋆ Aberrations

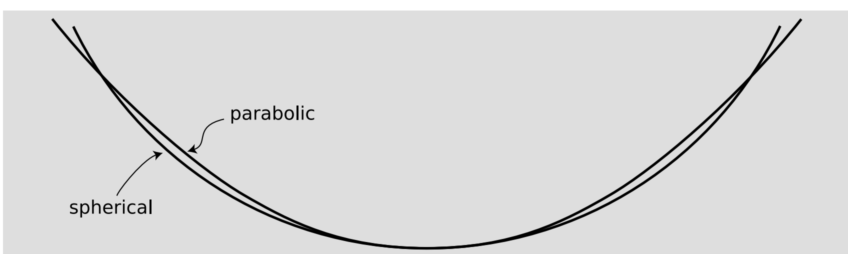
An imperfection or distortion in an image is called an aberration. An aberration can be produced by a flaw in a lens or mirror, but even with a perfect optical surface some degree of aberration is unavoidable. To see why, consider the mathematical approximation we've been making, which is that the depth 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

i / This photo was taken using a "fish-eye lens," which gives an extremely large field of view.

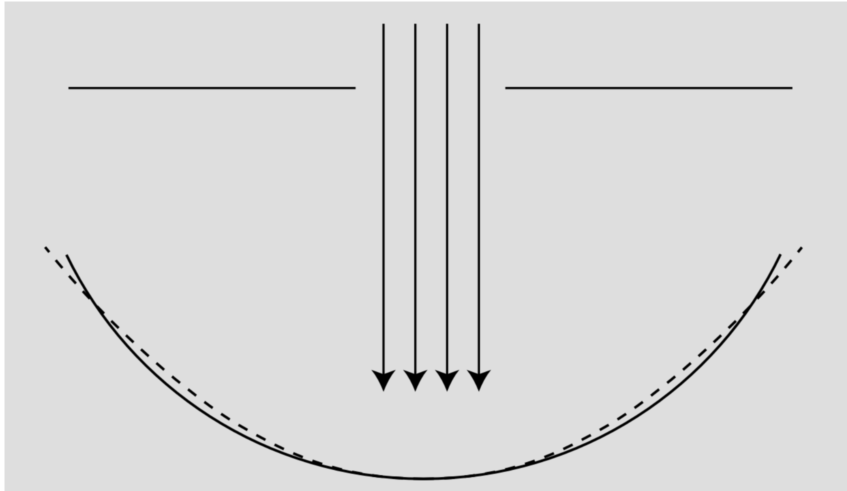


j / Spherical mirrors are the cheapest to make, but parabolic mirrors are better for making images of objects at infinity. A sphere has equal curvature everywhere, but a parabola has tighter curvature at its center and gentler curvature at the sides.

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.

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.

k / 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).



l / 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
θ_f	the focal angle, defined as $1/f$
θ_o	the object angle, defined as $1/d_o$
θ_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

$$\theta_f = \pm\theta_i \pm \theta_o$$
$$\frac{1}{f} = \pm\frac{1}{d_i} \pm \frac{1}{d_o} \qquad .$$

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 θ_i and θ_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. It is normally only physically possible for either θ_i or θ_o to be zero, not both. Imagine the case where that variable is zero. Since the left-hand side of the equation is positive by definition, the term on the right that we didn't eliminate must be the one that has a plus sign.

Once the correct form of the equation has been determined, the magnification can be found via the equation

$$M = \frac{d_i}{d_o} \qquad .$$

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

Problems

Key

- $\surd$ A computerized answer check is available online.
- $\int$ A problem that requires calculus.
- $\star$ A difficult problem.

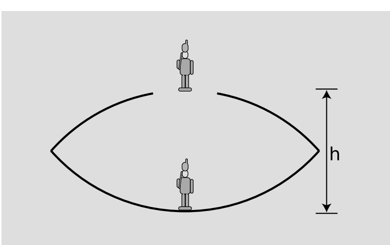
1 Apply the equation $M = d_i/d_o$ to the case of a flat mirror.

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. ▷ Solution, p. 108

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 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.

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.



Problem 5.

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

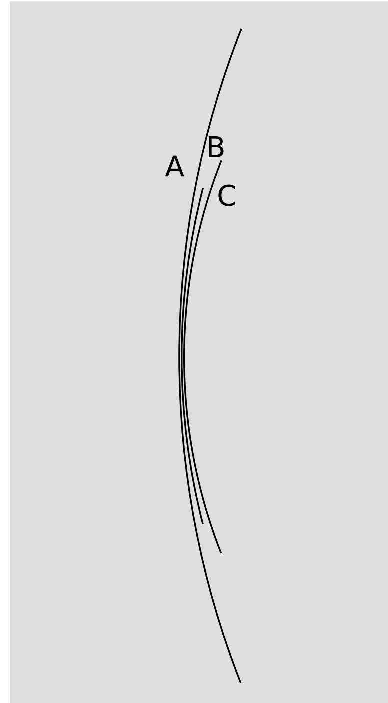
$$\frac{1}{f} = \frac{1}{h} + \frac{1}{h + \left(\frac{1}{h} - \frac{1}{f}\right)^{-1}} \qquad .$$

(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? $\star$

6 A hollowed-out surface that reflects sound waves can act just like an in-bending mirror. Suppose that, standing near such a surface, you are able to find 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?

7 Find the focal length of the mirror in problem 5 of chapter 1.

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

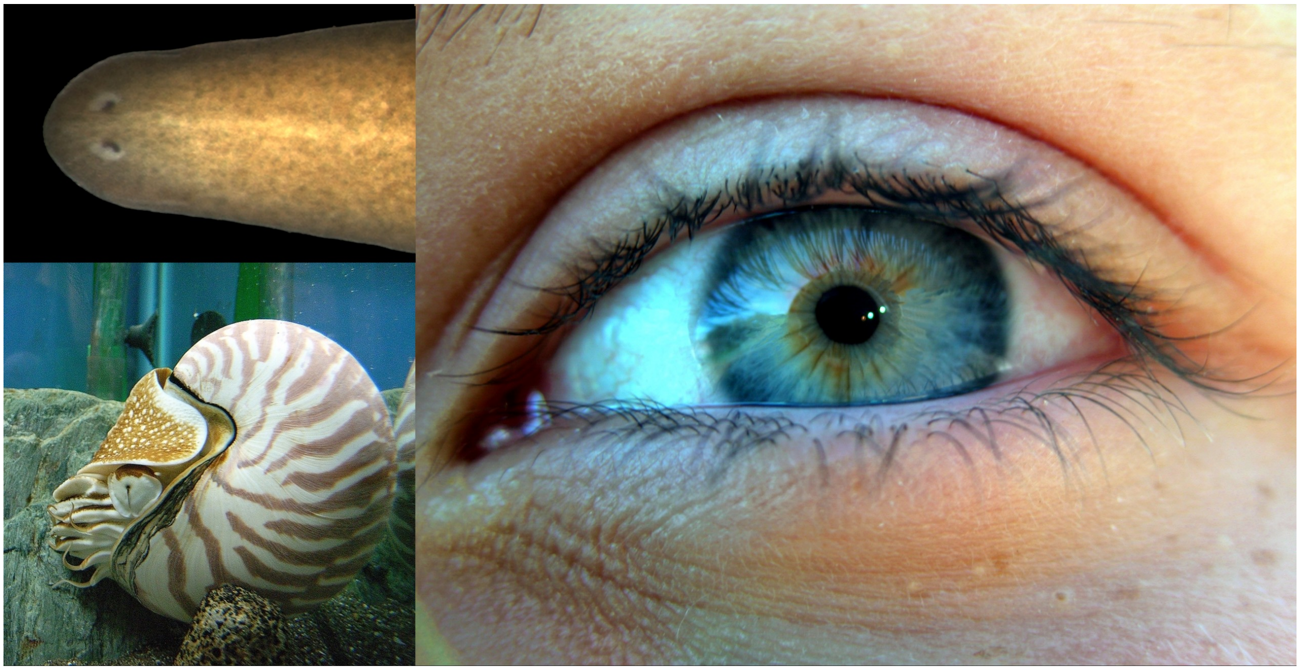


Problem 8.

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 $f = 20$ cm and $d_o = 30$ cm? (c) What if it was a *diverging* mirror with $f = 20$ cm and $d_o = 10$ cm? (d) A diverging mirror with $= 20$ cm and $d_o = 30$ cm?

▷ Solution, p. 108



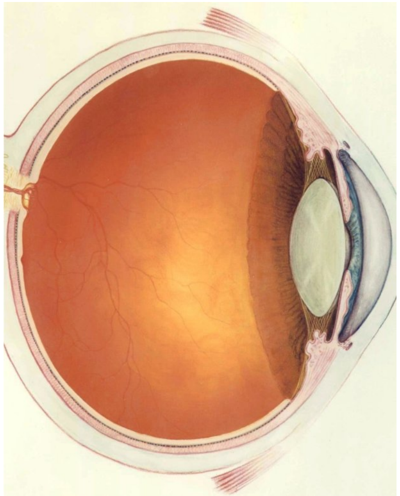
Three stages in the evolution of the eye. The flatworm has two eye pits. The nautilus's eyes are pinhole cameras. The human eye incorporates a lens.

Chapter 4

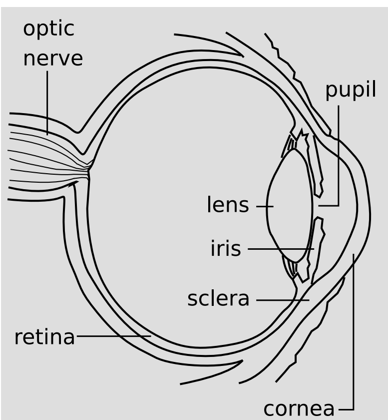
Refraction

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.

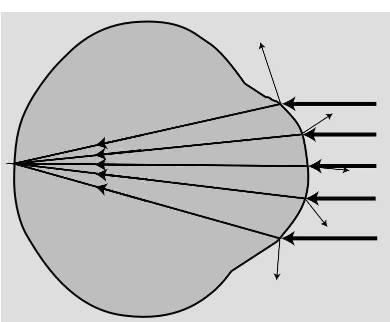
So great is the value attached to sight that some have imbued it with mystical aspects. Moses "had vision," George Bush did not. Christian fundamentalists who perceive a conflict between evolution and their religion have claimed that the eye is such a perfect device that it could never have arisen through 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 some-



a / A human eye.



b / The anatomy of the eye.



c / A simplified optical diagram of the eye. Light rays are bent when they cross from the air into the eye. (A little of the incident rays' energy goes into the reflected rays rather than the ones transmitted into the eye.)

where 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.

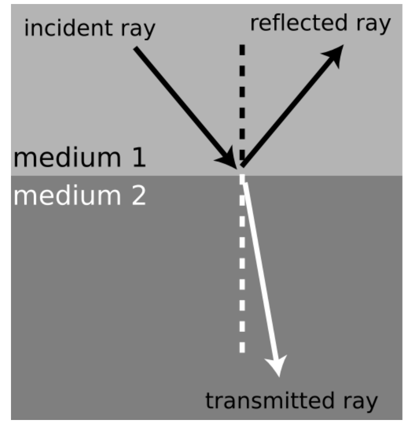
4.1 Refraction

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 again.

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.

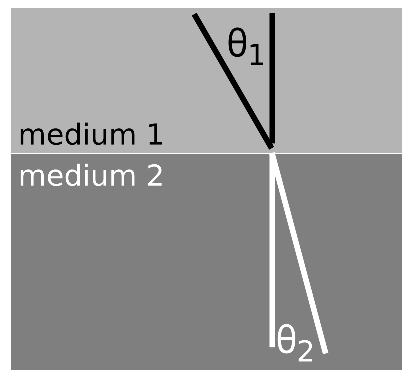


d / The incident, reflected, and transmitted (refracted) rays all lie in a plane that includes the normal (dashed line).

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.

