

same is true of 2, however. Only in curved regions such as 3 and 4 is an acceleration produced. In these examples, the vector sum of the two forces acting on the central mass is not zero. The important concept is that curvature makes force: the curved areas of a wave tend to experience forces resulting in an acceleration toward the mouth of the curve. Note, however, that an uncurved portion of the string need not remain motionless. It may move at constant velocity to either side.

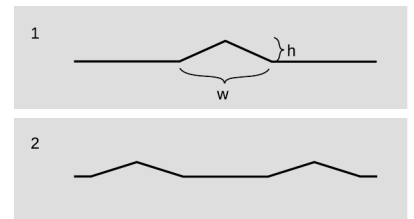
Approximate treatment

We now carry out an approximate treatment of the speed at which two pulses will spread out from an initial indentation on a string. For simplicity, we imagine a hammer blow that creates a triangular dent, n/1. We will estimate the amount of time, t , required until each of the pulses has traveled a distance equal to the width of the pulse itself. The velocity of the pulses is then $\pm w/t$.

As always, the velocity of a wave depends on the properties of the medium, in this case the string. The properties of the string can be summarized by two variables: the tension, T , and the mass per unit length, μ (Greek letter mu).

If we consider the part of the string encompassed by the initial dent as a single object, then this object has a mass of approximately μw (mass/length \times length = mass). (Here, and throughout the derivation, we assume that h is much less than w , so that we can ignore the fact that this segment of the string has a length slightly greater than w .) Although the downward acceleration of this segment of the string will be neither constant over time nor uniform across the string, we will pretend that it is constant for the sake of our simple estimate. Roughly speaking, the time interval between n/1 and 2 is the amount of time required for the initial dent to accelerate from rest and reach its normal, flattened position. Of course the tip of the triangle has a longer distance to travel than the edges, but again we ignore the complications and simply assume that the segment as a whole must travel a distance h . Indeed, it might seem surprising that the triangle would so neatly spring back to a perfectly flat shape. It is an experimental fact that it does, but our analysis is too crude to address such details.

The string is kinked, i.e., tightly curved, at the edges of the triangle, so it is here that there will be large forces that do not cancel out to zero. There are two forces acting on the triangular hump, one of magnitude T acting down and to the right, and one of the same magnitude acting down and to the left. If the angle of the sloping sides is θ , then the total force on the segment equals $2T \sin \theta$. Dividing the triangle into two right triangles, we see that $\sin \theta$ equals h divided by the length of one of the sloping sides. Since h is much less than w , the length of the sloping side is essentially



n / A triangular pulse spreads out.

the same as $w/2$, so we have $\sin \theta = 2h/w$, and $F = 4Th/w$. The acceleration of the segment (actually the acceleration of its center of mass) is

$$\begin{aligned} a &= F/m \\ &= 4Th/\mu w^2. \end{aligned}$$

The time required to move a distance h under constant acceleration a is found by solving $h = \frac{1}{2}at^2$ to yield

$$\begin{aligned} t &= \sqrt{\frac{2h}{a}} \\ &= w\sqrt{\frac{\mu}{2T}}. \end{aligned}$$

Our final result for the velocity of the pulses is

$$\begin{aligned} |v| &= \frac{w}{t} \\ &= \sqrt{\frac{2T}{\mu}}. \end{aligned}$$

The remarkable feature of this result is that the velocity of the pulses does not depend at all on w or h , i.e., any triangular pulse has the same speed. It is an experimental fact (and we will also prove rigorously in the following subsection) that any pulse of any kind, triangular or otherwise, travels along the string at the same speed. Of course, after so many approximations we cannot expect to have gotten all the numerical factors right. The correct result for the velocity of the pulses is

$$v = \sqrt{\frac{T}{\mu}}.$$

The importance of the above derivation lies in the insight it brings —that all pulses move with the same speed — rather than in the details of the numerical result. The reason for our too-high value for the velocity is not hard to guess. It comes from the assumption that the acceleration was constant, when actually the total force on the segment would diminish as it flattened out.

Rigorous derivation using calculus (optional)

After expending considerable effort for an approximate solution, we now display the power of calculus with a rigorous and completely general treatment that is nevertheless much shorter and easier. Let the flat position of the string define the x axis, so that y measures how far a point on the string is from equilibrium. The motion of the string is characterized by $y(x, t)$, a function of two variables. Knowing that the force on any small segment of string depends

on the curvature of the string in that area, and that the second derivative is a measure of curvature, it is not surprising to find that the infinitesimal force dF acting on an infinitesimal segment dx is given by

$$dF = T \frac{d^2 y}{dx^2} dx.$$

(This can be proved by vector addition of the two infinitesimal forces acting on either side.) The acceleration is then $a = dF/dm$, or, substituting $dm = \mu dx$,

$$\frac{d^2 y}{dt^2} = \frac{T}{\mu} \frac{d^2 y}{dx^2}.$$

The second derivative with respect to time is related to the second derivative with respect to position. This is no more than a fancy mathematical statement of the intuitive fact developed above, that the string accelerates so as to flatten out its curves.

Before even bothering to look for solutions to this equation, we note that it already proves the principle of superposition, because the derivative of a sum is the sum of the derivatives. Therefore the sum of any two solutions will also be a solution.

Based on experiment, we expect that this equation will be satisfied by any function $y(x, t)$ that describes a pulse or wave pattern moving to the left or right at the correct speed v . In general, such a function will be of the form $y = f(x - vt)$ or $y = f(x + vt)$, where f is any function of one variable. Because of the chain rule, each derivative with respect to time brings out a factor of $\pm v$. Evaluating the second derivatives on both sides of the equation gives

$$(\pm v)^2 f'' = \frac{T}{\mu} f''.$$

Squaring gets rid of the sign, and we find that we have a valid solution for any function f , provided that v is given by

$$v = \sqrt{\frac{T}{\mu}}.$$

Significance of the result

This specific result for the speed of waves on a string, $v = \sqrt{T/\mu}$, is utterly unimportant. Don't memorize it. Don't take notes on it. Try to erase it from your memory.

What *is* important about this result is that it is an example of two things that are usually true, at least approximately, for mechanical waves in general:

1. The speed at which a wave moves does not depend on the size or shape of the wave.
2. The speed of a mechanical wave depends on a combination of two properties of the medium: some measure of its *inertia* and some measure of its *tightness*, i.e., the strength of the force trying to bring the medium back toward equilibrium.

self-check B

- (a) What is it about the equation $v = \sqrt{T/\mu}$ that relates to fact 1 above?
(b) In the equation $v = \sqrt{T/\mu}$, which variable is a measure of inertia, and which is a measure of tightness? (c) Now suppose that we produce compressional wave pulses in a metal rod by tapping the end of the rod with a hammer. What physical properties of the rod would play the roles of inertia and tightness? How would you expect the speed of compressional waves in lead to compare with their speed in aluminum?
▷ Answer, p. 568

19.3 Sound and light waves

Sound waves

The phenomenon of sound is easily found to have all the characteristics we expect from a wave phenomenon:

- Sound waves obey superposition. Sounds do not knock other sounds out of the way when they collide, and we can hear more than one sound at once if they both reach our ear simultaneously.
- The medium does not move with the sound. Even standing in front of a titanic speaker playing earsplitting music, we do not feel the slightest breeze.
- The velocity of sound depends on the medium. Sound travels faster in helium than in air, and faster in water than in helium. Putting more energy into the wave makes it more intense, not faster. For example, you can easily detect an echo when you clap your hands a short distance from a large, flat wall, and the delay of the echo is no shorter for a louder clap.

Although not all waves have a speed that is independent of the shape of the wave, and this property therefore is irrelevant to our collection of evidence that sound is a wave phenomenon, sound does nevertheless have this property. For instance, the music in a large concert hall or stadium may take on the order of a second to reach someone seated in the nosebleed section, but we do not notice or care, because the delay is the same for every sound. Bass, drums, and vocals all head outward from the stage at 340 m/s, regardless of their differing wave shapes.

If sound has all the properties we expect from a wave, then what type of wave is it? It must be a vibration of a physical medium such as air, since the speed of sound is different in different media, such as helium or water. Further evidence is that we don't receive sound signals that have come to our planet through outer space. The roars and whooshes of Hollywood's space ships are fun, but scientifically wrong.¹

We can also tell that sound waves consist of compressions and expansions, rather than sideways vibrations like the shimmying of a snake. Only compressional vibrations would be able to cause your eardrums to vibrate in and out. Even for a very loud sound, the compression is extremely weak; the increase or decrease compared to normal atmospheric pressure is no more than a part per million. Our ears are apparently very sensitive receivers! Unlike a wave on a string, which vibrates in the direction perpendicular to the direction in which the wave pattern moves, a sound wave is a longitudinal wave, i.e., one in which the vibration is forward and backward along the direction of motion.

Light waves

Entirely similar observations lead us to believe that light is a wave, although the concept of light as a wave had a long and tortuous history. It is interesting to note that Isaac Newton very influentially advocated a contrary idea about light. The belief that matter was made of atoms was stylish at the time among radical thinkers (although there was no experimental evidence for their existence), and it seemed logical to Newton that light as well should be made of tiny particles, which he called corpuscles (Latin for "small objects"). Newton's triumphs in the science of mechanics, i.e., the study of matter, brought him such great prestige that nobody bothered to

¹Outer space is not a perfect vacuum, so it is possible for sound waves to travel through it. However, if we want to create a sound wave, we typically do it by creating vibrations of a physical object, such as the sounding board of a guitar, the reed of a saxophone, or a speaker cone. The lower the density of the surrounding medium, the less efficiently the energy can be converted into sound and carried away. An isolated tuning fork, left to vibrate in interstellar space, would dissipate the energy of its vibration into internal heat at a rate many orders of magnitude greater than the rate of sound emission into the nearly perfect vacuum around it.

question his incorrect theory of light for 150 years. One persuasive proof that light is a wave is that according to Newton's theory, two intersecting beams of light should experience at least some disruption because of collisions between their corpuscles. Even if the corpuscles were extremely small, and collisions therefore very infrequent, at least some dimming should have been measurable. In fact, very delicate experiments have shown that there is no dimming.

The wave theory of light was entirely successful up until the 20th century, when it was discovered that not all the phenomena of light could be explained with a pure wave theory. It is now believed that both light and matter are made out of tiny chunks which have *both* wave and particle properties. For now, we will content ourselves with the wave theory of light, which is capable of explaining a great many things, from cameras to rainbows.

If light is a wave, what is waving? What is the medium that wiggles when a light wave goes by? It isn't air. A vacuum is impenetrable to sound, but light from the stars travels happily through zillions of miles of empty space. Light bulbs have no air inside them, but that doesn't prevent the light waves from leaving the filament. For a long time, physicists assumed that there must be a mysterious medium for light waves, and they called it the aether (not to be confused with the chemical). Supposedly the aether existed everywhere in space, and was immune to vacuum pumps. The details of the story are more fittingly reserved for later in this course, but the end result was that a long series of experiments failed to detect any evidence for the aether, and it is no longer believed to exist. Instead, light can be explained as a wave pattern made up of electrical and magnetic fields.

19.4 Periodic waves

Period and frequency of a periodic wave

You choose a radio station by selecting a certain frequency. We have already defined period and frequency for vibrations, but what do they signify in the case of a wave? We can recycle our previous definition simply by stating it in terms of the vibrations that the wave causes as it passes a receiving instrument at a certain point in space. For a sound wave, this receiver could be an eardrum or a microphone. If the vibrations of the eardrum repeat themselves over and over, i.e., are periodic, then we describe the sound wave that caused them as periodic. Likewise we can define the period and frequency of a wave in terms of the period and frequency of the vibrations it causes. As another example, a periodic water wave would be one that caused a rubber duck to bob in a periodic manner as they passed by it.

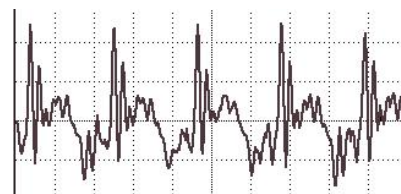
The period of a sound wave correlates with our sensory impression of musical pitch. A high frequency (short period) is a high note. The sounds that really define the musical notes of a song are only the ones that are periodic. It is not possible to sing a non-periodic sound like “sh” with a definite pitch.

The frequency of a light wave corresponds to color. Violet is the high-frequency end of the rainbow, red the low-frequency end. A color like brown that does not occur in a rainbow is not a periodic light wave. Many phenomena that we do not normally think of as light are actually just forms of light that are invisible because they fall outside the range of frequencies our eyes can detect. Beyond the red end of the visible rainbow, there are infrared and radio waves. Past the violet end, we have ultraviolet, x-rays, and gamma rays.

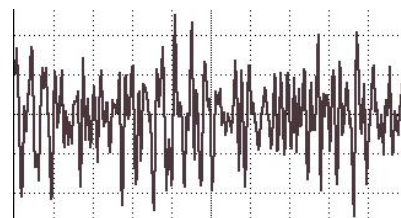
Graphs of waves as a function of position

Some waves, like sound waves, are easy to study by placing a detector at a certain location in space and studying the motion as a function of time. The result is a graph whose horizontal axis is time. With a water wave, on the other hand, it is simpler just to look at the wave directly. This visual snapshot amounts to a graph of the height of the water wave as a function of *position*. Any wave can be represented in either way.

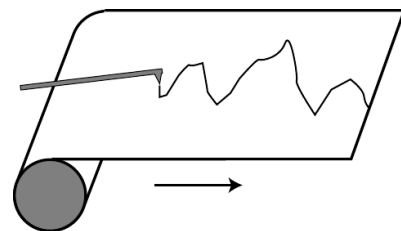
An easy way to visualize this is in terms of a strip chart recorder, an obsolescing device consisting of a pen that wiggles back and forth as a roll of paper is fed under it. It can be used to record a person’s electrocardiogram, or seismic waves too small to be felt as a noticeable earthquake but detectable by a seismometer. Taking the seismometer as an example, the chart is essentially a record of the ground’s wave motion as a function of time, but if the paper was set to feed at the same velocity as the motion of an earthquake wave, it



o / A graph of pressure versus time for a periodic sound wave, the vowel “ah.”

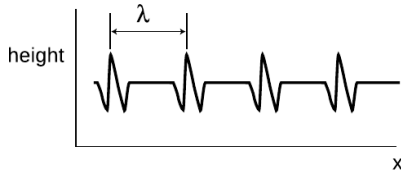


p / A similar graph for a non-periodic wave, “sh.”



q / A strip chart recorder.

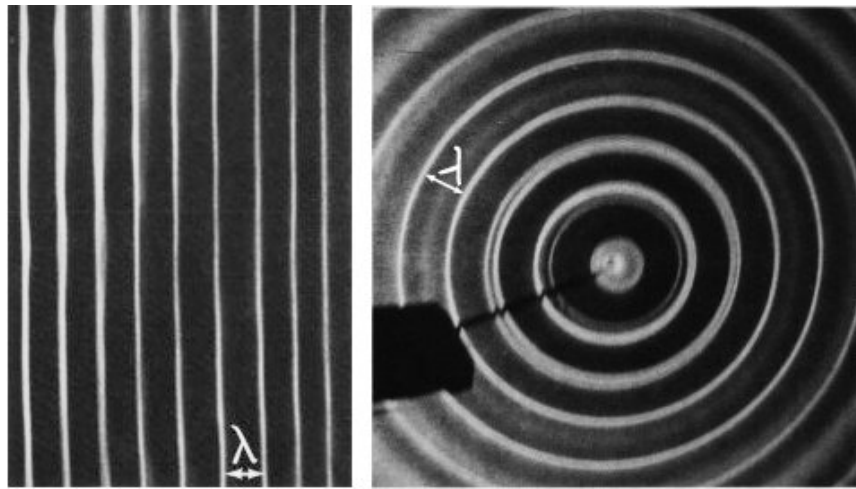
would also be a full-scale representation of the profile of the actual wave pattern itself. Assuming, as is usually the case, that the wave velocity is a constant number regardless of the wave's shape, knowing the wave motion as a function of time is equivalent to knowing it as a function of position.



r / A water wave profile created by a series of repeating pulses.

Wavelength

Any wave that is periodic will also display a repeating pattern when graphed as a function of position. The distance spanned by one repetition is referred to as one *wavelength*. The usual notation for wavelength is λ , the Greek letter lambda. Wavelength is to space as period is to time.



s / Wavelengths of linear and circular water waves.

Wave velocity related to frequency and wavelength

Suppose that we create a repetitive disturbance by kicking the surface of a swimming pool. We are essentially making a series of wave pulses. The wavelength is simply the distance a pulse is able to travel before we make the next pulse. The distance between pulses is λ , and the time between pulses is the period, T , so the speed of the wave is the distance divided by the time,

$$v = \lambda/T.$$

This important and useful relationship is more commonly written in terms of the frequency,

$$v = f\lambda.$$

Wavelength of radio waves

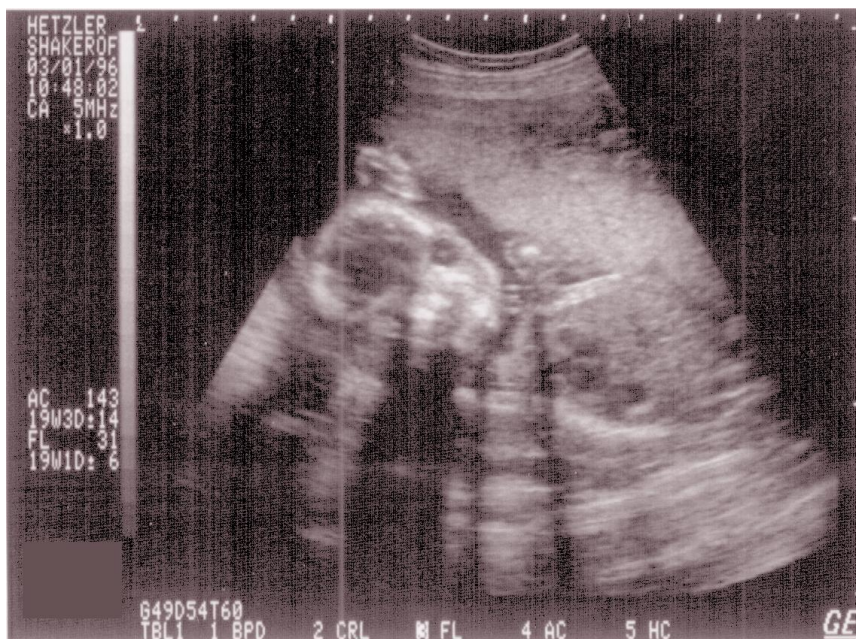
example 5

▷ The speed of light is 3.0×10^8 m/s. What is the wavelength of the radio waves emitted by KMHD, a station whose frequency is 89.1 MHz?

▷ Solving for wavelength, we have

$$\begin{aligned}\lambda &= v/f \\ &= (3.0 \times 10^8 \text{ m/s}) / (89.1 \times 10^6 \text{ s}^{-1}) \\ &= 3.4 \text{ m}\end{aligned}$$

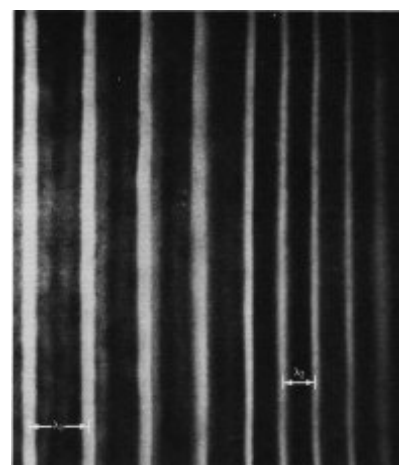
The size of a radio antenna is closely related to the wavelength of the waves it is intended to receive. The match need not be exact (since after all one antenna can receive more than one wavelength!), but the ordinary “whip” antenna such as a car’s is 1/4 of a wavelength. An antenna optimized to receive KMHD’s signal would have a length of $3.4 \text{ m}/4 = 0.85 \text{ m}$.



t / Ultrasound, i.e., sound with frequencies higher than the range of human hearing, was used to make this image of a fetus. The resolution of the image is related to the wavelength, since details smaller than about one wavelength cannot be resolved. High resolution therefore requires a short wavelength, corresponding to a high frequency.

The equation $v = f\lambda$ defines a fixed relationship between any two of the variables if the other is held fixed. The speed of radio waves in air is almost exactly the same for all wavelengths and frequencies (it is exactly the same if they are in a vacuum), so there is a fixed relationship between their frequency and wavelength. Thus we can say either “Are we on the same wavelength?” or “Are we on the same frequency?”

A different example is the behavior of a wave that travels from a region where the medium has one set of properties to an area where the medium behaves differently. The frequency is now fixed,



u / A water wave traveling into a region with a different depth changes its wavelength.

because otherwise the two portions of the wave would otherwise get out of step, causing a kink or discontinuity at the boundary, which would be unphysical. (A more careful argument is that a kink or discontinuity would have infinite curvature, and waves tend to flatten out their curvature. An infinite curvature would flatten out infinitely fast, i.e., it could never occur in the first place.) Since the frequency must stay the same, any change in the velocity that results from the new medium must cause a change in wavelength.

The velocity of water waves depends on the depth of the water, so based on $\lambda = v/f$, we see that water waves that move into a region of different depth must change their wavelength, as shown in figure u. This effect can be observed when ocean waves come up to the shore. If the deceleration of the wave pattern is sudden enough, the tip of the wave can curl over, resulting in a breaking wave.

A note on dispersive waves

The discussion of wave velocity given here is actually an oversimplification for a wave whose velocity depends on its frequency and wavelength. Such a wave is called a dispersive wave. Nearly all the waves we deal with in this course are non-dispersive, but the issue becomes important in quantum physics, as discussed in more detail in optional section 35.2.

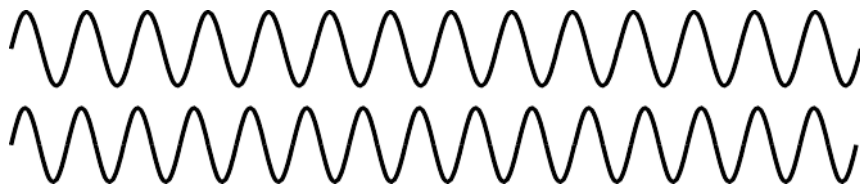
Sinusoidal waves

Sinusoidal waves are the most important special case of periodic waves. In fact, many scientists and engineers would be uncomfortable with defining a waveform like the “ah” vowel sound as having a definite frequency and wavelength, because they consider only sine waves to be pure examples of a certain frequency and wavelengths. Their bias is not unreasonable, since the French mathematician Fourier showed that any periodic wave with frequency f can be constructed as a superposition of sine waves with frequencies $f, 2f, 3f, \dots$. In this sense, sine waves are the basic, pure building blocks of all waves. (Fourier’s result so surprised the mathematical community of France that he was ridiculed the first time he publicly presented his theorem.)

However, what definition to use is a matter of utility. Our sense of hearing perceives any two sounds having the same period as possessing the same pitch, regardless of whether they are sine waves or not. This is undoubtedly because our ear-brain system evolved to be able to interpret human speech and animal noises, which are periodic but not sinusoidal. Our eyes, on the other hand, judge a color as pure (belonging to the rainbow set of colors) only if it is a sine wave.

Discussion question

A Suppose we superimpose two sine waves with equal amplitudes but slightly different frequencies, as shown in the figure. What will the superposition look like? What would this sound like if they were sound waves?



19.5 The Doppler effect

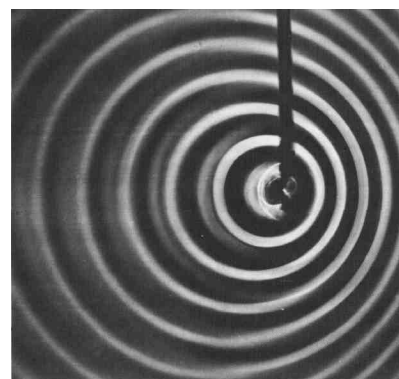
Figure v shows the wave pattern made by the tip of a vibrating rod which is moving across the water. If the rod had been vibrating in one place, we would have seen the familiar pattern of concentric circles, all centered on the same point. But since the source of the waves is moving, the wavelength is shortened on one side and lengthened on the other. This is known as the Doppler effect.