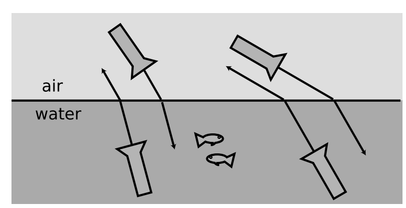
If you try shooting a beam of light at the boundary between two substances, say water and air, you'll find that regardless of the angle at which you send in the beam, the part of the beam in the water is always closer to the normal line, e. It doesn't matter if the ray is entering the water or leaving,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



e / The angles θ_1 and θ_2 are related to each other, and also depend on the properties of the two media. Because refraction is time-reversal symmetric, there is no need to label the rays with arrowheads.

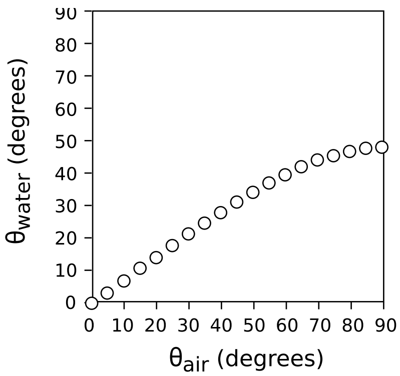
If, instead of water and air, you try another combination of substances, say plastic and gasoline, again you'll find that the ray's angle with respect to the normal is consistently smaller in one and larger in the other. Also, we find that if substance A has rays closer to normal than in B, and B has 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.



f / Refraction has time-reversal symmetry. Regardless of whether the light is going into or out of the water, the relationship between the two angles is the same, and the ray is closer to the normal while in the water.

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.



g / The relationship between the angles in refraction.

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

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

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

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

but rewriting it in the form

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

[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? ▷ Answer, p. 106

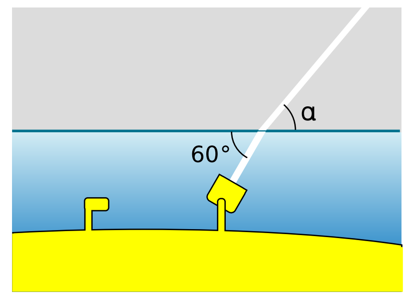
Finding an angle using Snell's law *example 1*

▷ A submarine shines its searchlight up toward the surface of the water. What is the angle α shown in the figure?

▷ 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 θ_1, we find

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

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 α is therefore 50°.



h / Example 1.

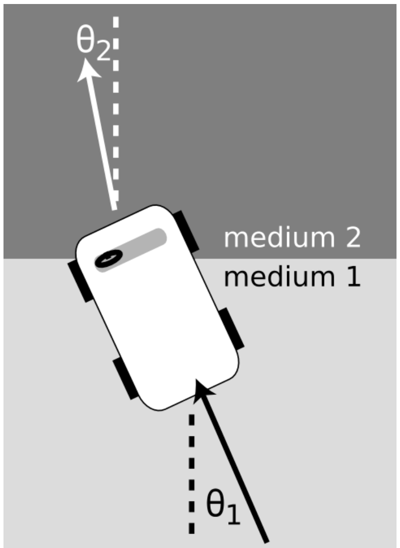
The index of refraction is related to the speed of light.

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

$$n = \frac{c}{v} \quad .$$

[n = medium's index of refraction, v = speed of light in that medium, c = 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.

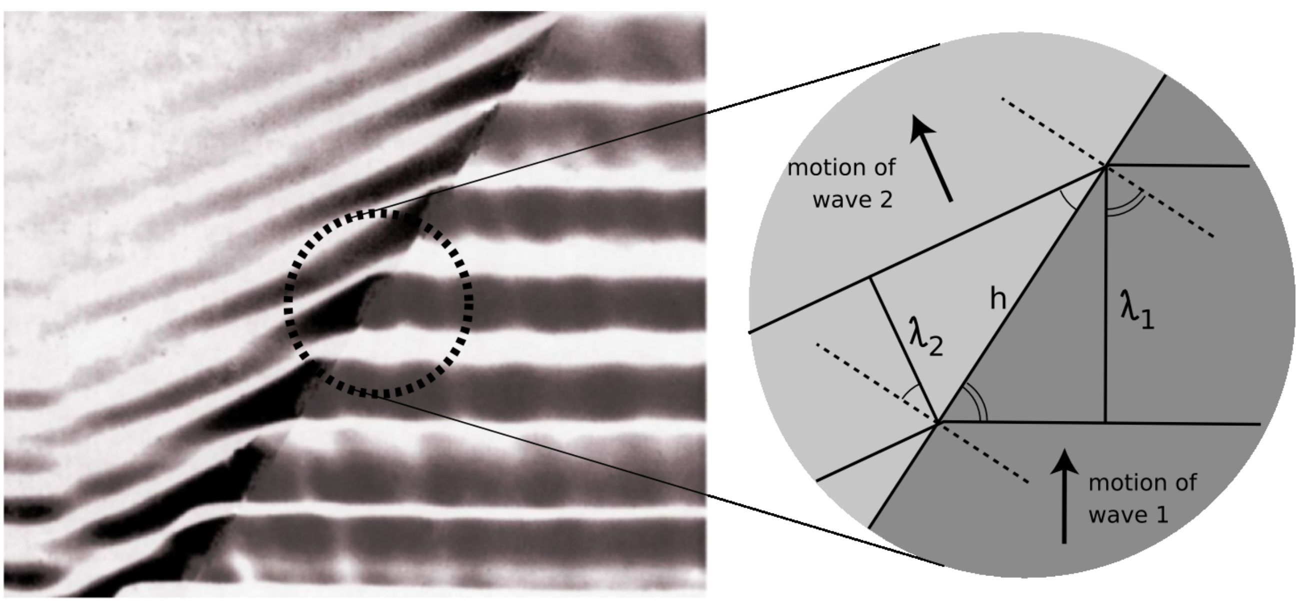


i / A mechanical model of refraction.

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 is 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.



j / A derivation of Snell's law.

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 θ_1, and the two on the left to θ_2.

Trigonometry gives

$$\sin\theta_1 = \lambda_1/h \qquad \text{and}$$
$$\sin\theta_2 = \lambda_2/h \qquad .$$

Eliminating h by dividing the equations, we find

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

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

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

which is one form of Snell's law.

Ocean waves near and far from shore *example 2*

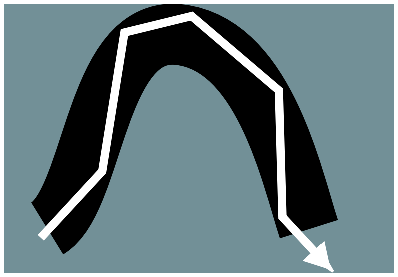
Ocean waves are formed by winds, typically on the open sea, and the wavefronts are perpendicular to the direction of the wind that formed them. At the beach, however, you have undoubtedly observed that waves tend come in with their wavefronts very nearly (but not exactly) parallel to the shoreline. This 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

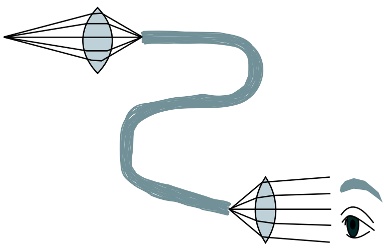
In general, the speed of light in a medium depends both on the medium and on the wavelength of the light. Another way of saying it is that a medium's index of refraction varies with wavelength. This is why a prism can be used to split up a beam of white light into a rainbow. Each wavelength of light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

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

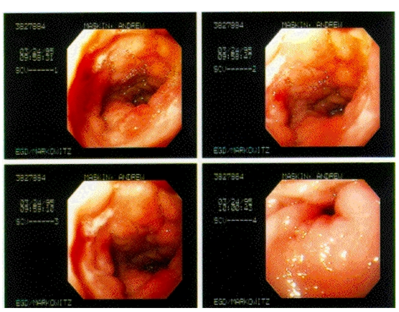
In book 3 we developed an equation for the percentage of the wave energy that is transmitted and the percentage reflected at a boundary between media. This was only done in the case of waves in one dimension, however, and rather than discuss the full three dimensional generalization it will be more useful 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,



k / Total internal reflection in a fiber-optic cable.



l / A simplified drawing of a surgical endoscope. The first lens forms a real image at one end of a bundle of optical fibers. The light is transmitted through the bundle, and is finally magnified by the eyepiece.



m / Endoscopic images of a duodenal ulcer.

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 θ_2, we find

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

and if n_1 is greater than n_2, then there will be large values of θ_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?

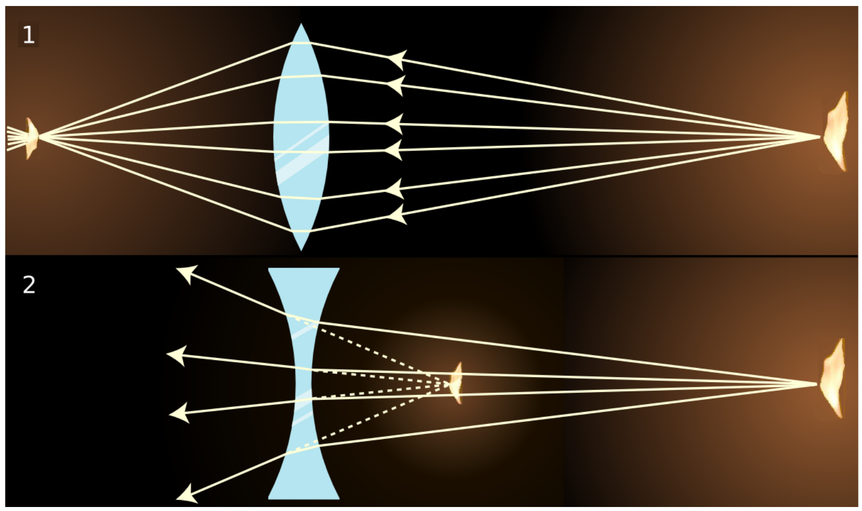
C A denser sample of a gas has a higher index of refraction than a less dense sample (i.e., a sample under lower pressure), but why would it not make sense for the index of refraction of a gas to be proportional to density?

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?

4.2 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.



n / 1. A converging lens forms an image of a candle flame. 2. A diverging lens.

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. 106

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.

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?

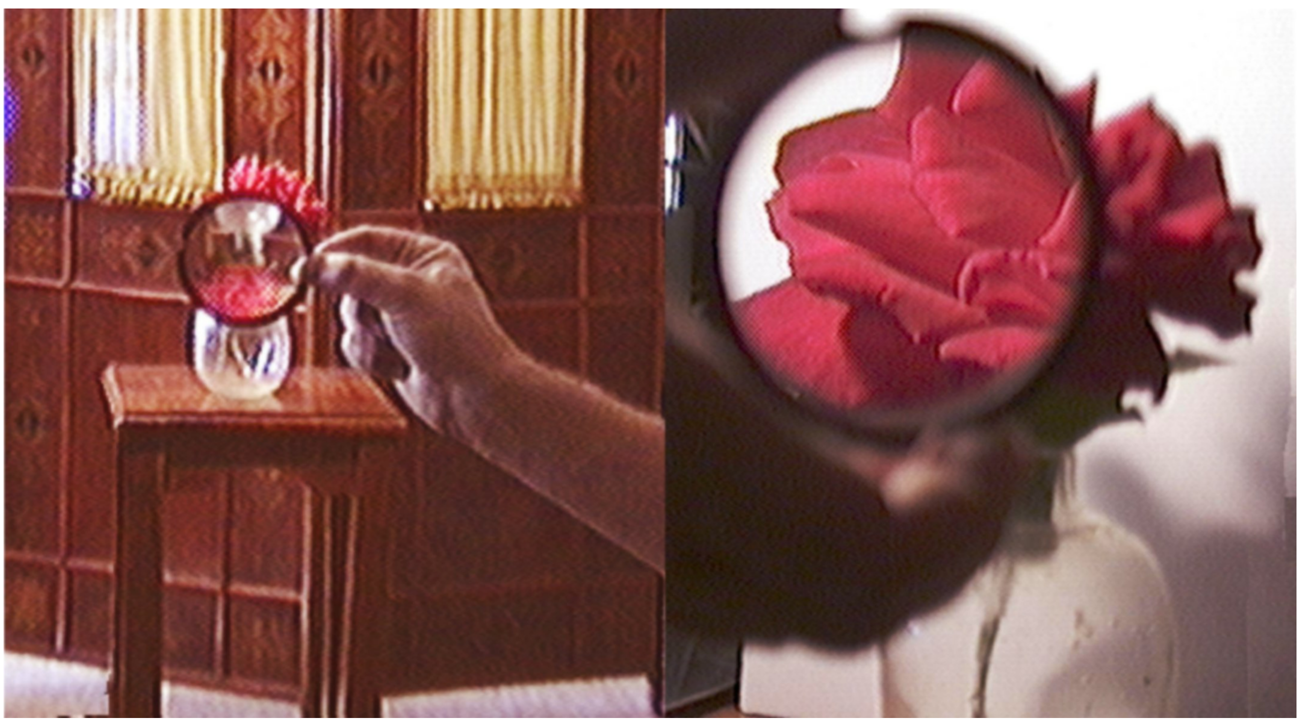
C There are three qualitatively different types of image formation that can occur with lenses, of which figures n/1 and n/2 exhaust only two. Figure out what the third possibility is. Which of the three possibilities can result in a magnification greater than one?