Note that the velocity of the waves is a fixed property of the medium, so for example the forward-going waves do not get an extra boost in speed as would a material object like a bullet being shot forward from an airplane.

We can also infer a change in frequency. Since the velocity is constant, the equation $v = f\lambda$ tells us that the change in wavelength must be matched by an opposite change in frequency: higher frequency for the waves emitted forward, and lower for the ones emitted backward. The frequency Doppler effect is the reason for the familiar dropping-pitch sound of a race car going by. As the car approaches us, we hear a higher pitch, but after it passes us we hear a frequency that is lower than normal.

The Doppler effect will also occur if the observer is moving but the source is stationary. For instance, an observer moving toward a stationary source will perceive one crest of the wave, and will then be surrounded by the next crest sooner than she otherwise would have, because she has moved toward it and hastened her encounter with it. Roughly speaking, the Doppler effect depends only the relative motion of the source and the observer, not on their absolute state of motion (which is not a well-defined notion in physics) or on their velocity relative to the medium.

Restricting ourselves to the case of a moving source, and to waves emitted either directly along or directly against the direction of motion, we can easily calculate the wavelength, or equivalently the frequency, of the Doppler-shifted waves. Let v be the velocity of the waves, and v_s the velocity of the source. The wavelength of the



v / The pattern of waves made by a point source moving to the right across the water. Note the shorter wavelength of the forward-emitted waves and the longer wavelength of the backward-going ones.

forward-emitted waves is shortened by an amount $v_s T$ equal to the distance traveled by the source over the course of one period. Using the definition $f = 1/T$ and the equation $v = f\lambda$, we find for the wavelength of the Doppler-shifted wave the equation

$$\lambda' = \left(1 - \frac{v_s}{v}\right) \lambda.$$

A similar equation can be used for the backward-emitted waves, but with a plus sign rather than a minus sign.

Doppler-shifted sound from a race car *example 6*

▷ If a race car moves at a velocity of 50 m/s, and the velocity of sound is 340 m/s, by what percentage are the wavelength and frequency of its sound waves shifted for an observer lying along its line of motion?

▷ For an observer whom the car is approaching, we find

$$1 - \frac{v_s}{v} = 0.85,$$

so the shift in wavelength is 15%. Since the frequency is inversely proportional to the wavelength for a fixed value of the speed of sound, the frequency is shifted upward by

$$1/0.85 = 1.18,$$

i.e., a change of 18%. (For velocities that are small compared to the wave velocities, the Doppler shifts of the wavelength and frequency are about the same.)

Doppler shift of the light emitted by a race car *example 7*

▷ What is the percent shift in the wavelength of the light waves emitted by a race car's headlights?

▷ Looking up the speed of light, $v = 3.0 \times 10^8$ m/s, we find

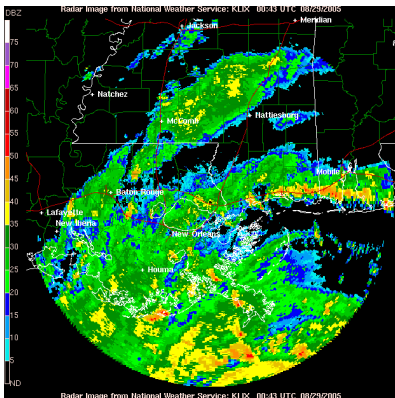
$$1 - \frac{v_s}{v} = 0.99999983,$$

i.e., the percentage shift is only 0.000017%.

The second example shows that under ordinary earthbound circumstances, Doppler shifts of light are negligible because ordinary things go so much slower than the speed of light. It's a different story, however, when it comes to stars and galaxies, and this leads us to a story that has profound implications for our understanding of the origin of the universe.

Doppler radar *example 8*

The first use of radar was by Britain during World War II: antennas on the ground sent radio waves up into the sky, and detected the echoes when the waves were reflected from German planes. Later, air forces wanted to mount radar antennas on airplanes,



w / Example 8. A Doppler radar image of Hurricane Katrina, in 2005.

but then there was a problem, because if an airplane wanted to detect another airplane at a lower altitude, it would have to aim its radio waves downward, and then it would get echoes from the ground. The solution was the invention of Doppler radar, in which echoes from the ground were differentiated from echoes from other aircraft according to their Doppler shifts. A similar technology is used by meteorologists to map out rainclouds without being swamped by reflections from the ground, trees, and buildings.

Optional topic: Doppler shifts of light

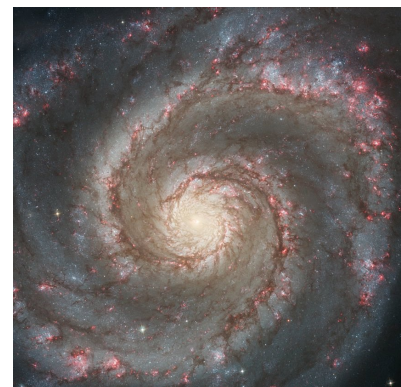
If Doppler shifts depend only on the relative motion of the source and receiver, then there is no way for a person moving with the source and another person moving with the receiver to determine who is moving and who isn't. Either can blame the Doppler shift entirely on the other's motion and claim to be at rest herself. This is entirely in agreement with the principle stated originally by Galileo that all motion is relative.

On the other hand, a careful analysis of the Doppler shifts of water or sound waves shows that it is only approximately true, at low speeds, that the shifts just depend on the relative motion of the source and observer. For instance, it is possible for a jet plane to keep up with its own sound waves, so that the sound waves appear to stand still to the pilot of the plane. The pilot then knows she is moving at exactly the speed of sound. The reason this doesn't disprove the relativity of motion is that the pilot is not really determining her absolute motion but rather her motion relative to the air, which is the medium of the sound waves.

Einstein realized that this solved the problem for sound or water waves, but would not salvage the principle of relative motion in the case of light waves, since light is not a vibration of any physical medium such as water or air. Beginning by imagining what a beam of light would look like to a person riding a motorcycle alongside it, Einstein eventually came up with a radical new way of describing the universe, in which space and time are distorted as measured by observers in different states of motion. As a consequence of this theory of relativity, he showed that light waves would have Doppler shifts that would exactly, not just approximately, depend only on the relative motion of the source and receiver. The resolution of the motorcycle paradox is given in example 7 on p. 712, and a quantitative discussion of Doppler shifts of light is given on p. 716.

The Big Bang

As soon as astronomers began looking at the sky through telescopes, they began noticing certain objects that looked like clouds in deep space. The fact that they looked the same night after night meant that they were beyond the earth's atmosphere. Not knowing what they really were, but wanting to sound official, they called them "nebulae," a Latin word meaning "clouds" but sounding more impressive. In the early 20th century, astronomers realized that although some really were clouds of gas (e.g., the middle "star" of Orion's sword, which is visibly fuzzy even to the naked eye when conditions are good), others were what we now call galaxies: virtual



x / The galaxy M51. Under high magnification, the milky clouds reveal themselves to be composed of trillions of stars.



y / How do astronomers know what mixture of wavelengths a star emitted originally, so that they can tell how much the Doppler shift was? This image (obtained by the author with equipment costing about \$5, and no telescope) shows the mixture of colors emitted by the star Sirius. (If you have the book in black and white, blue is on the left and red on the right.) The star appears white or bluish-white to the eye, but any light looks white if it contains roughly an equal mixture of the rainbow colors, i.e., of all the pure sinusoidal waves with wavelengths lying in the visible range. Note the black “gap teeth.” These are the fingerprint of hydrogen in the outer atmosphere of Sirius. These wavelengths are selectively absorbed by hydrogen. Sirius is in our own galaxy, but similar stars in other galaxies would have the whole pattern shifted toward the red end, indicating they are moving away from us.



z / The telescope at Mount Wilson used by Hubble.

island universes consisting of trillions of stars (for example the Andromeda Galaxy, which is visible as a fuzzy patch through binoculars). Three hundred years after Galileo had resolved the Milky Way into individual stars through his telescope, astronomers realized that the universe is made of galaxies of stars, and the Milky Way is simply the visible part of the flat disk of our own galaxy, seen from inside.

This opened up the scientific study of cosmology, the structure and history of the universe as a whole, a field that had not been seriously attacked since the days of Newton. Newton had realized that if gravity was always attractive, never repulsive, the universe would have a tendency to collapse. His solution to the problem was to posit a universe that was infinite and uniformly populated with matter, so that it would have no geometrical center. The gravitational forces in such a universe would always tend to cancel out by symmetry, so there would be no collapse. By the 20th century, the belief in an unchanging and infinite universe had become conventional wisdom in science, partly as a reaction against the time that had been wasted trying to find explanations of ancient geological phenomena based on catastrophes suggested by biblical events like Noah’s flood.

In the 1920’s astronomer Edwin Hubble began studying the Doppler shifts of the light emitted by galaxies. A former college football player with a serious nicotine addiction, Hubble did not set out to change our image of the beginning of the universe. His autobiography seldom even mentions the cosmological discovery for which he is now remembered. When astronomers began to study the Doppler shifts of galaxies, they expected that each galaxy’s direction and velocity of motion would be essentially random. Some would be approaching us, and their light would therefore be Doppler-shifted to the blue end of the spectrum, while an equal number would be expected to have red shifts. What Hubble discovered instead was that except for a few very nearby ones, all the galaxies had red shifts, indicating that they were receding from us at a hefty fraction of the speed of light. Not only that, but the ones farther away were receding more quickly. The speeds were directly proportional to their distance from us.

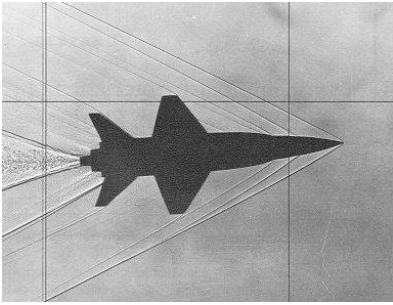
Did this mean that the earth (or at least our galaxy) was the center of the universe? No, because Doppler shifts of light only depend on the relative motion of the source and the observer. If we see a distant galaxy moving away from us at 10% of the speed of light, we can be assured that the astronomers who live in that galaxy will see ours receding from them at the same speed in the opposite direction. The whole universe can be envisioned as a rising loaf of raisin bread. As the bread expands, there is more and more space between the raisins. The farther apart two raisins are, the greater the speed with which they move apart.

Extrapolating backward in time using the known laws of physics, the universe must have been denser and denser at earlier and earlier times. At some point, it must have been extremely dense and hot, and we can even detect the radiation from this early fireball, in the form of microwave radiation that permeates space. The phrase Big Bang was originally coined by the doubters of the theory to make it sound ridiculous, but it stuck, and today essentially all astronomers accept the Big Bang theory based on the very direct evidence of the red shifts and the cosmic microwave background radiation.

What the Big Bang is not

Finally it should be noted what the Big Bang theory is not. It is not an explanation of *why* the universe exists. Such questions belong to the realm of religion, not science. Science can find ever simpler and ever more fundamental explanations for a variety of phenomena, but ultimately science takes the universe as it is according to observations.

Furthermore, there is an unfortunate tendency, even among many scientists, to speak of the Big Bang theory as a description of the very first event in the universe, which caused everything after it. Although it is true that time may have had a beginning (Einstein's theory of general relativity admits such a possibility), the methods of science can only work within a certain range of conditions such as temperature and density. Beyond a temperature of about 10^9 degrees C, the random thermal motion of subatomic particles becomes so rapid that its velocity is comparable to the speed of light. Early enough in the history of the universe, when these temperatures existed, Newtonian physics becomes less accurate, and we must describe nature using the more general description given by Einstein's theory of relativity, which encompasses Newtonian physics as a special case. At even higher temperatures, beyond about 10^{33} degrees, physicists know that Einstein's theory as well begins to fall apart, but we don't know how to construct the even more general theory of nature that would work at those temperatures. No matter how far physics progresses, we will never be able to describe nature at infinitely high temperatures, since there is a limit to the temperatures we can explore by experiment and observation in order to guide us to the right theory. We are confident that we understand the basic physics involved in the evolution of the universe starting a few minutes after the Big Bang, and we may be able to push back to milliseconds or microseconds after it, but we cannot use the methods of science to deal with the beginning of time itself.



aa / Shock waves from the X-15 rocket plane, flying at 3.5 times the speed of sound.



ab / As in figure aa, this plane shows a shock wave. The sudden decompression of the air causes water droplets to condense, forming a cloud.