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

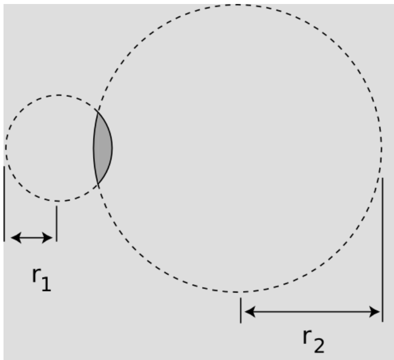
E In figures n/1 and n/2, the only rays drawn were those that happened to enter the lenses. Discuss this in relation to figure o.

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?

G In general, the index of refraction depends on the color of the light. What effect would this have on images formed by lenses?



o / Two images of a rose created by the same lens and recorded with the same camera.



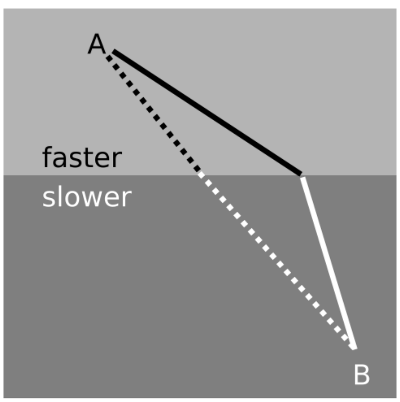
p / The radii of curvature appearing in the lensmaker's equation.

4.3 ⋆ 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

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

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 curved outward or both are curved inward; otherwise a negative sign applies. The proof of this equation is left as an exercise to those readers who are sufficiently brave and motivated.



q / The principle of least time applied to refraction.

4.4 ⋆ 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.

Summary

Selected Vocabulary

refraction	the change in direction that occurs when a wave encounters the interface between two media
index of refraction	an optical property of matter; the speed of light in a vacuum divided by the speed of light in the substance in question

Notation

n	the index of refraction

Summary

Refraction is a change in direction that occurs when a wave encounters the interface between two media. Together, refraction and reflection account for the basic principles behind nearly all optical devices.

Snell discovered the equation for refraction,

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \qquad ,$$

[angles measured with respect to the normal]

through experiments with light rays, long before light was proven to be a wave. Snell's law can be proven based on the geometrical behavior of waves. Here n is the index of refraction. Snell invented this quantity to describe the refractive properties of various substances, but it was later found to be related to the speed of light in the substance,

$$n = \frac{c}{v} \qquad ,$$

where c is the speed of light in a vacuum. In general a material's index of refraction is different for different wavelengths of light.

As discussed in the third book of this series, any wave is partially transmitted and partially reflected at the boundary between two media in which its speeds are different. It is not particularly important to know the equation that tells what fraction is transmitted (and thus refracted), but important technologies such as fiber optics are based on the fact that this fraction becomes *zero* for sufficiently oblique angles. This phenomenon is referred to as total internal reflection. It occurs when there is no angle that satisfies Snell's law.

Problems

Key

- $\checkmark$ A computerized answer check is available online.
- $\int$ A problem that requires calculus.
- $\star$ A difficult problem.

1 Suppose a converging lens is constructed of a type of plastic whose index of refraction is less than that of water. How will the lens's behavior be different if it is placed underwater?

2 There are two main types of telescopes, refracting (using lenses) and reflecting (using mirrors). (Some telescopes use a mixture of the two types of elements: the light first encounters a large curved mirror, and then goes through an eyepiece that is a lens.) What implications would the color-dependence of focal length have for the relative merits of the two types of telescopes? What would happen with white starlight, for example?

3 Based on Snell's law, explain why rays of light passing through the edges of a converging lens are bent more than rays passing through parts closer to the center. It might seem like it should be the other way around, since the rays at the edge pass through less glass — shouldn't they be affected less? In your answer:

- Include a ray diagram showing a huge close-up view of the relevant part of the lens.
- Make use of the fact that the front and back surfaces aren't always parallel; a lens in which the front and back surfaces *are* always parallel doesn't focus light at all, so if your explanation doesn't make use of this fact, your argument must be incorrect.
- Make sure your argument still works even if the rays don't come in parallel to the axis.

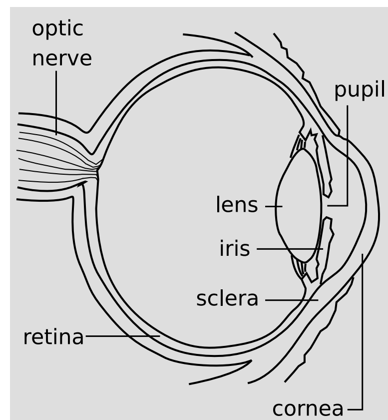
4 When you take pictures with a camera, the distance between the lens and the film has to be adjusted, depending on the distance at which you want to focus. This is done by moving the lens. If you want to change your focus so that you can take a picture of something farther away, which way do you have to move the lens? Explain using ray diagrams. [Based on a problem by Eric Mazur.]

5 (a) Light is being reflected diffusely from an object 1.000 m under water. The light that comes up to the surface is refracted at the water-air interface. If the refracted rays all appear to come from the same point, then there will be a virtual image of the object in the water, above the object's actual position, which will be visible to an observer above the water. Consider three rays, A, B and C, whose angles in the water with respect to the normal are $\theta_i = 0.000\,°$, $1.000\,°$and $20.000\,°$respectively. Find the depth of the point at which the refracted parts of A and B appear to have intersected, and do the same for A and C. Show that the intersections are at nearly the same depth, but not quite. [Check: The difference in depth should be about 4 cm.]

(b) Since all the refracted rays do not quite appear to have come from the same point, this is technically not a virtual image. In practical terms, what effect would this have on what you see?

(c) In the case where the angles are all small, use algebra and trig to show that the refracted rays do appear to come from the same point, and find an equation for the depth of the virtual image. Do not put in any numerical values for the angles or for the indices of refraction — just keep them as symbols. You will need the approximation $\sin\theta \approx \tan\theta \approx \theta$, which is valid for small angles measured in radians. ⋆

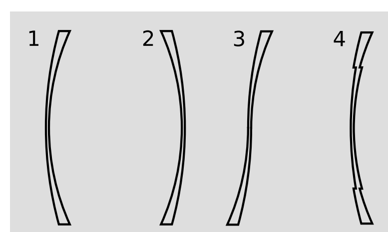
6 The drawing shows the anatomy of the human eye, at twice life size. Find the radius of curvature of the outer surface of the cornea by measurements on the figure, and then derive the focal length of the air-cornea interface, where almost all the focusing of light occurs. You will need to use physical reasoning to modify the lensmaker's equation for the case where there is only a single refracting surface. Assume that the index of refraction of the cornea is essentially that of water. ⋆



Problem 6.

7 When swimming underwater, why is your vision made much clearer by wearing goggles with flat pieces of glass that trap air behind them? [Hint: You can simplify your reasoning by considering the special case where you are looking at an object far away, and along the optic axis of the eye.]

8 The figure shows four lenses. Lens 1 has two spherical surfaces. Lens 2 is the same as lens 1 but turned around. Lens 3 is made by cutting through lens 1 and turning the bottom around. Lens 4 is made by cutting a central circle out of lens 1 and recessing it.



Problem 8.

(a) A parallel beam of light enters lens 1 from the left, parallel to its axis. Reasoning based on Snell's law, will the beam emerging from the lens be bent inward or outward, or will it remain parallel to the axis? Explain your reasoning. As part of your answer, make an huge drawing of one small part of the lens, and apply Snell's law at both interfaces. Recall that rays are bent more if they come to the interface at a larger angle with respect to the normal.

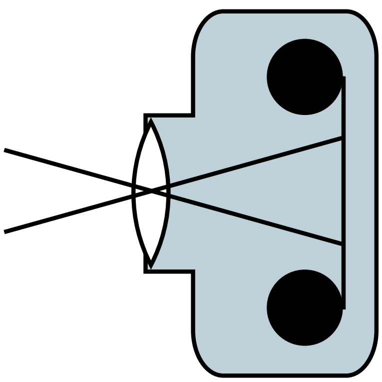
(b) What will happen with lenses 2, 3, and 4? Explain. Drawings are not necessary.

9 Prove that the principle of least time leads to Snell's law. ⋆

10 An object is more than one focal length from a converging lens. (a) Draw a ray diagram. (b) Using reasoning like that developed in chapter 3, determine the positive and negative signs in the equation $1/f = \pm 1/d_i \pm 1/d_o$. (c) The images of the rose in section 4.2 were made using a lens with a focal length of 23 cm. If the lens is placed 80 cm from the rose, locate the image. $\surd$

11 An object is less than one focal length from a converging lens. (a) Draw a ray diagram. (b) Using reasoning like that developed in chapter 3, determine the positive and negative signs in the equation $1/f = \pm 1/d_i \pm 1/d_o$. (c) The images of the rose in section 4.2 were made using a lens with a focal length of 23 cm. If the lens is placed 10 cm from the rose, locate the image. $\surd$

12 Nearsighted people wear glasses whose lenses are diverging. (a) Draw a ray diagram. For simplicity pretend that there is no eye behind the glasses. (b) Using reasoning like that developed in chapter 3, determine the positive and negative signs in the equation $1/f = \pm 1/d_i \pm 1/d_o$. (c) If the focal length of the lens is 50.0 cm, and the person is looking at an object at a distance of 80.0 cm, locate the image. $\surd$



Problem 13.

13 Two standard focal lengths for camera lenses are 50 mm (standard) and 28 mm (wide-angle). To see how the focal lengths relate to the angular size of the field of view, it is helpful to visualize things as represented in the figure. Instead of showing many rays coming from the same point on the same object, as we normally do, the figure shows two rays from two different objects. Although the lens will intercept infinitely many rays from each of these points, we have shown only the ones that pass through the center of the lens, so that they suffer no angular deflection. (Any angular deflection at the front surface of the lens is canceled by an opposite deflection at the back, since the front and back surfaces are parallel at the lens's center.) What is special about these two rays is that they are aimed at the edges of one 35-mm-wide frame of film; that is, they show the limits of the field of view. Throughout this problem, we assume that d_o is much greater than d_i. (a) Compute the angular width of the camera's field of view when these two lenses are used. (b) Use small-angle approximations to find a simplified equation for the angular width of the field of view, θ, in terms of the focal length, f, and the width of the film, w. Your equation should not have any trig functions in it. Compare the results of this approximation with your answers from part a. (c) Suppose that we are holding constant the aperture (amount of surface area of the lens being used to collect light). When switching from a 50-mm lens to a 28-mm lens, how many times longer or shorter must the exposure be in order to make a properly developed picture, i.e., one that is not under- or overexposed? [Based on a problem by Arnold Arons.]

▷ Solution, p. 108

14 A nearsighted person is one whose eyes focus light too strongly, and who is therefore unable to relax the lens inside her eye sufficiently to form an image on her retina of an object that is too far away.

(a) Draw a ray diagram showing what happens when the person tries, with uncorrected vision, to focus at infinity.