Discussion questions

A If an airplane travels at exactly the speed of sound, what would be the wavelength of the forward-emitted part of the sound waves it emitted? How should this be interpreted, and what would actually happen? What happens if it's going faster than the speed of sound? Can you use this to explain what you see in figure aa?

B If bullets go slower than the speed of sound, why can a supersonic fighter plane catch up to its own sound, but not to its own bullets?

C If someone inside a plane is talking to you, should their speech be Doppler shifted?

D The plane in figure ab was photographed when it was traveling at a speed close to the speed of sound. Comparing figures aa and ab, how can we tell from the angles of the cones that the speed is much lower in figure ab?

Summary

Selected vocabulary

superposition . . .	the adding together of waves that overlap with each other
medium	a physical substance whose vibrations constitute a wave
wavelength	the distance in space between repetitions of a periodic wave
Doppler effect . . .	the change in a wave's frequency and wavelength due to the motion of the source or the observer or both

Notation

λ	wavelength (Greek letter lambda)
---------------------	----------------------------------

Summary

Wave motion differs in three important ways from the motion of material objects:

(1) Waves obey the principle of superposition. When two waves collide, they simply add together.

(2) The medium is not transported along with the wave. The motion of any given point in the medium is a vibration about its equilibrium location, not a steady forward motion.

(3) The velocity of a wave depends on the medium, not on the amount of energy in the wave. (For some types of waves, notably water waves, the velocity may also depend on the shape of the wave.)

Sound waves consist of increases and decreases (typically very small ones) in the density of the air. Light is a wave, but it is a vibration of electric and magnetic fields, not of any physical medium. Light can travel through a vacuum.

A periodic wave is one that creates a periodic motion in a receiver as it passes it. Such a wave has a well-defined period and frequency, and it will also have a wavelength, which is the distance in space between repetitions of the wave pattern. The velocity, frequency, and wavelength of a periodic wave are related by the equation

$$v = f\lambda.$$

A wave emitted by a moving source will be shifted in wavelength and frequency. The shifted wavelength is given by the equation

$$\lambda' = \left(1 - \frac{v_s}{v}\right) \lambda,$$

where v is the velocity of the waves and v_s is the velocity of the source, taken to be positive or negative so as to produce a Doppler-lengthened wavelength if the source is receding and a Doppler-shortened one if it approaches. A similar shift occurs if the observer

is moving, and in general the Doppler shift depends approximately only on the relative motion of the source and observer if their velocities are both small compared to the waves' velocity. (This is not just approximately but exactly true for light waves, and as required by Einstein's theory of relativity.)

Problems

Key

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

1 The following is a graph of the height of a water wave as a function of *position*, at a certain moment in time.



Trace this graph onto another piece of paper, and then sketch below it the corresponding graphs that would be obtained if

- (a) the amplitude and frequency were doubled while the velocity remained the same;
- (b) the frequency and velocity were both doubled while the amplitude remained unchanged;
- (c) the wavelength and amplitude were reduced by a factor of three while the velocity was doubled.

Explain all your answers. [Problem by Arnold Arons.]

2 (a) The graph shows the height of a water wave pulse as a function of position. Draw a graph of height as a function of time for a specific point on the water. Assume the pulse is traveling to the right.

- (b) Repeat part a, but assume the pulse is traveling to the left.
- (c) Now assume the original graph was of height as a function of time, and draw a graph of height as a function of position, assuming the pulse is traveling to the right.
- (d) Repeat part c, but assume the pulse is traveling to the left.

Explain all your answers. [Problem by Arnold Arons.]



Problem 2.



Problem 3.

3 The figure shows one wavelength of a steady sinusoidal wave traveling to the right along a string. Define a coordinate system in which the positive x axis points to the right and the positive y axis up, such that the flattened string would have $y = 0$. Copy the figure, and label with $y = 0$ all the appropriate parts of the string. Similarly, label with $v = 0$ all parts of the string whose velocities are zero, and with $a = 0$ all parts whose accelerations are zero. There is more than one point whose velocity is of the greatest magnitude. Pick one of these, and indicate the direction of its velocity vector. Do the same for a point having the maximum magnitude of acceleration. Explain all your answers.

[Problem by Arnold Arons.]

4 (a) Find an equation for the relationship between the Doppler-shifted frequency of a wave and the frequency of the original wave, for the case of a stationary observer and a source moving directly toward or away from the observer. ✓

(b) Check that the units of your answer make sense.

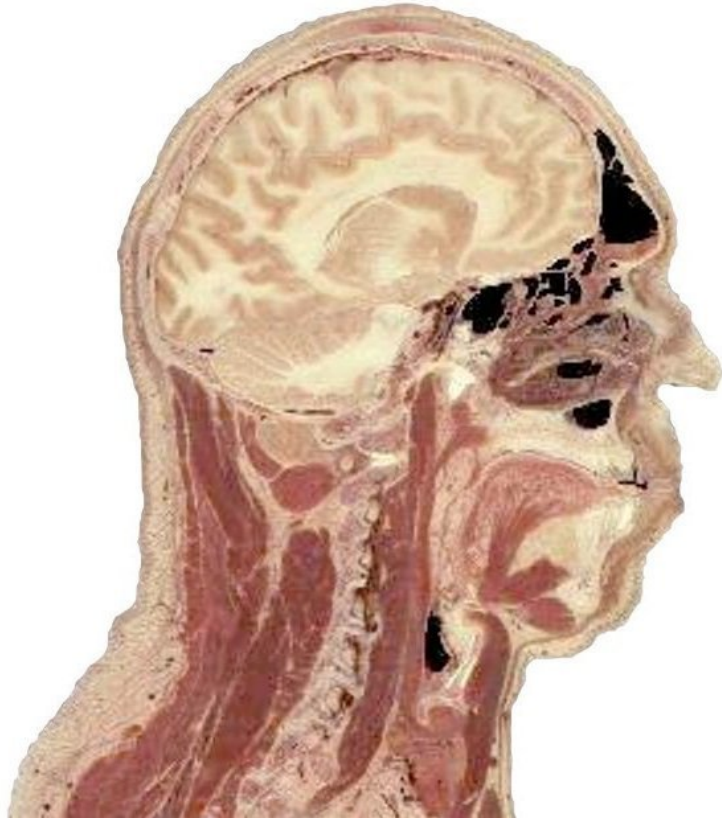
(c) Check that the dependence on v_s makes sense.

5 Suggest a quantitative experiment to look for any deviation from the principle of superposition for surface waves in water. Make it simple and practical.

6 The musical note middle C has a frequency of 262 Hz. What are its period and wavelength? ✓

7 Singing that is off-pitch by more than about 1% sounds bad. How fast would a singer have to be moving relative to the rest of a band to make this much of a change in pitch due to the Doppler effect? ✓

8 In section 19.2, we saw that the speed of waves on a string depends on the ratio of T/μ , i.e., the speed of the wave is greater if the string is under more tension, and less if it has more inertia. This is true in general: the speed of a mechanical wave always depends on the medium's inertia in relation to the restoring force (tension, stiffness, resistance to compression,...). Based on these ideas, explain why the speed of sound in air is significantly greater on a hot day, while the speed of sound in liquids and solids shows almost no variation with temperature.



A cross-sectional view of a human body, showing the vocal tract.

Chapter 20

Bounded Waves

Speech is what separates humans most decisively from animals. No other species can master syntax, and even though chimpanzees can learn a vocabulary of hand signs, there is an unmistakable difference between a human infant and a baby chimp: starting from birth, the human experiments with the production of complex speech sounds.

Since speech sounds are instinctive for us, we seldom think about them consciously. How do we control sound waves so skillfully? Mostly we do it by changing the shape of a connected set of hollow cavities in our chest, throat, and head. Somehow by moving the boundaries of this space in and out, we can produce all the vowel sounds. Up until now, we have been studying only those properties of waves that can be understood as if they existed in an infinite, open space. In this chapter we address what happens when a wave is confined within a certain space, or when a wave pattern encounters the boundary between two different media, as when a light wave moving through air encounters a glass windowpane.

a / A diver photographed this fish, and its reflection, from underwater. The reflection is the one on top, and is formed by light waves that went up to the surface of the water, but were then reflected back down into the water.



20.1 Reflection, transmission, and absorption

Reflection and transmission

Sound waves can echo back from a cliff, and light waves are reflected from the surface of a pond. We use the word reflection, normally applied only to light waves in ordinary speech, to describe any such case of a wave rebounding from a barrier. Figure b shows a circular water wave being reflected from a straight wall. In this chapter, we will concentrate mainly on reflection of waves that move in one dimension, as in figure c.

Wave reflection does not surprise us. After all, a material object such as a rubber ball would bounce back in the same way. But waves are not objects, and there are some surprises in store.

First, only part of the wave is usually reflected. Looking out through a window, we see light waves that passed through it, but a person standing outside would also be able to see her reflection in the glass. A light wave that strikes the glass is partly *reflected* and partly *transmitted* (passed) by the glass. The energy of the original wave is split between the two. This is different from the behavior of the rubber ball, which must go one way or the other, not both.

Second, consider what you see if you are swimming underwater and you look up at the surface. You see your own reflection. This is utterly counterintuitive, since we would expect the light waves to burst forth to freedom in the wide-open air. A material projectile shot up toward the surface would never rebound from the water-air

boundary! Figure a shows a similar example.

What is it about the difference between two media that causes waves to be partly reflected at the boundary between them? Is it their density? Their chemical composition? Typically what matters is the speed of the wave in the two media.¹ *A wave is partially reflected and partially transmitted at the boundary between media in which it has different speeds.* For example, the speed of light waves in window glass is about 30% less than in air, which explains why windows always make reflections. Figures d/1 and 2 show examples of wave pulses being reflected at the boundary between two coil springs of different weights, in which the wave speed is different.

Reflections such as b and c, where a wave encounters a massive fixed object, can usually be understood on the same basis as cases like d/1 and 2 later in this section, where two media meet. Example c, for instance, is like a more extreme version of example d/1. If the heavy coil spring in d/1 was made heavier and heavier, it would end up acting like the fixed wall to which the light spring in c has been attached.

self-check A

In figure c, the reflected pulse is upside-down, but its depth is just as big as the original pulse's height. How does the energy of the reflected pulse compare with that of the original? ▷ Answer, p. 568

Fish have internal ears. *example 1*

Why don't fish have ear-holes? The speed of sound waves in a fish's body is not much different from their speed in water, so sound waves are not strongly reflected from a fish's skin. They pass right through its body, so fish can have internal ears.

Whale songs traveling long distances *example 2*

Sound waves travel at drastically different speeds through rock, water, and air. Whale songs are thus strongly reflected at both the bottom and the surface. The sound waves can travel hundreds of miles, bouncing repeatedly between the bottom and the surface, and still be detectable. Sadly, noise pollution from ships has nearly shut down this cetacean version of the internet.

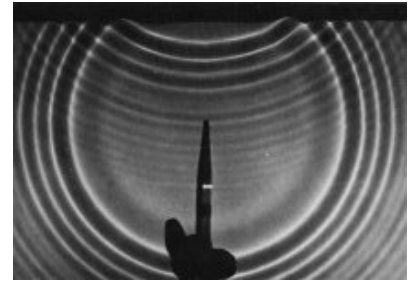
Long-distance radio communication. *example 3*

Radio communication can occur between stations on opposite sides of the planet. The mechanism is similar to the one explained in example 2, but the three media involved are the earth, the atmosphere, and the ionosphere.

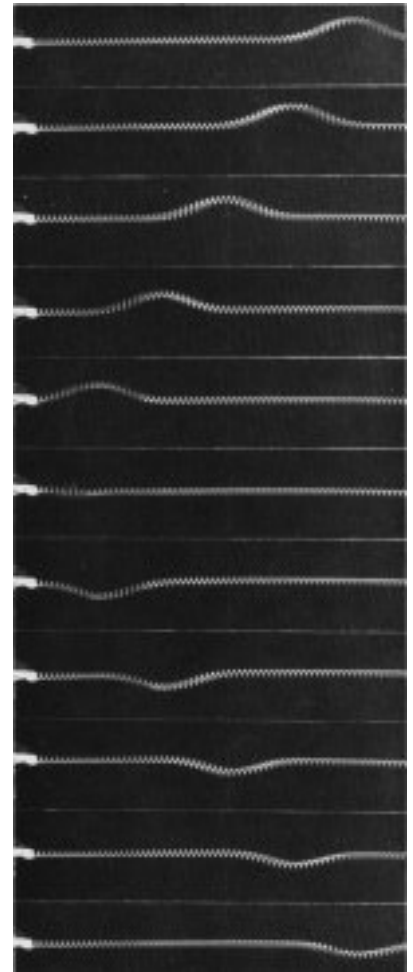
self-check B

Sonar is a method for ships and submarines to detect each other by producing sound waves and listening for echoes. What properties would an underwater object have to have in order to be invisible to sonar? ▷

¹Some exceptions are described in sec. 20.5, p. 539.



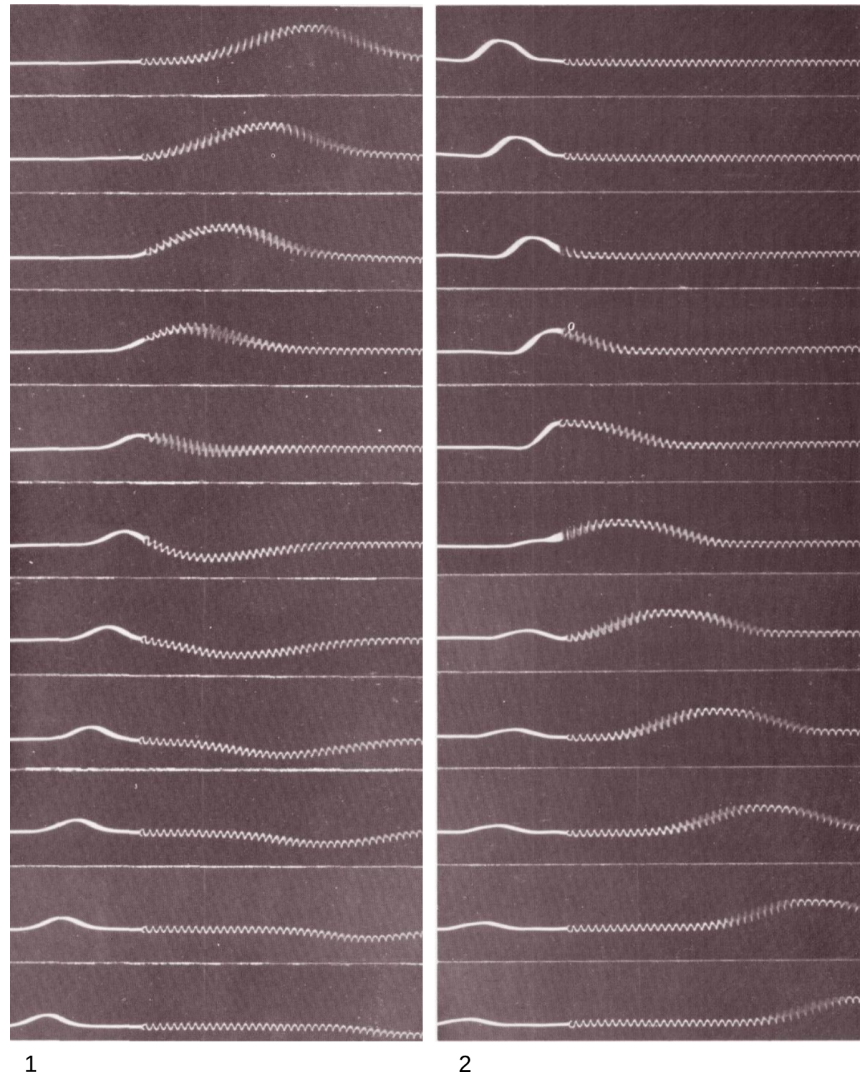
b / Circular water waves are reflected from a boundary on the top.



c / A wave on a spring, initially traveling to the left, is reflected from the fixed end.

Answer, p. 568

The use of the word “reflection” naturally brings to mind the creation of an image by a mirror, but this might be confusing, because we do not normally refer to “reflection” when we look at surfaces that are not shiny. Nevertheless, reflection is how we see the surfaces of all objects, not just polished ones. When we look at a sidewalk, for example, we are actually seeing the reflecting of the sun from the concrete. The reason we don’t see an image of the sun at our feet is simply that the rough surface blurs the image so drastically.



d / 1. A wave in the lighter spring, where the wave speed is greater, travels to the left and is then partly reflected and partly transmitted at the boundary with the heavier coil spring, which has a lower wave speed. The reflection is inverted. 2. A wave moving to the right in the heavier spring is partly reflected at the boundary with the lighter spring. The reflection is uninverted.

Inverted and uninverted reflections

Notice how the pulse reflected back to the right in example d/1 comes back upside-down, whereas the one reflected back to the left in 2 returns in its original upright form. This is true for other waves as well. In general, there are two possible types of reflections, a reflection back into a faster medium and a reflection back into a slower medium. One type will always be an inverting reflection and one noninverting.

It's important to realize that when we discuss inverted and uninverted reflections on a string, we are talking about whether the wave is flipped across the direction of motion (i.e., upside-down in these drawings). The reflected pulse will always be reversed front to back, as shown in figure e. This is because it is traveling in the other direction. The leading edge of the pulse is what gets reflected first, so it is still ahead when it starts back to the left — it's just that "ahead" is now in the opposite direction.

Absorption

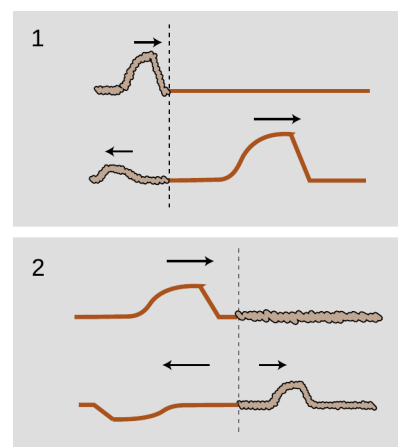
So far we have tacitly assumed that wave energy remains as wave energy, and is not converted to any other form. If this was true, then the world would become more and more full of sound waves, which could never escape into the vacuum of outer space. In reality, any mechanical wave consists of a traveling pattern of vibrations of some physical medium, and vibrations of matter always produce heat, as when you bend a coat-hangar back and forth and it becomes hot. We can thus expect that in mechanical waves such as water waves, sound waves, or waves on a string, the wave energy will gradually be converted into heat. This is referred to as *absorption*.

The wave suffers a decrease in amplitude, as shown in figure f. The decrease in amplitude amounts to the same fractional change for each unit of distance covered. For example, if a wave decreases from amplitude 2 to amplitude 1 over a distance of 1 meter, then after traveling another meter it will have an amplitude of $1/2$. That is, the reduction in amplitude is exponential. This can be proven as follows. By the principle of superposition, we know that a wave of amplitude 2 must behave like the superposition of two identical waves of amplitude 1. If a single amplitude-1 wave would die down to amplitude $1/2$ over a certain distance, then two amplitude-1 waves superposed on top of one another to make amplitude $1 + 1 = 2$ must die down to amplitude $1/2 + 1/2 = 1$ over the same distance.

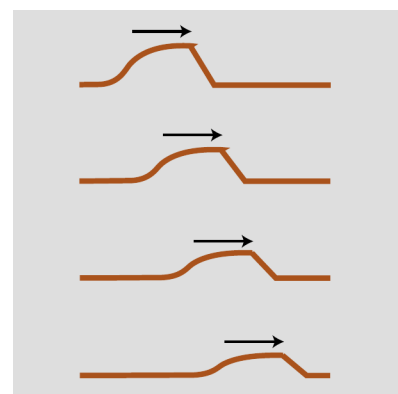
self-check C

As a wave undergoes absorption, it loses energy. Does this mean that it slows down? ▷ Answer, p. 568

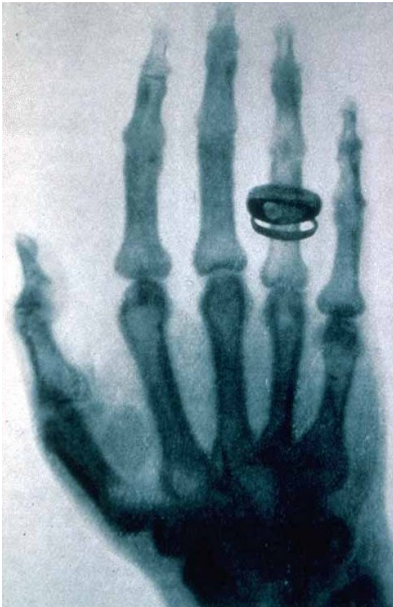
In many cases, this frictional heating effect is quite weak. Sound waves in air, for instance, dissipate into heat extremely slowly, and the sound of church music in a cathedral may reverberate for as much



e / 1. An uninverted reflection. The reflected pulse is reversed front to back, but is not upside-down. 2. An inverted reflection. The reflected pulse is reversed both front to back and top to bottom.



f / A pulse traveling through a highly absorptive medium.



g / X-rays are light waves with a very high frequency. They are absorbed strongly by bones, but weakly by flesh.

as 3 or 4 seconds before it becomes inaudible. During this time it has traveled over a kilometer! Even this very gradual dissipation of energy occurs mostly as heating of the church's walls and by the leaking of sound to the outside (where it will eventually end up as heat). Under the right conditions (humid air and low frequency), a sound wave in a straight pipe could theoretically travel hundreds of kilometers before being noticeably attenuated.

In general, the absorption of mechanical waves depends a great deal on the chemical composition and microscopic structure of the medium. Ripples on the surface of antifreeze, for instance, die out extremely rapidly compared to ripples on water. For sound waves and surface waves in liquids and gases, what matters is the viscosity of the substance, i.e., whether it flows easily like water or mercury or more sluggishly like molasses or antifreeze. This explains why our intuitive expectation of strong absorption of sound in water is incorrect. Water is a very weak absorber of sound (viz. whale songs and sonar), and our incorrect intuition arises from focusing on the wrong property of the substance: water's high density, which is irrelevant, rather than its low viscosity, which is what matters.

Light is an interesting case, since although it can travel through matter, it is not itself a vibration of any material substance. Thus we can look at the star Sirius, 10^{14} km away from us, and be assured that none of its light was absorbed in the vacuum of outer space during its 9-year journey to us. The Hubble Space Telescope routinely observes light that has been on its way to us since the early history of the universe, billions of years ago. Of course the energy of light can be dissipated if it does pass through matter (and the light from distant galaxies is often absorbed if there happen to be clouds of gas or dust in between).

Soundproofing

example 4

Typical amateur musicians setting out to soundproof their garages tend to think that they should simply cover the walls with the densest possible substance. In fact, sound is not absorbed very strongly even by passing through several inches of wood. A better strategy for soundproofing is to create a sandwich of alternating layers of materials in which the speed of sound is very different, to encourage reflection.

The classic design is alternating layers of fiberglass and plywood. The speed of sound in plywood is very high, due to its stiffness, while its speed in fiberglass is essentially the same as its speed in air. Both materials are fairly good sound absorbers, but sound waves passing through a few inches of them are still not going to be absorbed sufficiently. The point of combining them is that a sound wave that tries to get out will be strongly reflected at each of the fiberglass-plywood boundaries, and will bounce back and forth many times like a ping pong ball. Due to all the back-

and-forth motion, the sound may end up traveling a total distance equal to ten times the actual thickness of the soundproofing before it escapes. This is the equivalent of having ten times the thickness of sound-absorbing material.