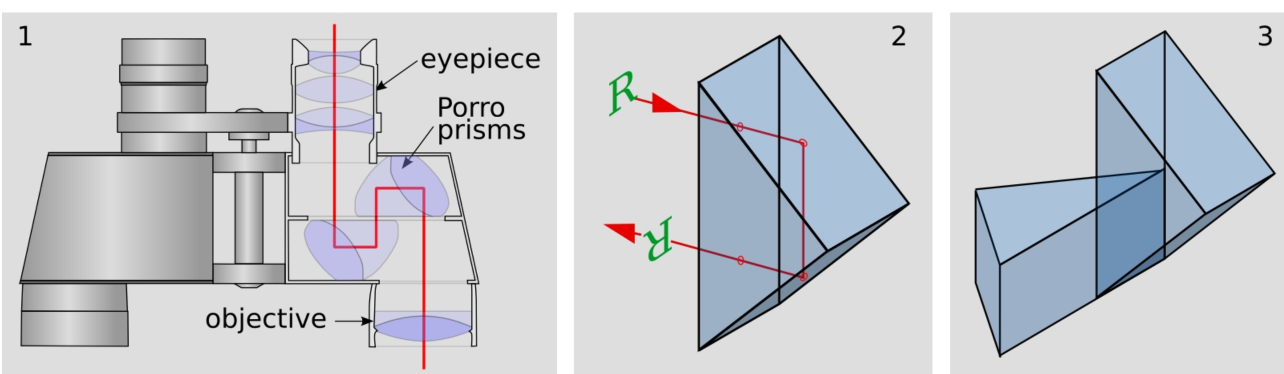
(b) What type of lenses do her glasses have? Explain.

(c) Draw a ray diagram showing what happens when she wears glasses. Locate both the image formed by the glasses and the final image.

(d) Suppose she sometimes uses contact lenses instead of her glasses. Does the focal length of her contacts have to be less than, equal to, or greater than that of her glasses? Explain.

15 Diamond has an index of refraction of 2.42, and part of the reason diamonds sparkle is that this encourages a light ray to undergo many total internal reflections before it emerges. Calculate the critical angle at which total internal reflection occurs in diamond. Explain the interpretation of your result: Is it measured from the normal, or from the surface? Is it a minimum, or a maximum? How would the critical angle have been different for a substance such as glass or plastic, with a lower index of refraction?

16 Fred's eyes are able to focus on things as close as 5.0 cm. Fred holds a magnifying glass with a focal length of 3.0 cm at a height of 2.0 cm above a flatworm. (a) Locate the image, and find the magnification. (b) Without the magnifying glass, from what distance would Fred want to view the flatworm to see its details as well as possible? With the magnifying glass? (c) Compute the angular magnification.



Problem 17.

17 Panel 1 of the figure shows the optics inside a pair of binoculars. They are essentially a pair of telescopes, one for each eye.

But to make them more compact, and allow the eyepieces to be the right distance apart for a human face, they incorporate a set of eight prisms, which fold the light path. In addition, the prisms make the image upright. Panel 2 shows one of these prisms, known as a Porro prism. The light enters along a normal, undergoes two total internal reflections at angles of 45 degrees with respect to the back surfaces, and exits along a normal. The image of the letter R has been flipped across the horizontal. Panel 3 shows a pair of these prisms glued together. The image will be flipped across both the horizontal and the vertical, which makes it oriented the right way for the user of the binoculars.
(a) Find the minimum possible index of refraction for the glass used in the prisms.
(b) For a material of this minimal index of refraction, find the fraction of the incoming light that will be lost to reflection in the four Porro prisms on a each side of a pair of binoculars. (See chapter 4 or *Vibrations and Waves*, or section 6.2 of *Simple Nature*.) In real, high-quality binoculars, the optical surfaces of the prisms have antireflective coatings, but carry out your calculation for the case where there is no such coating.
(c) Discuss the reasons why a designer of binoculars might or might not want to use a material with exactly the index of refraction found in part a. ⋆

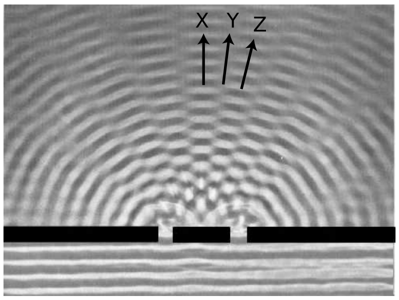
This image of the Pleiades star cluster shows haloes around the stars due to the wave nature of light.

Chapter 5
Wave Optics

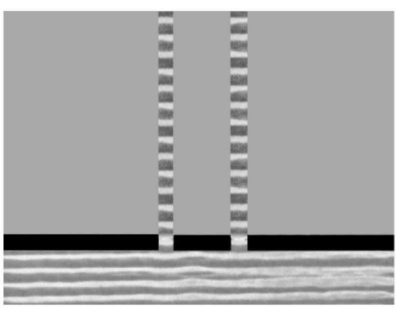
Electron microscopes can make images of individual atoms, but why will a visible-light microscope never be able to? Stereo speakers create the illusion of music that comes from a band arranged in your living room, but why doesn't the stereo illusion work with bass notes? Why are computer chip manufacturers investing billions of dollars in equipment to etch chips with x-rays instead of visible light?

The answers to all of these questions have to do with the subject of wave optics. So far this book has discussed the interaction of light waves with matter, and its practical applications to optical devices like mirrors, but we have used the ray model of light almost exclusively. Hardly ever have we explicitly made use of the fact that light is an electromagnetic wave. We were able to get away with the simple ray model because the chunks of matter we were discussing, such as lenses and mirrors, were thousands of times larger than a wavelength of light. We now turn to phenomena and devices that can only be understood using the wave model of light.

5.1 Diffraction



a / In this view from overhead, a straight, sinusoidal water wave encounters a barrier with two gaps in it. Strong wave vibration occurs at angles X and Z, but there is none at all at angle Y. (The figure has been retouched from a real photo of water waves. In reality, the waves beyond the barrier would be much weaker than the ones before it, and they would therefore be difficult to see.)



b / This doesn't happen.

Figure a shows a typical problem in wave optics, enacted with water waves. It may seem surprising that we don't get a simple pattern like figure b, but the pattern would only be that simple if the wavelength was hundreds of times shorter than the distance between the gaps in the barrier and the widths of the gaps.

Wave optics is a broad subject, but this example will help us to pick out a reasonable set of restrictions to make things more manageable:

(1) We restrict ourselves to cases in which a wave travels through a uniform medium, encounters a certain area in which the medium has different properties, and then emerges on the other side into a second uniform region.

(2) We assume that the incoming wave is a nice tidy sine-wave pattern with wavefronts that are lines (or, in three dimensions, planes).

(3) In figure a we can see that the wave pattern immediately beyond the barrier is rather complex, but farther on it sorts itself out into a set of wedges separated by gaps in which the water is still. We will restrict ourselves to studying the simpler wave patterns that occur farther away, so that the main question of interest is how intense the outgoing wave is at a given angle.

The kind of phenomenon described by restriction (1) is called *diffraction*. Diffraction can be defined as the behavior of a wave when it encounters an obstacle or a nonuniformity in its medium. In general, diffraction causes a wave to bend around obstacles and make patterns of strong and weak waves radiating out beyond the obstacle. Understanding diffraction is the central problem of wave optics. If you understand diffraction, even the subset of diffraction problems that fall within restrictions (2) and (3), the rest of wave optics is icing on the cake.

Diffraction can be used to find the structure of an unknown diffracting object: even if the object is too small to study with ordinary imaging, it may be possible to work backward from the diffraction pattern to learn about the object. The structure of a crystal, for example, can be determined from its x-ray diffraction pattern.

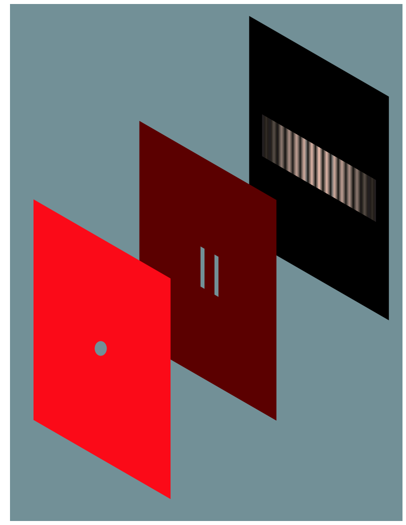
Diffraction can also be a bad thing. In a telescope, for example, light waves are diffracted by all the parts of the instrument. This will cause the image of a star to appear fuzzy even when the focus has been adjusted correctly. By understanding diffraction, one can learn how a telescope must be designed in order to reduce this problem — essentially, it should have the biggest possible diameter.

There are two ways in which restriction (2) might commonly be violated. First, the light might be a mixture of wavelengths. If we simply want to observe a diffraction pattern or to use diffraction as a technique for studying the object doing the diffracting (e.g., if the object is too small to see with a microscope), then we can pass the light through a colored filter before diffracting it.

A second issue is that light from sources such as the sun or a lightbulb does not consist of a nice neat plane wave, except over very small regions of space. Different parts of the wave are out of step with each other, and the wave is referred to as *incoherent*. One way of dealing with this is shown in figure c. After filtering to select a certain wavelength of red light, we pass the light through a small pinhole. The region of the light that is intercepted by the pinhole is so small that one part of it is not out of step with another. Beyond the pinhole, light spreads out in a spherical wave; this is analogous to what happens when you speak into one end of a paper towel roll and the sound waves spread out in all directions from the other end. By the time the spherical wave gets to the double slit it has spread out and reduced its curvature, so that we can now think of it as a simple plane wave.

If this seems laborious, you may be relieved to know that modern technology gives us an easier way to produce a single-wavelength, coherent beam of light: the laser.

The parts of the final image on the screen in c are called diffraction fringes. The center of each fringe is a point of maximum brightness, and halfway between two fringes is a minimum.



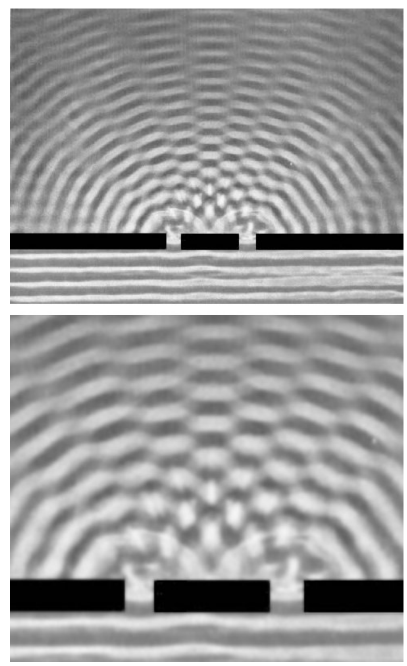
c / A practical, low-tech setup for observing diffraction of light.

Discussion Question

A Why would x-rays rather than visible light be used to find the structure of a crystal? Sound waves are used to make images of fetuses in the womb. What would influence the choice of wavelength?

5.2 Scaling of Diffraction

This chapter has "optics" in its title, so it is nominally about light, but we started out with an example involving water waves. Water waves are certainly easier to visualize, but is this a legitimate comparison? In fact the analogy works quite well, despite the fact that a light wave has a wavelength about a million times shorter. This is because diffraction effects scale uniformly. That is, if we enlarge or reduce the whole diffraction situation by the same factor, including both the wavelengths and the sizes of the obstacles the wave encounters, the result is still a valid solution.



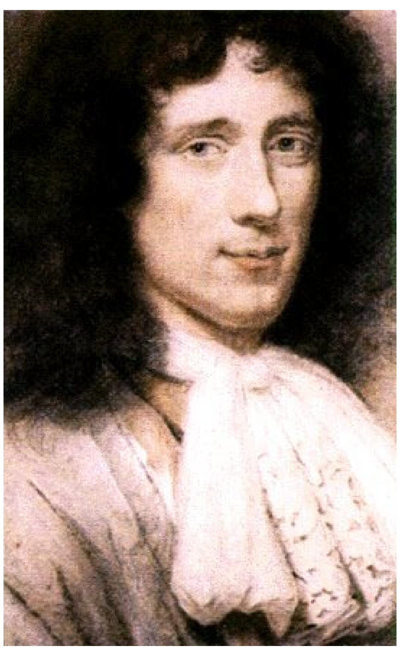
d / The bottom figure is simply a copy of the middle portion of the top one, scaled up by a factor of two. All the angles are the same. Physically, the angular pattern of the diffraction fringes can't be any different if we scale both λ and d by the same factor, leaving λ/d unchanged.

This is unusually simple behavior! In the first book of this series we saw many examples of more complex scaling, such as the impossibility of bacteria the size of dogs, or the need for an elephant to eliminate heat through its ears because of its small surface-to-volume ratio, whereas a tiny shrew's life-style centers around conserving its body heat.

Of course water waves and light waves differ in many ways, not just in scale, but the general facts you will learn about diffraction are applicable to all waves. In some ways it might have been more appropriate to insert this chapter at the end of book 3, Vibrations and Waves, but many of the important applications are to light waves, and you would probably have found these much more difficult without any background in optics.

Another way of stating the simple scaling behavior of diffraction is that the diffraction angles we get depend only on the unitless ratio λ/d, where λ is the wavelength of the wave and d is some dimension of the diffracting objects, e.g., the center-to-center spacing between the slits in figure a. If, for instance, we scale up both λ and d by a factor of 37, the ratio λ/d will be unchanged.

5.3 The Correspondence Principle



e / Christiaan Huygens (1629-1695).

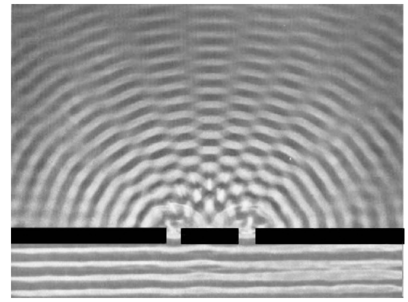
The only reason we don't usually notice diffraction of light in everyday life is that we don't normally deal with objects that are comparable in size to a wavelength of visible light, which is about a millionth of a meter. Does this mean that wave optics contradicts ray optics, or that wave optics sometimes gives wrong results? No. If you hold three fingers out in the sunlight and cast a shadow with them, *either* wave optics or ray optics can be used to predict the straightforward result: a shadow pattern with two bright lines where the light has gone through the gaps between your fingers. Wave optics is a more general theory than ray optics, so in any case where ray optics is valid, the two theories will agree. This is an example of a general idea enunciated by the physicist Niels Bohr, called the *correspondence principle:* when flaws in a physical theory lead to the creation of a new and more general theory, the new theory must still agree with the old theory within its more restricted area of applicability. After all, a theory is only created as a way of describing experimental observations. If the original theory had not worked in any cases at all, it would never have become accepted.

In the case of optics, the correspondence principle tells us that when λ/d is small, both the ray and the wave model of light must give approximately the same result. Suppose you spread your fingers and cast a shadow with them using a coherent light source. The quantity λ/d is about 10 -4, so the two models will agree very closely. (To be specific, the shadows of your fingers will be outlined by a series of light and dark fringes, but the angle subtended by a fringe will be on the order of 10^{-4} radians, so they will be invisible and washed out by the natural fuzziness of the edges of sun-shadows, caused by the finite size of the sun.)

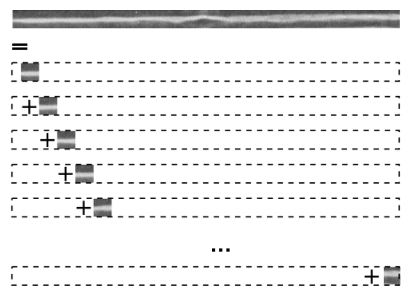
self-check A

What kind of wavelength would an electromagnetic wave have to have in order to diffract dramatically around your body? Does this contradict the correspondence principle? ▷ Answer, p. 106

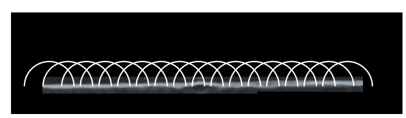
5.4 Huygens' Principle



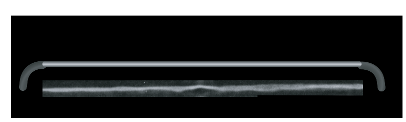
f / Double-slit diffraction.



g / A wavefront can be analyzed by the principle of superposition, breaking it down into many small parts.



h / If it was by itself, each of the parts would spread out as a circular ripple.



i / Adding up the ripples produces a new wavefront.

Returning to the example of double-slit diffraction, f, note the strong visual impression of two overlapping sets of concentric semicircles. This is an example of *Huygens' principle*, named after a Dutch physicist and astronomer. (The first syllable rhymes with "boy.") Huygens' principle states that any wavefront can be broken down into many small side-by-side wave peaks, g, which then spread out as circular ripples, h, and by the principle of superposition, the result of adding up these sets of ripples must give the same result as allowing the wave to propagate forward, i. In the case of sound or light waves, which propagate in three dimensions, the "ripples" are actually spherical rather than circular, but we can often imagine things in two dimensions for simplicity.

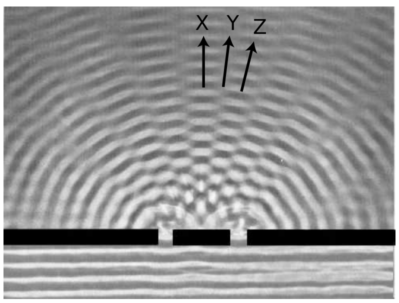
In double-slit diffraction the application of Huygens' principle is visually convincing: it is as though all the sets of ripples have been blocked except for two. It is a rather surprising mathematical fact, however, that Huygens' principle gives the right result in the case of an unobstructed linear wave, h and i. A theoretically infinite number of circular wave patterns somehow conspire to add together and produce the simple linear wave motion with which we are familiar.

Since Huygens' principle is equivalent to the principle of superposition, and superposition is a property of waves, what Huygens had created was essentially the first wave theory of light. However, he imagined light as a series of pulses, like hand claps, rather than as a sinusoidal wave.

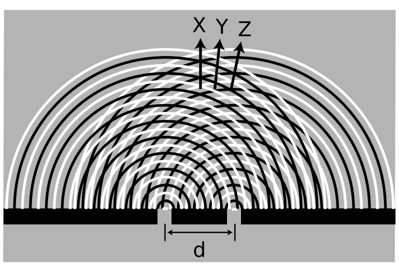
The history is interesting. Isaac Newton loved the atomic theory of matter so much that he searched enthusiastically for evidence that light was also made of tiny particles. The paths of his light particles would correspond to rays in our description; the only significant difference between a ray model and a particle model of light would occur if one could isolate individual particles and show that light had a "graininess" to it. Newton never did this, so although he thought of his model as a particle model, it is more accurate to say he was one of the builders of the ray model.

Almost all that was known about reflection and refraction of light could be interpreted equally well in terms of a particle model or a wave model, but Newton had one reason for strongly opposing

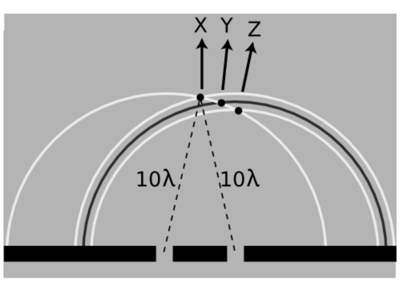
j / Thomas Young



k / Double-slit diffraction.



l / Use of Huygens' principle.



m / Constructive interference along the center-line.

Huygens' wave theory. Newton knew that waves exhibited diffraction, but diffraction of light is difficult to observe, so Newton believed that light did not exhibit diffraction, and therefore must not be a wave. Although Newton's criticisms were fair enough, the debate also took on the overtones of a nationalistic dispute between England and continental Europe, fueled by English resentment over Leibniz's supposed plagiarism of Newton's calculus. Newton wrote a book on optics, and his prestige and political prominence tended to discourage questioning of his model.

Thomas Young (1773-1829) was the person who finally, a hundred years later, did a careful search for wave interference effects with light and analyzed the results correctly. He observed double-slit diffraction of light as well as a variety of other diffraction effects, all of which showed that light exhibited wave interference effects, and that the wavelengths of visible light waves were extremely short. The crowning achievement was the demonstration by the experimentalist Heinrich Hertz and the theorist James Clerk Maxwell that light was an *electromagnetic* wave. Maxwell is said to have related his discovery to his wife one starry evening and told her that she was the only person in the world who knew what starlight was.

5.5 Double-Slit Diffraction

Let's now analyze double-slit diffraction, k, using Huygens' principle. The most interesting question is how to compute the angles such as X and Z where the wave intensity is at a maximum, and the in-between angles like Y where it is minimized. Let's measure all our angles with respect to the vertical center line of the figure, which was the original direction of propagation of the wave.

If we assume that the width of the slits is small (on the order of the wavelength of the wave or less), then we can imagine only a single set of Huygens ripples spreading out from each one, l. White lines represent peaks, black ones troughs. The only dimension of the diffracting slits that has any effect on the geometric pattern of the overlapping ripples is then the center-to-center distance, d, between the slits.

We know from our discussion of the scaling of diffraction that there must be some equation that relates an angle like θ_Z to the ratio λ/d,

$$\frac{\lambda}{d} \leftrightarrow \theta_Z \qquad .$$

If the equation for θ_Z depended on some other expression such as $\lambda + d$ or λ^2/d, then it would change when we scaled λ and d by the same factor, which would violate what we know about the scaling of diffraction.

Along the central maximum line, X, we always have positive waves coinciding with positive ones and negative waves coinciding with negative ones. (I have arbitrarily chosen to take a snapshot of the pattern at a moment when the waves emerging from the slit are experiencing a positive peak.) The superposition of the two sets of ripples therefore results in a doubling of the wave amplitude along this line. There is constructive interference. This is easy to explain, because by symmetry, each wave has had to travel an equal number of wavelengths to get from its slit to the center line, m: Because both sets of ripples have ten wavelengths to cover in order to reach the point along direction X, they will be in step when they get there.

At the point along direction Y shown in the same figure, one wave has traveled ten wavelengths, and is therefore at a positive extreme, but the other has traveled only nine and a half wavelengths, so it at a negative extreme. There is perfect cancellation, so points along this line experience no wave motion.

But the distance traveled does not have to be equal in order to get constructive interference. At the point along direction Z, one wave has gone nine wavelengths and the other ten. They are both at a positive extreme.

self-check B

At a point half a wavelength below the point marked along direction X, carry out a similar analysis. ▷ Answer, p. 107

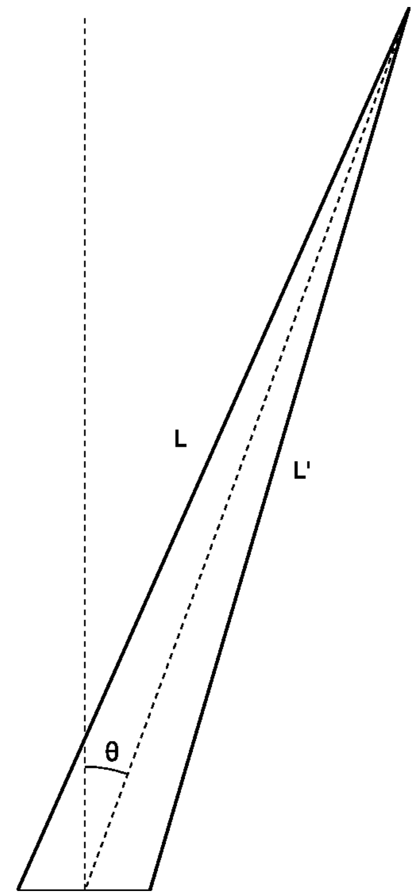
To summarize, we will have perfect constructive interference at any point where the distance to one slit differs from the distance to the other slit by an integer number of wavelengths. Perfect destructive interference will occur when the number of wavelengths of path length difference equals an integer plus a half.

Now we are ready to find the equation that predicts the angles of the maxima and minima. The waves travel different distances to get to the same point in space, n. We need to find whether the waves are in phase (in step) or out of phase at this point in order to predict whether there will be constructive interference, destructive interference, or something in between.