The swim bladder

example 5

The swim bladder of a fish, which was first discussed in homework problem 2 in chapter 18, is often located right next to the fish's ear. As discussed in example 1 on page 523, the fish's body is nearly transparent to sound, so it's actually difficult to get any of the sound wave energy to deposit itself in the fish so that the fish can hear it! The physics here is almost exactly the same as the physics of example 4 above, with the gas-filled swim bladder playing the role of the low-density material.

Radio transmission

example 6

A radio transmitting station, such as a commercial station or an amateur "ham" radio station, must have a length of wire or cable connecting the amplifier to the antenna. The cable and the antenna act as two different media for radio waves, and there will therefore be partial reflection of the waves as they come from the cable to the antenna. If the waves bounce back and forth many times between the amplifier and the antenna, a great deal of their energy will be absorbed. There are two ways to attack the problem. One possibility is to design the antenna so that the speed of the waves in it is as close as possible to the speed of the waves in the cable; this minimizes the amount of reflection. The other method is to connect the amplifier to the antenna using a type of wire or cable that does not strongly absorb the waves. Partial reflection then becomes irrelevant, since all the wave energy will eventually exit through the antenna.

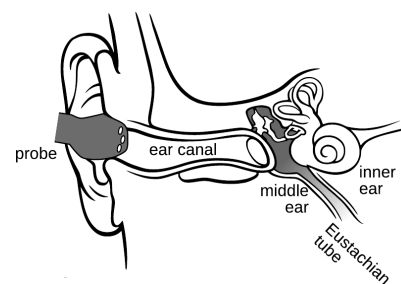
The tympanogram

example 7

The tympanogram is a medical procedure used to diagnose problems with the middle ear.

The middle ear is a chamber, normally filled with air, lying between the eardrum (tympanic membrane) and the inner ear. It contains a tiny set of bones that act as a system of levers to amplify the motion of the eardrum and transmit it to the inner ear. The air pressure in the inner ear is normally equalized via the Eustachian tube, which connects to the throat; when you feel uncomfortable pressure in your ear while flying, it's because the pressure has not yet equalized. Ear infections or allergies can cause the middle ear to become filled with fluid, and the Eustachian tube can also become blocked, so that the pressure in the inner ear cannot become equalized.

The tympanometer has a probe that is inserted into the ear, with several holes. One hole is used to send a 226 Hz sound wave into



h / A tympanometer, example 7.

the ear canal. The ear has evolved so as to transmit a maximum amount of wave motion to the inner ear. Any change in its physical properties will change its behavior from its normal optimum, so that more sound energy than normal is reflected back. A second hole in the probe senses the reflected wave. If the reflection is stronger than normal, there is probably something wrong with the inner ear.

The full physical analysis is fairly complex. The middle ear has some of the characteristics of a mass oscillating on a spring, but it also has some of the characteristics of a medium that carries waves. Crudely, we could imagine that an infected, fluid-filled middle ear would act as a medium that differed greatly from the air in the outer ear, causing a large amount of reflection.

Equally crudely, we could forget about the wave ideas and think of the middle ear purely as a mass on a spring. We expect resonant behavior, and there is in fact such a resonance, which is typically at a frequency of about 600 Hz in adults, so the 226 Hz frequency emitted by the probe is actually quite far from resonance. If the mechanisms of the middle ear are jammed and cannot vibrate, then it is not possible for energy of the incoming sound wave to be turned into energy of vibration in the middle ear, and therefore by conservation of energy we would expect all of the sound to be reflected.

Sometimes the middle ear's mechanisms can get jammed because of abnormally high or low pressure, because the Eustachian tube is blocked and cannot equalize the pressure with the outside environment. Diagnosing such a condition is the purpose of the third hole in the probe, which is used to vary the pressure in the ear canal. The amount of reflection is measured as a function of this pressure. If the reflection is minimized for some value of the pressure that is different than atmospheric pressure, it indicates that that is the value of the pressure in the middle ear; when the pressures are equalized, the forces on the eardrum cancel out, and it can relax to its normal position, unjamming the middle ear's mechanisms.

Discussion question

A A sound wave that underwent a pressure-inverting reflection would have its compressions converted to expansions and vice versa. How would its energy and frequency compare with those of the original sound? Would it sound any different? What happens if you swap the two wires where they connect to a stereo speaker, resulting in waves that vibrate in the opposite way?

20.2 Quantitative treatment of reflection

In this section we use the example of waves on a string to analyze the reasons why a reflection occurs at the boundary between media, predict quantitatively the intensities of reflection and transmission, and discuss how to tell which reflections are inverting and which are noninverting. Some more technical aspects of the discussion are relegated to sec. 20.5, p. 539.

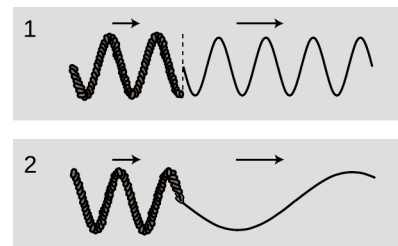
Why reflection occurs

To understand the fundamental reasons for what does occur at the boundary between media, let's first discuss what doesn't happen. For the sake of concreteness, consider a sinusoidal wave on a string. If the wave progresses from a heavier portion of the string, in which its velocity is low, to a lighter-weight part, in which it is high, then the equation $v = f\lambda$ tells us that it must change its frequency, or its wavelength, or both. If only the frequency changed, then the parts of the wave in the two different portions of the string would quickly get out of step with each other, producing a discontinuity in the wave, $i/1$. This is unphysical, so we know that the wavelength must change while the frequency remains constant, 2.

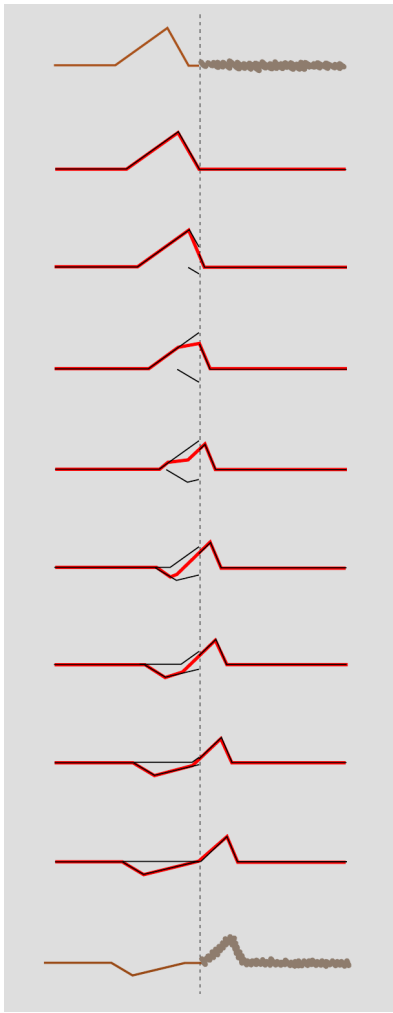
But there is still something unphysical about figure 2. The sudden change in the shape of the wave has resulted in a sharp kink at the boundary. This can't really happen, because the medium tends to accelerate in such a way as to eliminate curvature. A sharp kink corresponds to an infinite curvature at one point, which would produce an infinite acceleration, which would not be consistent with the smooth pattern of wave motion envisioned in figure 2. Waves can have kinks, but not stationary kinks.

We conclude that without positing partial reflection of the wave, we cannot simultaneously satisfy the requirements of (1) continuity of the wave, and (2) no sudden changes in the slope of the wave. (The student who has studied calculus will recognize this as amounting to an assumption that both the wave and its derivative are continuous functions.)

Does this amount to a proof that reflection occurs? Not quite. We have only proven that certain types of wave motion are not valid solutions. In the following subsection, we prove that a valid solution can always be found in which a reflection occurs. Now in physics, we normally assume (but seldom prove formally) that the equations of motion have a unique solution, since otherwise a given set of initial conditions could lead to different behavior later on, but the Newtonian universe is supposed to be deterministic. Since the solution must be unique, and we derive below a valid solution involving a reflected pulse, we will have ended up with what amounts to a proof of reflection.



1. A change in frequency without a change in wavelength would produce a discontinuity in the wave. 2. A simple change in wavelength without a reflection would result in a sharp kink in the wave.



j / A pulse being partially reflected and partially transmitted at the boundary between two strings in which the speed of waves is different. The top drawing shows the pulse heading to the right, toward the heavier string. For clarity, all but the first and last drawings are schematic. Once the reflected pulse begins to emerge from the boundary, it adds together with the trailing parts of the incident pulse. Their sum, shown as a wider line, is what is actually observed.

Intensity of reflection

We will now show, in the case of waves on a string, that it is possible to satisfy the physical requirements given above by constructing a reflected wave, and as a bonus this will produce an equation for the proportions of reflection and transmission and a prediction as to which conditions will lead to inverted and which to uninverted reflection. We assume only that the principle of superposition holds, which is a good approximation for waves on a string of sufficiently small amplitude.

Let the unknown amplitudes of the reflected and transmitted waves be R and T , respectively. An inverted reflection would be represented by a negative value of R . We can without loss of generality take the incident (original) wave to have unit amplitude. Superposition tells us that if, for instance, the incident wave had double this amplitude, we could immediately find a corresponding solution simply by doubling R and T .

Just to the left of the boundary, the height of the wave is given by the height 1 of the incident wave, plus the height R of the part of the reflected wave that has just been created and begun heading back, for a total height of $1 + R$. On the right side immediately next to the boundary, the transmitted wave has a height T . To avoid a discontinuity, we must have

$$1 + R = T.$$

Next we turn to the requirement of equal slopes on both sides of the boundary. Let the slope of the incoming wave be s immediately to the left of the junction. If the wave was 100% reflected, and without inversion, then the slope of the reflected wave would be $-s$, since the wave has been reversed in direction. In general, the slope of the reflected wave equals $-sR$, and the slopes of the superposed waves on the left side add up to $s - sR$. On the right, the slope depends on the amplitude, T , but is also changed by the stretching or compression of the wave due to the change in speed. If, for example, the wave speed is twice as great on the right side, then the slope is cut in half by this effect. The slope on the right is therefore $s(v_1/v_2)T$, where v_1 is the velocity in the original medium and v_2 the velocity in the new medium. Equality of slopes gives $s - sR = s(v_1/v_2)T$, or

$$1 - R = \frac{v_1}{v_2}T.$$

Solving the two equations for the unknowns R and T gives

$$R = \frac{v_2 - v_1}{v_2 + v_1} \quad \text{and} \quad T = \frac{2v_2}{v_2 + v_1}.$$

The first equation shows that there is no reflection unless the two wave speeds are different, and that the reflection is inverted in reflection back into a fast medium.

The energies of the transmitted and reflected waves always add up to the same as the energy of the original wave. There is never any abrupt loss (or gain) in energy when a wave crosses a boundary. (Conversion of wave energy to heat occurs for many types of waves, but it occurs throughout the medium.) The equation for T , surprisingly, allows the amplitude of the transmitted wave to be greater than 1, i.e., greater than that of the incident wave. This does not violate conservation of energy, because this occurs when the second string is less massive, reducing its kinetic energy, and the transmitted pulse is broader and less strongly curved, which lessens its potential energy. In other words, the constant of proportionality in $E \propto A^2$ is different in the two different media.

We have attempted to develop some general facts about wave reflection by using the specific example of a wave on a string, which raises the question of whether these facts really are general. These issues are discussed in more detail in optional section 20.5, p. 539, but here is a brief summary.

The following facts are more generally true for wave reflection in one dimension.

- The wave is partially reflected and partially transmitted, with the reflected and transmitted parts sharing the energy.
- For an interface between media 1 and 2, there are two possible reflections: back into 1, and back into 2. One of these is inverting ($R < 0$) and the other is noninverting ($R > 0$).

The following aspects of our analysis may need to be modified for different types of waves.

- In some cases, the expressions for the reflected and transmitted amplitudes depend not on the ratio v_1/v_2 but on some more complicated ratio $v_1 \dots / v_2 \dots$, where \dots stands for some additional property of the medium.
- The sign of R , depends not just on this ratio but also on the type of the wave and on what we choose as a measure of amplitude.

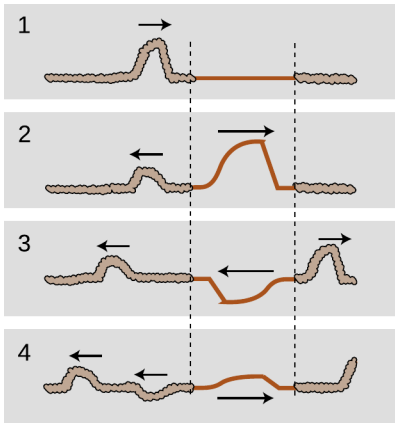


k / Seen from this angle, the optical coating on the lenses of these binoculars appears purple and green. (The color varies depending on the angle from which the coating is viewed, and the angle varies across the faces of the lenses because of their curvature.)

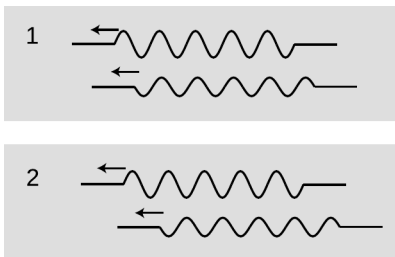
20.3 Interference effects

If you look at the front of a pair of high-quality binoculars, you will notice a greenish-blue coating on the lenses. This is advertised as a coating to prevent reflection. Now reflection is clearly undesirable — we want the light to go *in* the binoculars — but so far I've described reflection as an unalterable fact of nature, depending only on the properties of the two wave media. The coating can't change the speed of light in air or in glass, so how can it work? The key is that the coating itself is a wave medium. In other words, we have a three-layer sandwich of materials: air, coating, and glass. We will analyze the way the coating works, not because optical coatings are an important part of your education but because it provides a good example of the general phenomenon of wave interference effects.

There are two different interfaces between media: an air-coating boundary and a coating-glass boundary. Partial reflection and partial transmission will occur at each boundary. For ease of visualization let's start by considering an equivalent system consisting of three dissimilar pieces of string tied together, and a wave pattern consisting initially of a single pulse. Figure l/1 shows the incident pulse moving through the heavy rope, in which its velocity is low. When it encounters the lighter-weight rope in the middle, a faster medium, it is partially reflected and partially transmitted. (The transmitted pulse is bigger, but nevertheless has only part of the original energy.) The pulse transmitted by the first interface is then partially reflected and partially transmitted by the second boundary, 3. In figure 4, two pulses are on the way back out to the left, and a single pulse is heading off to the right. (There is still a weak pulse caught between the two boundaries, and this will rattle back and forth, rapidly getting too weak to detect as it leaks energy to the outside with each partial reflection.)



l / A rope consisting of three sections, the middle one being lighter.



m / Two reflections, are superimposed. One reflection is inverted.

Note how, of the two reflected pulses in 4, one is inverted and one uninverted. One underwent reflection at the first boundary (a reflection back into a slower medium is uninverted), but the other was reflected at the second boundary (reflection back into a faster medium is inverted).

Now let's imagine what would have happened if the incoming wave pattern had been a long sinusoidal wave train instead of a single pulse. The first two waves to reemerge on the left could be in phase, m/1, or out of phase, 2, or anywhere in between. The amount of lag between them depends entirely on the width of the middle segment of string. If we choose the width of the middle string segment correctly, then we can arrange for destructive interference to occur, 2, with cancellation resulting in a very weak reflected wave.

This whole analysis applies directly to our original case of optical coatings. Visible light from most sources does consist of a stream

of short sinusoidal wave-trains such as the ones drawn above. The only real difference between the waves-on-a-rope example and the case of an optical coating is that the first and third media are air and glass, in which light does not have the same speed. However, the general result is the same as long as the air and the glass have light-wave speeds that are either both greater than the coating's or both less than the coating's.

The business of optical coatings turns out to be a very arcane one, with a plethora of trade secrets and “black magic” techniques handed down from master to apprentice. Nevertheless, the ideas you have learned about waves in general are sufficient to allow you to come to some definite conclusions without any further technical knowledge. The self-check and discussion questions will direct you along these lines of thought.

The example of an optical coating was typical of a wide variety of wave interference effects. With a little guidance, you are now ready to figure out for yourself other examples such as the rainbow pattern made by a compact disc, a layer of oil on a puddle, or a soap bubble.

self-check D

1. Color corresponds to wavelength of light waves. Is it possible to choose a thickness for an optical coating that will produce destructive interference for all colors of light?
2. How can you explain the rainbow colors on the soap bubble in figure n?

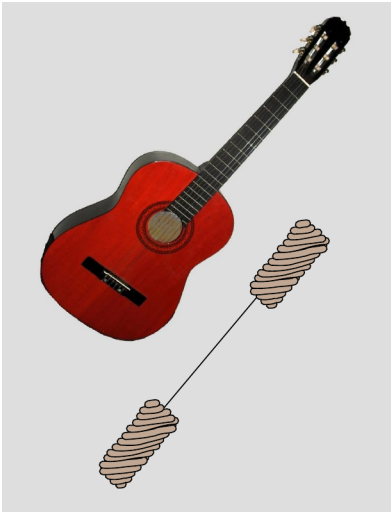
▷ Answer, p. 568

Discussion questions

- A** Is it possible to get *complete* destructive interference in an optical coating, at least for light of one specific wavelength?
- B** Sunlight consists of sinusoidal wave-trains containing on the order of a hundred cycles back-to-back, for a length of something like a tenth of a millimeter. What happens if you try to make an optical coating thicker than this?
- C** Suppose you take two microscope slides and lay one on top of the other so that one of its edges is resting on the corresponding edge of the bottom one. If you insert a sliver of paper or a hair at the opposite end, a wedge-shaped layer of air will exist in the middle, with a thickness that changes gradually from one end to the other. What would you expect to see if the slides were illuminated from above by light of a single color? How would this change if you gradually lifted the lower edge of the top slide until the two slides were finally parallel?
- D** An observation like the one described in discussion question C was used by Newton as evidence *against* the wave theory of light! If Newton didn't know about inverting and noninverting reflections, what would have seemed inexplicable to him about the region where the air layer had zero or nearly zero thickness?



n / A soap bubble displays interference effects.

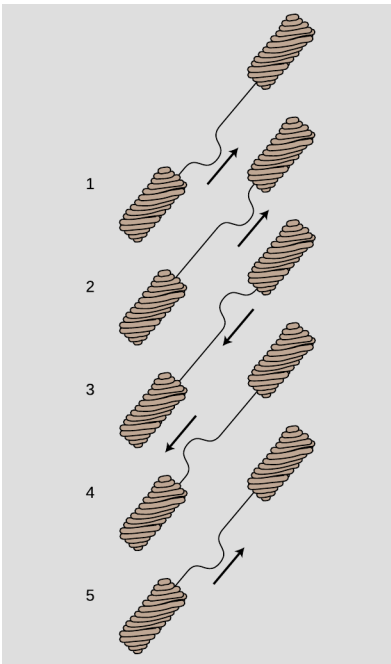


o / A model of a guitar string.

20.4 Waves bounded on both sides

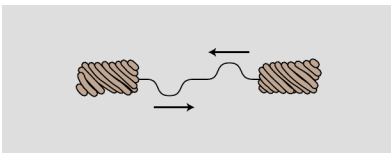
In the examples discussed in section 20.3, it was theoretically true that a pulse would be trapped permanently in the middle medium, but that pulse was not central to our discussion, and in any case it was weakening severely with each partial reflection. Now consider a guitar string. At its ends it is tied to the body of the instrument itself, and since the body is very massive, the behavior of the waves when they reach the end of the string can be understood in the same way as if the actual guitar string was attached on the ends to strings that were extremely massive, ∞ . Reflections are most intense when the two media are very dissimilar. Because the wave speed in the body is so radically different from the speed in the string, we should expect nearly 100% reflection.

Although this may seem like a rather bizarre physical model of the actual guitar string, it already tells us something interesting about the behavior of a guitar that we would not otherwise have understood. The body, far from being a passive frame for attaching the strings to, is actually the exit path for the wave energy in the strings. With every reflection, the wave pattern on the string loses a tiny fraction of its energy, which is then conducted through the body and out into the air. (The string has too little cross-section to make sound waves efficiently by itself.) By changing the properties of the body, moreover, we should expect to have an effect on the manner in which sound escapes from the instrument. This is clearly demonstrated by the electric guitar, which has an extremely massive, solid wooden body. Here the dissimilarity between the two wave media is even more pronounced, with the result that wave energy leaks out of the string even more slowly. This is why an electric guitar with no electric pickup can hardly be heard at all, and it is also the reason why notes on an electric guitar can be sustained for longer than notes on an acoustic guitar.



p / The motion of a pulse on the string.

If we initially create a disturbance on a guitar string, how will the reflections behave? In reality, the finger or pick will give the string a triangular shape before letting it go, and we may think of this triangular shape as a very broad “dent” in the string which will spread out in both directions. For simplicity, however, let’s just imagine a wave pattern that initially consists of a single, narrow pulse traveling up the neck, $p/1$. After reflection from the top end, it is inverted, 3. Now something interesting happens: figure 5 is identical to figure 1. After two reflections, the pulse has been inverted twice and has changed direction twice. It is now back where it started. The motion is periodic. This is why a guitar produces sounds that have a definite sensation of pitch.



q / A tricky way to double the frequency.

self-check E

Notice that from $p/1$ to $p/5$, the pulse has passed by every point on the string exactly twice. This means that the total distance it has traveled

equals $2L$, where L is the length of the string. Given this fact, what are the period and frequency of the sound it produces, expressed in terms of L and v , the velocity of the wave? ▷ Answer, p. 569

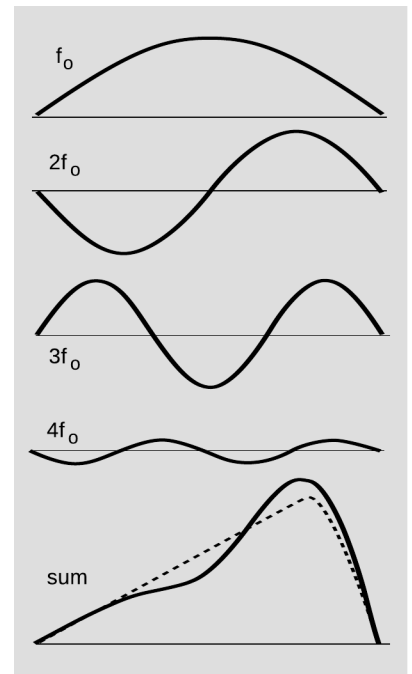
Note that if the waves on the string obey the principle of superposition, then the velocity must be independent of amplitude, and the guitar will produce the same pitch regardless of whether it is played loudly or softly. In reality, waves on a string obey the principle of superposition approximately, but not exactly. The guitar, like just about any acoustic instrument, is a little out of tune when played loudly. (The effect is more pronounced for wind instruments than for strings, but wind players are able to compensate for it.)

Now there is only one hole in our reasoning. Suppose we somehow arrange to have an initial setup consisting of two identical pulses heading toward each other, as in figure q. They will pass through each other, undergo a single inverting reflection, and come back to a configuration in which their positions have been exactly interchanged. This means that the period of vibration is half as long. The frequency is twice as high.

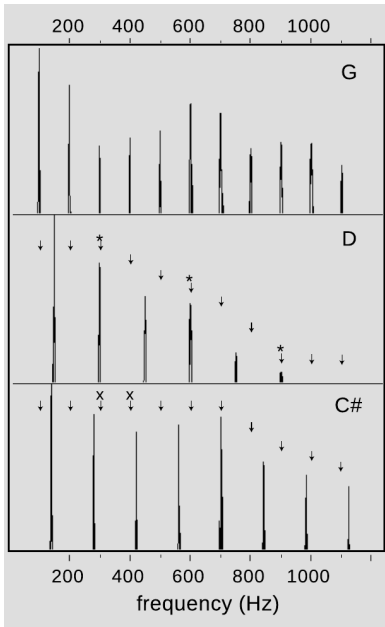
This might seem like a purely academic possibility, since nobody actually plays the guitar with two picks at once! But in fact it is an example of a very general fact about waves that are bounded on both sides. A mathematical theorem called Fourier's theorem states that any wave can be created by superposing sine waves. Figure r shows how even by using only four sine waves with appropriately chosen amplitudes, we can arrive at a sum which is a decent approximation to the realistic triangular shape of a guitar string being plucked. The one-hump wave, in which half a wavelength fits on the string, will behave like the single pulse we originally discussed. We call its frequency f_o . The two-hump wave, with one whole wavelength, is very much like the two-pulse example. For the reasons discussed above, its frequency is $2f_o$. Similarly, the three-hump and four-hump waves have frequencies of $3f_o$ and $4f_o$.