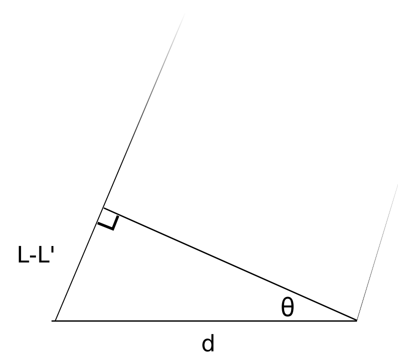
One of our basic assumptions in this chapter is that we will only be dealing with the diffracted wave in regions very far away from the object that diffracts it, so the triangle is long and skinny. Most real-world examples with diffraction of light, in fact, would have triangles with even skinner proportions than this one. The two long sides are therefore very nearly parallel, and we are justified in drawing the right triangle shown in figure o, labeling one leg of the right triangle as the difference in path length , $L-L'$, and labeling the acute angle as θ. (In reality this angle is a tiny bit greater than the one labeled θ in figure n.)

The difference in path length is related to d and θ by the equation

$$\frac{L-L'}{d}=\sin\theta \qquad .$$



n / The waves travel distances L_1 and L_2 from the two slits to get to the same point in space, at an angle θ from the center line.



o / A close-up view of figure n, showing how the path length difference $L-L'$ is related to d and to the angle θ.

Constructive interference will result in a maximum at angles for which $L - L'$ is an integer number of wavelengths,

$$L - L' = m\lambda \qquad . \qquad \text{[condition for a maximum; } m \text{ is an integer]}$$

Here m equals 0 for the central maximum, -1 for the first maximum to its left, $+2$ for the second maximum on the right, etc. Putting all the ingredients together, we find $m\lambda/d = \sin\theta$, or

$$\frac{\lambda}{d} = \frac{\sin\theta}{m} \qquad . \qquad \text{[condition for a maximum; } m \text{ is an integer]}$$

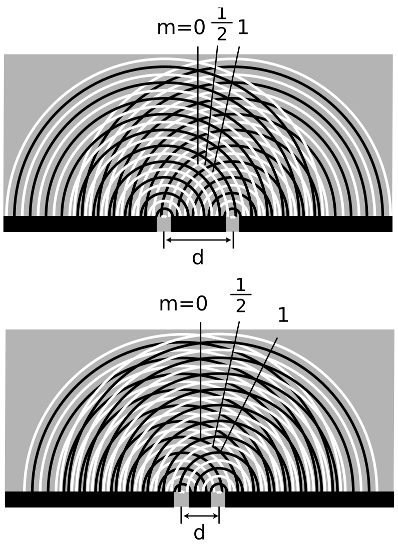
Similarly, the condition for a minimum is

$$\frac{\lambda}{d} = \frac{\sin\theta}{m} \qquad . \qquad \text{[condition for a minimum; } m \text{ is an integer plus } 1/2\text{]}$$

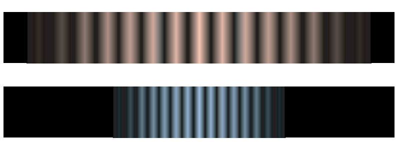
That is, the minima are about halfway between the maxima.

As expected based on scaling, this equation relates angles to the unitless ratio λ/d. Alternatively, we could say that we have proven the scaling property in the special case of double-slit diffraction. It was inevitable that the result would have these scaling properties, since the whole proof was geometric, and would have been equally valid when enlarged or reduced on a photocopying machine!

Counterintuitively, this means that a diffracting object with smaller dimensions produces a bigger diffraction pattern, p.



p / Cutting d in half doubles the angles of the diffraction fringes.



q / Double-slit diffraction patterns of long-wavelength red light (top) and short-wavelength blue light (bottom).

Double-slit diffraction of blue and red light *example 1*

Blue light has a shorter wavelength than red. For a given double-slit spacing d, the smaller value of λ/d for leads to smaller values of $\sin\theta$, and therefore to a more closely spaced set of diffraction fringes, (g)

The correspondence principle *example 2*

Let's also consider how the equations for double-slit diffraction relate to the correspondence principle. When the ratio λ/d is very small, we should recover the case of simple ray optics. Now if λ/d is small, $\sin\theta$ must be small as well, and the spacing between the diffraction fringes will be small as well. Although we have not proven it, the central fringe is always the brightest, and the fringes get dimmer and dimmer as we go farther from it. For small values of λ/d, the part of the diffraction pattern that is bright enough to be detectable covers only a small range of angles. This is exactly what we would expect from ray optics: the rays passing through the two slits would remain parallel, and would continue moving in the $\theta = 0$ direction. (In fact there would be images of the two separate slits on the screen, but our analysis was all in terms of angles, so we should not expect it to address the issue of whether there is structure within a set of rays that are all traveling in the $\theta = 0$ direction.)

Spacing of the fringes at small angles — *example 3*

At small angles, we can use the approximation $\sin\theta \approx \theta$, which is valid if θ is measured in radians. The equation for double-slit diffraction becomes simply

$$\frac{\lambda}{d} = \frac{\theta}{m} \quad ,$$

which can be solved for θ to give

$$\theta = \frac{m\lambda}{d} \quad .$$

The difference in angle between successive fringes is the change in θ that results from changing m by plus or minus one,

$$\Delta\theta = \frac{\lambda}{d} \quad .$$

For example, if we write θ_7 for the angle of the seventh bright fringe on one side of the central maximum and θ_8 for the neighboring one, we have

$$\begin{aligned} \theta_8 - \theta_7 &= \frac{8\lambda}{d} - \frac{7\lambda}{d} \\ &= \frac{\lambda}{d} \quad , \end{aligned}$$

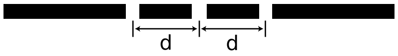
and similarly for any other neighboring pair of fringes.

Although the equation $\lambda/d = \sin\theta/m$ is only valid for a double slit, it is can still be a guide to our thinking even if we are observing diffraction of light by a virus or a flea's leg: it is always true that

(1) large values of λ/d lead to a broad diffraction pattern, and

(2) diffraction patterns are repetitive.

In many cases the equation looks just like $\lambda/d = \sin\theta/m$ but with an extra numerical factor thrown in, and with d interpreted as some other dimension of the object, e.g., the diameter of a piece of wire.

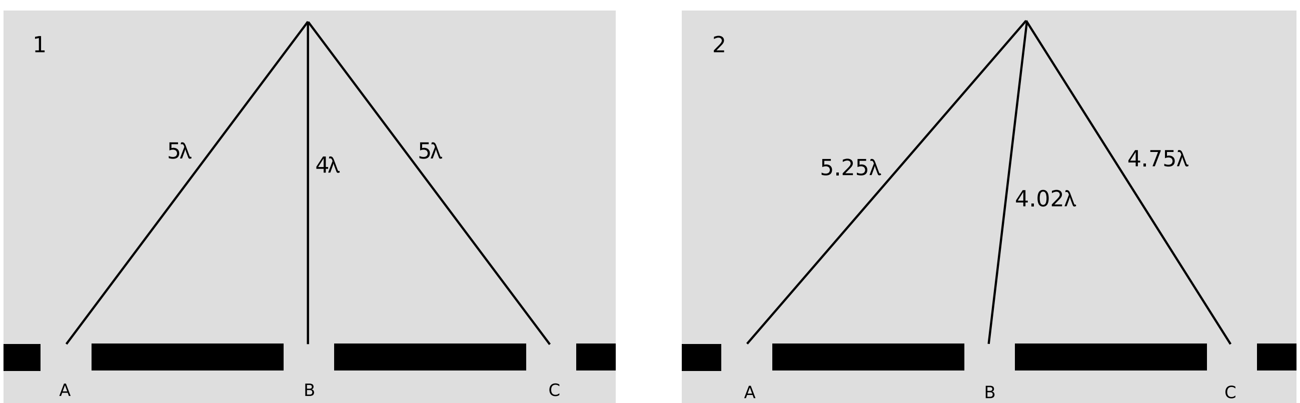


r / A triple slit.

5.6 Repetition

Suppose we replace a double slit with a triple slit, r. We can think of this as a third repetition of the structures that were present in the double slit. Will this device be an improvement over the double slit for any practical reasons?

The answer is yes, as can be shown using figure s. For ease of visualization, I have violated our usual rule of only considering points very far from the diffracting object. The scale of the drawing is such that a wavelengths is one cm. In s/1, all three waves travel an integer number of wavelengths to reach the same point, so there is a bright central spot, as we would expect from our experience with the double slit. In figure s/2, we show the path lengths to a new point. This point is farther from slit A by a quarter of a wavelength, and correspondingly closer to slit C. The distance from slit B has hardly changed at all. Because the paths lengths traveled from slits A and C differ from half a wavelength, there will be perfect destructive interference between these two waves. There is still some uncanceled wave intensity because of slit B, but the amplitude will be three times less than in figure s/1, resulting in a factor of 9 decrease in brightness. Thus, by moving off to the right a little, we have gone from the bright central maximum to a point that is quite dark.



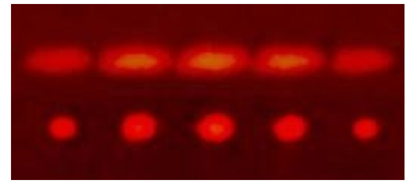
s / 1. There is a bright central maximum. 2. At this point just off the central maximum, the path lengths traveled by the three waves have changed.

Now let's compare with what would have happened if slit C had been covered, creating a plain old double slit. The waves coming from slits A and B would have been out of phase by 0.23 wavelengths, but this would not have caused very severe interference. The point in figure s/2 would have been quite brightly lit up.

To summarize, we have found that adding a third slit narrows down the central fringe dramatically. The same is true for all the other fringes as well, and since the same amount of energy is concentrated in narrower diffraction fringes, each fringe is brighter and easier to see, t.

This is an example of a more general fact about diffraction: if some feature of the diffracting object is repeated, the locations of the maxima and minima are unchanged, but they become narrower.

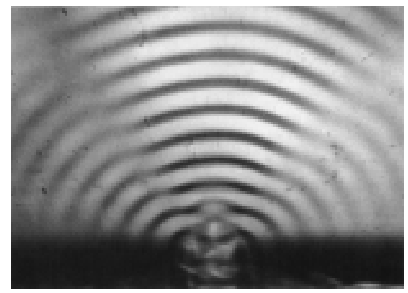
Taking this reasoning to its logical conclusion, a diffracting object with thousands of slits would produce extremely narrow fringes. Such an object is called a diffraction grating.



t / A double-slit diffraction pattern (top), and a pattern made by five slits (bottom).

5.7 Single-Slit Diffraction

If we use only a single slit, is there diffraction? If the slit is not wide compared to a wavelength of light, then we can approximate its behavior by using only a single set of Huygens ripples. There are no other sets of ripples to add to it, so there are no constructive or destructive interference effects, and no maxima or minima. The result will be a uniform spherical wave of light spreading out in all directions, like what we would expect from a tiny lightbulb. We could call this a diffraction pattern, but it is a completely featureless one, and it could not be used, for instance, to determine the wavelength of the light, as other diffraction patterns could.



u / Single-slit diffraction of water waves.

All of this, however, assumes that the slit is narrow compared to a wavelength of light. If, on the other hand, the slit is broader, there will indeed be interference among the sets of ripples spreading out from various points along the opening. Figure u shows an example with water waves, and figure v with light.

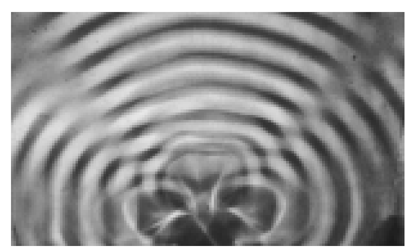
v / Single-slit diffraction of red light. Note the double width of the central maximum.

self-check C

How does the wavelength of the waves compare with the width of the slit in figure u? ▷ Answer, p. 107

We will not go into the details of the analysis of single-slit diffraction, but let us see how its properties can be related to the general things we've learned about diffraction. We know based on scaling arguments that the angular sizes of features in the diffraction pattern must be related to the wavelength and the width, a, of the slit by some relationship of the form

$$\frac{\lambda}{a} \leftrightarrow \theta \qquad .$$



w / A pretty good simulation of the single-slit pattern of figure u, made by using three motors to produce overlapping ripples from three neighboring points in the water.

This is indeed true, and for instance the angle between the maximum of the central fringe and the maximum of the next fringe on one side equals $1.5\lambda/a$. Scaling arguments will never produce factors such as the 1.5, but they tell us that the answer must involve λ/a, so all the familiar qualitative facts are true. For instance, shorter-wavelength light will produce a more closely spaced diffraction pattern.

x / An image of the Pleiades star cluster. The circular rings around the bright stars are due to single-slit diffraction at the mouth of the telescope's tube.

y / A radio telescope.

An important scientific example of single-slit diffraction is in telescopes. Images of individual stars, as in figure x, are a good way to examine diffraction effects, because all stars except the sun are so far away that no telescope, even at the highest magnification, can image their disks or surface features. Thus any features of a star's image must be due purely to optical effects such as diffraction. A prominent cross appears around the brightest star, and dimmer ones surround the dimmer stars. Something like this is seen in most telescope photos, and indicates that inside the tube of the telescope there were two perpendicular struts or supports. Light diffracted around these struts. You might think that diffraction could be eliminated entirely by getting rid of all obstructions in the tube, but the circles around the stars are diffraction effects arising from single-slit diffraction at the mouth of the telescope's tube! (Actually we have not even talked about diffraction through a circular opening, but the idea is the same.) Since the angular sizes of the diffracted images depend on λ/a, the only way to improve the resolution of the images is to increase the diameter, a, of the tube. This is one of the main reasons (in addition to light-gathering power) why the best telescopes must be very large in diameter.

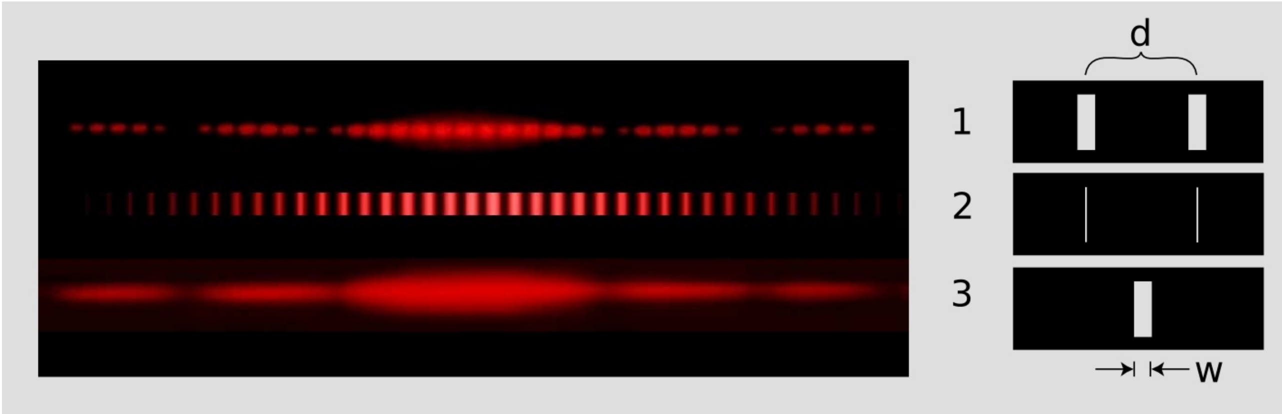
self-check D

What would this imply about radio telescopes as compared with visible-light telescopes? ▷ Answer, p. 107

Double-slit diffraction is easier to understand conceptually than single-slit diffraction, but if you do a double-slit diffraction experiment in real life, you are likely to encounter a complicated pattern like figure z/1, rather than the simpler one, 2, you were expecting. This is because the slits are fairly big compared to the wavelength of the light being used. We really have two different distances in our pair of slits: d, the distance between the slits, and w, the width of each slit. Remember that smaller distances on the object the light diffracts around correspond to larger features of the diffraction pattern. The pattern 1 thus has two spacings in it: a short spacing corresponding to the large distance d, and a long spacing that relates to the small dimension w.

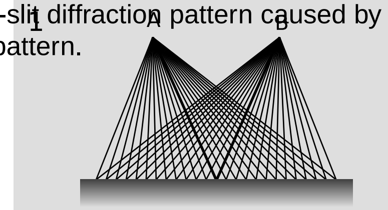
Discussion Question

A Why is it optically impossible for bacteria to evolve eyes that use visible light to form images?



z / 1. A diffraction pattern formed by a real double slit. The width of each slit is fairly big compared to the wavelength of the light. This is a real photo. 2. This idealized pattern is not likely to occur in real life. To get it, you would need each slit to be so narrow that its width was comparable to the wavelength of the light, but that's not usually possible. This is not a real photo. 3. A real photo of a single-slit diffraction pattern caused by a slit whose width is the same as the widths of the slits used to make the top pattern.

5.8 $\int\star$ The Principle of Least Time



2

12.25
11.91
11.59
11.30
11.03
10.80
10.59
10.41
10.26
10.15
10.07
10.02
10.00
10.02
10.07
10.15
10.26
10.41
10.59
10.80
11.03
11.30
11.59
11.91
12.25

aa / Light could take many different paths from A to B.

In sections 1.5 and 4.4, we saw how in the ray model of light, both refraction and reflection can be described in an elegant and beautiful way by a single principle, the principle of least time. We can now justify the principle of least time based on the wave model of light. Consider an example involving reflection, aa. Starting at point A, Huygens' principle for waves tells us that we can think of the wave as spreading out in all directions. Suppose we imagine all the possible ways that a ray could travel from A to B. We show this by drawing 25 possible paths, of which the central one is the shortest. Since the principle of least time connects the wave model to the ray model, we should expect to get the most accurate results when the wavelength is much shorter than the distances involved — for the sake of this numerical example, let's say that a wavelength is 1/10 of the shortest reflected path from A to B. The table, 2, shows the distances traveled by the 25 rays.

Note how similar are the distances traveled by the group of 7 rays, indicated with a bracket, that come closest to obeying the principle of least time. If we think of each one as a wave, then all 7 are again nearly in phase at point B. However, the rays that are farther from satisfying the principle of least time show more rapidly changing distances; on reuniting at point B, their phases are a random jumble, and they will very nearly cancel each other out. Thus, almost none of the wave energy delivered to point B goes by these longer paths. Physically we find, for instance, that a wave pulse emitted at A is observed at B after a time interval corresponding very nearly to the shortest possible path, and the pulse is not very "smeared out" when it gets there. The shorter the wavelength compared to the dimensions of the figure, the more accurate these approximate statements become.

Instead of drawing a finite number of rays, such 25, what happens if we think of the angle, θ, of emission of the ray as a continuously varying variable? Minimizing the distance L requires

$$\frac{\mathrm{d}L}{\mathrm{d}\theta} = 0 \qquad .$$

Because L is changing slowly in the vicinity of the angle that satisfies the principle of least time, all the rays that come out close to this angle have very nearly the same L, and remain very nearly in phase when they reach B. This is the basic reason why the discrete table, aa/2, turned out to have a group of rays that all traveled nearly the same distance.

As discussed in section 1.5, the principle of least time is really a principle of least *or greatest* time. This makes perfect sense, since $\mathrm{d}L/\mathrm{d}\theta = 0$ can in general describe either a minimum or a maximum

The principle of least time is very general. It does not apply just to refraction and reflection — it can even be used to prove that light rays travel in a straight line through empty space, without taking detours! This general approach to wave motion was used by Richard Feynman, one of the pioneers who in the 1950's reconciled quantum mechanics with relativity (book 6 in this series). A very readable explanation is given in a book Feynman wrote for laypeople, QED: The Strange Theory of Light and Matter.

Summary

Selected Vocabulary

diffraction . . . . the behavior of a wave when it encounters an obstacle or a nonuniformity in its medium; in general, diffraction causes a wave to bend around obstacles and make patterns of strong and weak waves radiating out beyond the obstacle.

coherent . . . . . a light wave whose parts are all in phase with each other

Other Terminology and Notation

wavelets . . . . . the ripples in Huygens' principle

Summary

Wave optics is a more general theory of light than ray optics. When light interacts with material objects that are much larger then one wavelength of the light, the ray model of light is approximately correct, but in other cases the wave model is required.

Huygens' principle states that, given a wavefront at one moment in time, the future behavior of the wave can be found by breaking the wavefront up into a large number of small, side-by-side wave peaks, each of which then creates a pattern of circular or spherical ripples. As these sets of ripples add together, the wave evolves and moves through space. Since Huygens' principle is a purely geometrical construction, diffraction effects obey a simple scaling rule: the behavior is unchanged if the wavelength and the dimensions of the diffracting objects are both scaled up or down by the same factor. If we wish to predict the angles at which various features of the diffraction pattern radiate out, scaling requires that these angles depend only on the unitless ratio λ/d, where d is the size of some feature of the diffracting object.

Double-slit diffraction is easily analyzed using Huygens' principle if the slits are narrower than one wavelength. We need only construct two sets of ripples, one spreading out from each slit. The angles of the maxima (brightest points in the bright fringes) and minima (darkest points in the dark fringes) are given by the equation

$$\frac{\lambda}{d} = \frac{\sin\theta}{m} \qquad ,$$

where d is the center-to-center spacing of the slits, and m is an integer at a maximum or an integer plus 1/2 at a minimum.

If some feature of a diffracting object is repeated, the diffraction fringes remain in the same places, but become narrower with each repetition. By repeating a double-slit pattern hundreds or thousands of times, we obtain a diffraction grating.

A single slit can produce diffraction fringes if it is larger than one wavelength. Many practical instances of diffraction can be interpreted as single-slit diffraction, e.g., diffraction in telescopes. The main thing to realize about single-slit diffraction is that it exhibits the same kind of relationship between λ, d, and angles of fringes as in any other type of diffraction.

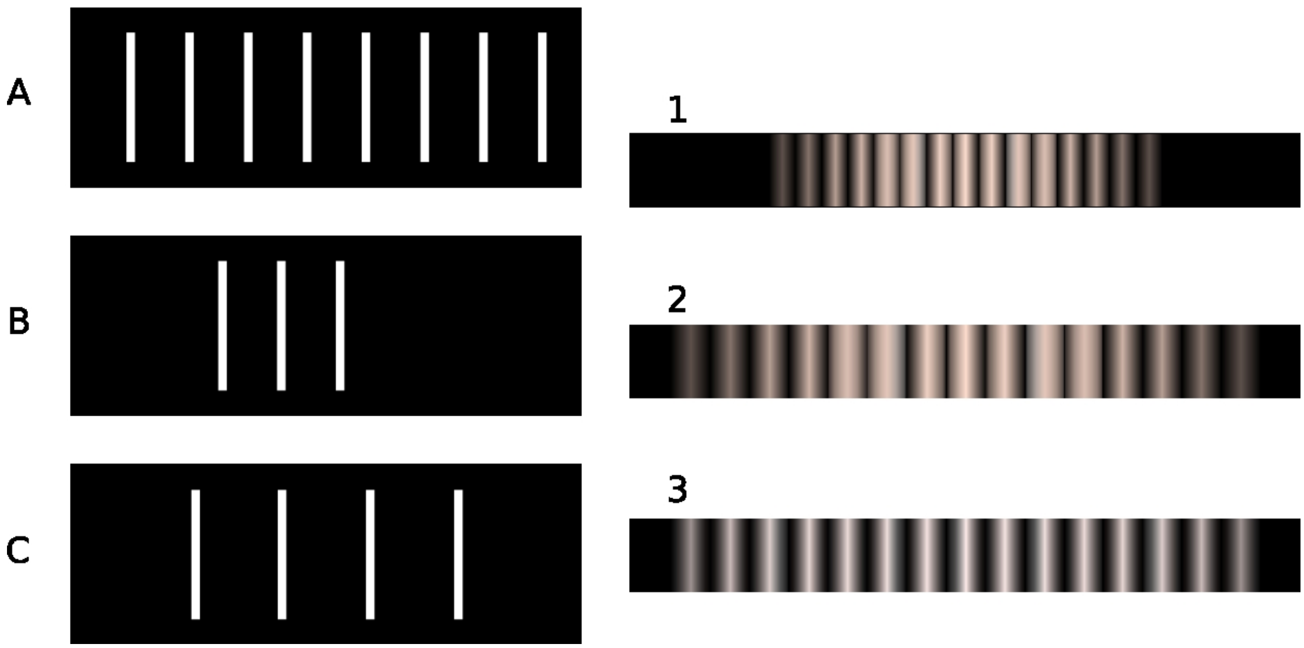
Problems

Key

- √ A computerized answer check is available online.
- ∫ A problem that requires calculus.
- ⋆ A difficult problem.

1 Why would blue or violet light be the best for microscopy?

2 Match gratings A-C with the diffraction patterns 1-3 that they produce. Explain.



3 The beam of a laser passes through a diffraction grating, fans out, and illuminates a wall that is perpendicular to the original beam, lying at a distance of 2.0 m from the grating. The beam is produced by a helium-neon laser, and has a wavelength of 694.3 nm. The grating has 2000 lines per centimeter. (a) What is the distance on the wall between the central maximum and the maxima immediately to its right and left? (b) How much does your answer change when you use the approximation ? √

4 When white light passes through a diffraction grating, what is the smallest value of m for which the visible spectrum of order m overlaps the next one, of order $m+1$? (The visible spectrum runs from about 400 nm to about 700 nm.)

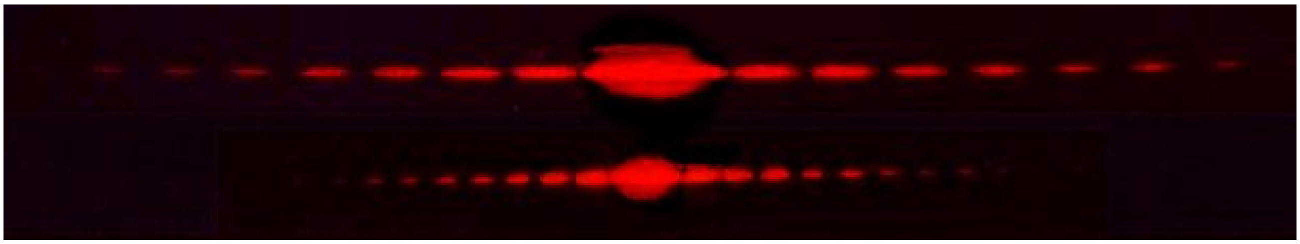
5 Ultrasound, i.e., sound waves with frequencies too high to be audible, can be used for imaging fetuses in the womb or for breaking up kidney stones so that they can be eliminated by the body. Consider the latter application. Lenses can be built to focus sound waves, but because the wavelength of the sound is not all that small compared to the diameter of the lens, the sound will not be concentrated exactly at the geometrical focal point. Instead, a diffraction pattern will be created with an intense central spot surrounded by fainter rings. About 85% of the power is concentrated within the central spot. The angle of the first minimum (surrounding the central spot) is given by $\sin\theta = \lambda/b$, where b is the diameter of the lens. This is similar to the corresponding equation for a single slit, but with a factor of 1.22 in front which arises from the circular shape of the aperture. Let the distance from the lens to the patient's kidney stone be $L = 20$ cm. You will want $f > 20$ kHz, so that the sound is inaudible. Find values of b and f that would result in a usable design, where the central spot is small enough to lie within a kidney stone 1 cm in diameter.

6 For star images such as the ones in figure 4.4, estimate the angular width of the diffraction spot due to diffraction at the mouth of the telescope. Assume a telescope with a diameter of 10 meters (the largest currently in existence), and light with a wavelength in the middle of the visible range. Compare with the actual angular size of a star of diameter 10^9 m seen from a distance of $10^{17}m$. What does this tell you?

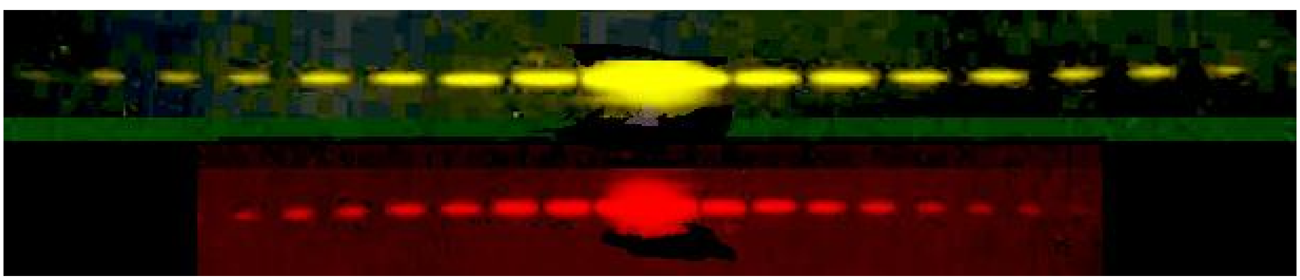
7 Under what circumstances could one get a mathematically undefined result by solving the double-slit diffraction equation for θ? Give a physical interpretation of what would actually be observed.

8 When ultrasound is used for medical imaging, the frequency may be as high as 5-20 MHz. Another medical application of ultrasound is for therapeutic heating of tissues inside the body; here, the frequency is typically 1-3 MHz. What fundamental physical reasons could you suggest for the use of higher frequencies for imaging?

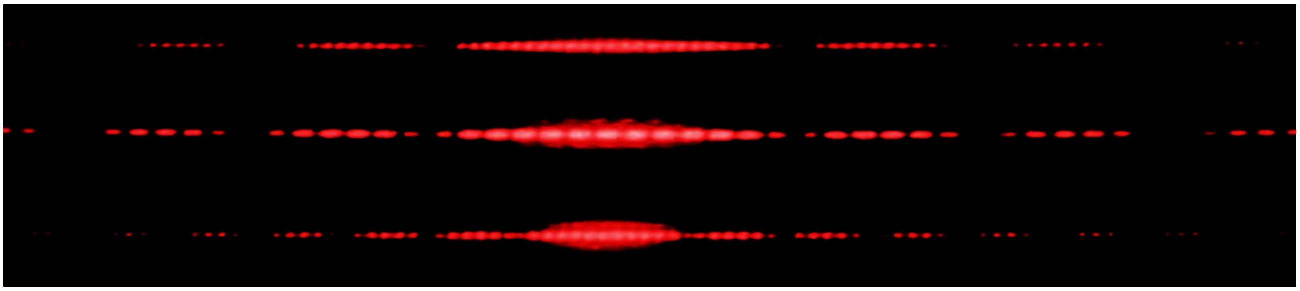
9 The figure below shows two diffraction patterns, both made with the same wavelength of red light. (a) What type of slits made the patterns? Is it a single slit, double slits, or something else? Explain. (b) Compare the dimensions of the slits used to make the top and bottom pattern. Give a numerical ratio, and state which way the ratio is, i.e., which slit pattern was the larger one. Explain.

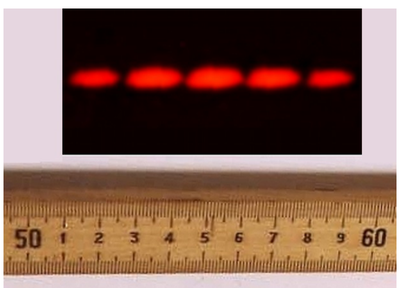


10 The figure below shows two diffraction patterns. The top one was made with yellow light, and the bottom one with red. Could the slits used to make the two patterns have been the same?



11 The figure below shows three diffraction patterns. All were made under identical conditions, except that a different set of double slits was used for each one. The slits used to make the top pattern had a center-to-center separation $d = 0.50$ mm, and each slit was $w = 0.04$ mm wide. (a) Determine d and w for the slits used to make the pattern in the middle. (b) Do the same for the slits used to make the bottom pattern.





Problems 12 and 13.

12 The figure shows a diffraction pattern made by a double slit, along with an image of a meter stick to show the scale. The slits were 146 cm away from the screen on which the diffraction pattern was projected. The spacing of the slits was 0.050 mm. What was the wavelength of the light?

13 The figure shows a diffraction pattern made by a double slit, along with an image of a meter stick to show the scale. Sketch the diffraction pattern from the figure on your paper. Now consider the four variables in the equation $\lambda/d = \sin\theta/m$. Which of these are the same for all five fringes, and which are different for each fringe? Which variable would you naturally use in order to label which fringe was which? Label the fringes on your sketch using the values of that variable.

Appendix 1: Exercises

Exercise 2A: Exploring Images With a Curved Mirror

Equipment:

curved mirrors like the ones described in this chapter

curved mirrors that bulge outward (for part 6 only)

1. Obtain a curved mirror from your instructor. If it is silvered on both sides, make sure you're working with the hollowed-out 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. How do you explain your observations?

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? Explain what's happening.

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.

Write down a prediction of what will happen. Will you be able to see the image with your eye very close to the mirror?

Prediction:____________________________

Now test your prediction. If your prediction was incorrect, can you explain your results?

4. 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?

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 curved mirror that bulges outward, and that therefore bends light rays away from the central axis. Draw a typical ray diagram. Is the image real, or virtual? Will there be more than one type of image?

Prediction:____________________________

Test your prediction with the new type of mirror.

Exercise 3A: Object and Image Distances

Equipment:

optical benches

inbending 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.

Exercise 4A: How strong are your glasses?

This exercise was created by Dan MacIsaac.

Equipment:

eyeglasses

outbending lenses for students who don't wear glasses, or who use inbending glasses

rulers and metersticks

scratch paper

marking pens

Most people who wear glasses have glasses whose lenses are outbending, which allows them to focus on objects far away. Such a lens cannot form a real image, so its focal length cannot be measured as easily as that of an inbending lens. In this exercise you will determine the focal length of your own glasses by taking them off, holding them at a distance from your face, and looking through them at a set of parallel lines on a piece of paper. The lines will be reduced (the lens's magnification is less than one), and by adjusting the distance between the lens and the paper, you can make the magnification equal 1/2 exactly, so that two spaces between lines as seen through the lens fit into one space as seen simultaneously to the side of the lens. This object distance can be used in order to find the focal length of the lens.

1. Use a marker to draw three evenly spaced parallel lines on the paper. (A spacing of a few cm works well.)

2. Does this technique really measure magnification or does it measure angular magnification? What can you do in your experiment in order to make these two quantities nearly the same, so the math is simpler?

3. Before taking any numerical data, use algebra to find the focal length of the lens in terms of d_o, the object distance that results in a magnification of 1/2.

4. Measure the object distance that results in a magnification of 1/2, and determine the focal length of your lens.

Exercise 5A: Double-Source Interference

Equipment:

ripple tank

1. Observe the wave pattern formed by a single source. Try adjusting the frequency at which the motor runs. What do you have to do to the frequency in order to increase the wavelength, and what do you have to do to decrease it?

2. Observe the interference pattern formed by two sources. For convenience, try to get your wavelength as close as possible to 1 cm. We'll call this setup, with $\lambda = 1$ cm and $d = 2.5$ cm, the default setup.

3. Imagine that you were to double the wavelength and double the distance between the sources. How would a snapshot of this wave pattern compare with a snapshot of the pattern made by the default setup? Based on this, how do you predict the angles of the maxima and minima will compare? ________

Test your predictions.

4. On a piece of paper, make a life-size drawing of the two sources in the default setup, and locate the following points:

A. The point that is 10 wavelengths from source #1 and 10 wavelengths from source #2.

B. The point that is 11 wavelengths from #1 and 11 from #2.

C. The point that is 10 wavelengths from #1 and 10.5 from #2.

D. The point that is 11 wavelengths from #1 and 11.5 from #2.

E. The point that is 10 wavelengths from #1 and 11 from #2.

F. The point that is 11 wavelengths from #1 and 12 from #2.

You can do this either using a compass or by putting the next page under your paper and tracing.

What do these points correspond to in the real wave pattern?

5. Make a fresh copy of your drawing, showing only point E and the two sources, which form a long, skinny triangle. Now suppose you were to change the default setup by doubling d, while leaving λ the same. Realistically this involves moving one peg over one hole, while leaving the other peg in the same place, but it's easier to understand what's happening on the drawing if you move both sources outward, keeping the center fixed. Based on your drawing, what will happen to the position of point E when you double d? How has the angle of point E changed?__________

Test your prediction.

6. In the previous part of the exercise, you saw the effect of doubling d while leaving λthe same. Now what do you think would happen to your angles if, starting from the standard setup, you doubled λ while leaving d the same?_________

Try it.

7. Suppose λ was a millionth of a centimeter, while d was still as in the standard setup. What would happen to the angles? What does this tell you about observing diffraction of light?