Theoretically we would need to add together infinitely many such wave patterns to describe the initial triangular shape of the string exactly, although the amplitudes required for the very high frequency parts would be very small, and an excellent approximation could be achieved with as few as ten waves.

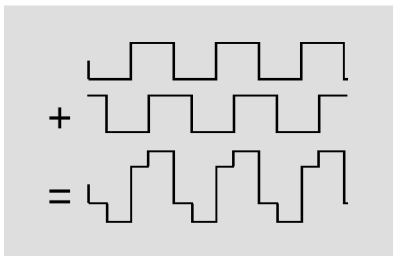
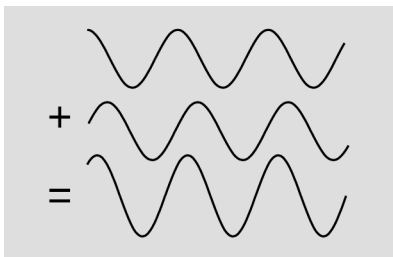
We thus arrive at the following very general conclusion. Whenever a wave pattern exists in a medium bounded on both sides by media in which the wave speed is very different, the motion can be broken down into the motion of a (theoretically infinite) series of sine waves, with frequencies $f_o, 2f_o, 3f_o, \dots$ Except for some technical details, to be discussed below, this analysis applies to a vast range of sound-producing systems, including the air column within the human vocal tract. Because sounds composed of this kind of pattern of frequencies are so common, our ear-brain system has evolved so



r / Using the sum of four sine waves to approximate the triangular initial shape of a plucked guitar string.



s / Graphs of loudness versus frequency for the vowel “ah,” sung as three different musical notes. G is consonant with D, since every overtone of G that is close to an overtone of D (*) is at exactly the same frequency. G and C# are dissonant together, since some of the overtones of G (x) are close to, but not right on top of, those of C#.



t / Sine waves add to make sine waves. Other functions don’t have this property.

as to perceive them as a single, fused sensation of tone.

Musical applications

Many musicians claim to be able to pick out by ear several of the frequencies $2f_o$, $3f_o$, ..., called overtones or *harmonics* of the fundamental f_o , but they are kidding themselves. In reality, the overtone series has two important roles in music, neither of which depends on this fictitious ability to “hear out” the individual overtones.

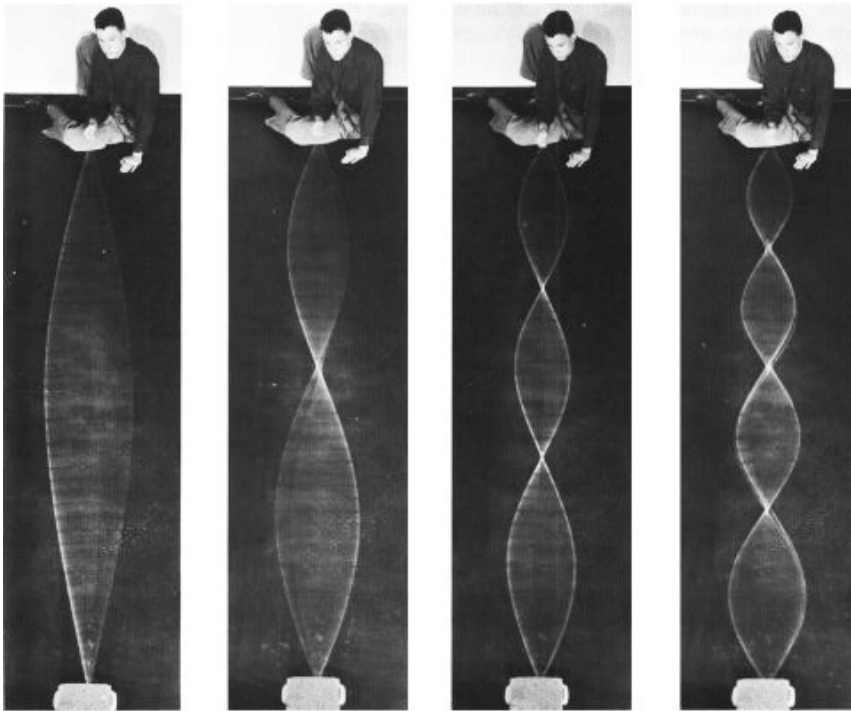
First, the relative strengths of the overtones is an important part of the personality of a sound, called its timbre (rhymes with “amber”). The characteristic tone of the brass instruments, for example, is a sound that starts out with a very strong harmonic series extending up to very high frequencies, but whose higher harmonics die down drastically as the attack changes to the sustained portion of the note.

Second, although the ear cannot separate the individual harmonics of a single musical tone, it is very sensitive to clashes between the overtones of notes played simultaneously, i.e., in harmony. We tend to perceive a combination of notes as being dissonant if they have overtones that are close but not the same. Roughly speaking, strong overtones whose frequencies differ by more than 1% and less than 10% cause the notes to sound dissonant. It is important to realize that the term “dissonance” is not a negative one in music. No matter how long you search the radio dial, you will never hear more than three seconds of music without at least one dissonant combination of notes. Dissonance is a necessary ingredient in the creation of a musical cycle of tension and release. Musically knowledgeable people don’t use the word “dissonant” as a criticism of music, although dissonance can be used in a clumsy way, or without providing any contrast between dissonance and consonance.

Standing waves

Figure u shows sinusoidal wave patterns made by shaking a rope. I used to enjoy doing this at the bank with the pens on chains, back in the days when people actually went to the bank. You might think that I and the person in the photos had to practice for a long time in order to get such nice sine waves. In fact, a sine wave is the only shape that can create this kind of wave pattern, called a standing wave, which simply vibrates back and forth in one place without moving. The sine wave just creates itself automatically when you find the right frequency, because no other shape is possible.

If you think about it, it’s not even obvious that sine waves should be able to do this trick. After all, waves are supposed to travel at a set speed, aren’t they? The speed isn’t supposed to be zero! Well, we can actually think of a standing wave as a superposition of a moving



u / Standing waves on a spring.

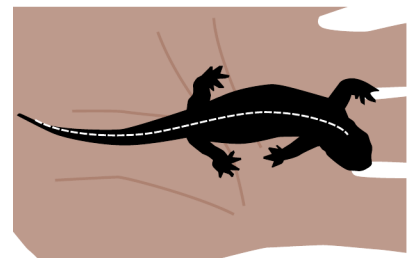
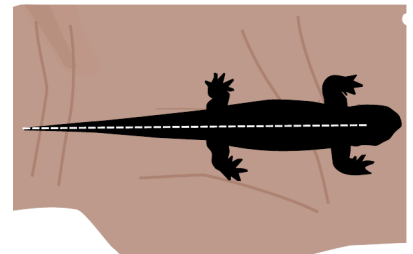
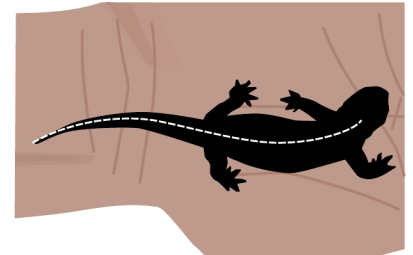
sine wave with its own reflection, which is moving the opposite way. Sine waves have the unique mathematical property, t , that the sum of sine waves of equal wavelength is simply a new sine wave with the same wavelength. As the two sine waves go back and forth, they always cancel perfectly at the ends, and their sum appears to stand still.

Standing wave patterns are rather important, since atoms are really standing-wave patterns of electron waves. You are a standing wave!

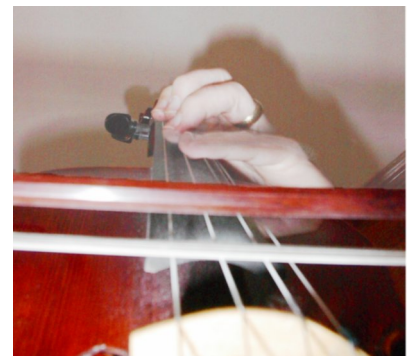
Harmonics on string instruments

example 8

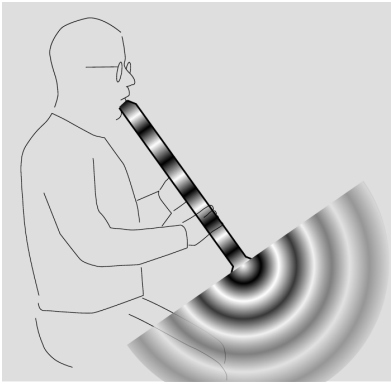
Figure w shows a violist playing what string players refer to as a natural harmonic. The term “harmonic” is used here in a somewhat different sense than in physics. The musician’s pinkie is pressing very lightly against the string — not hard enough to make it touch the fingerboard — at a point precisely at the center of the string’s length. As shown in the diagram, this allows the string to vibrate at frequencies $2f_0, 4f_0, 6f_0, \dots$, which have stationary points at the center of the string, but not at the odd multiples $f_0, 3f_0, \dots$. Since all the overtones are multiples of $2f_0$, the ear perceives $2f_0$ as the basic frequency of the note. In musical terms, doubling the frequency corresponds to raising the pitch by an octave. The technique can be used in order to make it easier to play high notes in rapid passages, or for its own sake, because of the change in timbre.



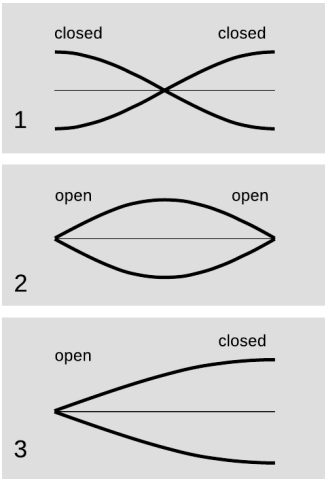
v / A salamander crawls across a person’s palm. Its spine oscillates as a standing wave.



w / Example 8.



x / Surprisingly, sound waves undergo partial reflection at the open ends of tubes as well as closed ones.



y / Graphs of excess density versus position for the lowest-frequency standing waves of three types of air columns. Points on the axis have normal air density.

Standing-wave patterns of air columns

The air column inside a wind instrument behaves very much like the wave-on-a-string example we've been concentrating on so far, the main difference being that we may have either inverting or noninverting reflections at the ends.

Some organ pipes are closed at both ends. The speed of sound is different in metal than in air, so there is a strong reflection at the closed ends, and we can have standing waves. These reflections are both density-noninverting, so we get symmetric standing-wave patterns, such as the one shown in figure y/1.

Figure x shows the sound waves in and around a bamboo Japanese flute called a shakuhachi, which is *open* at both ends of the air column. We can only have a standing wave pattern if there are reflections at the ends, but that is very counterintuitive — why is there any reflection at all, if the sound wave is free to emerge into open space, and there is no change in medium? Recall the reason why we got reflections at a change in medium: because the wavelength changes, so the wave has to readjust itself from one pattern to another, and the only way it can do that without developing a kink is if there is a reflection. Something similar is happening here. The only difference is that the wave is adjusting from being a plane wave to being a spherical wave. The reflections at the open ends are density-inverting, y/2, so the wave pattern is pinched off at the ends. Comparing panels 1 and 2 of the figure, we see that although the wave patterns are different, in both cases the wavelength is the same: in the lowest-frequency standing wave, half a wavelength fits inside the tube. Thus, it isn't necessary to memorize which type of reflection is inverting and which is uninverting. It's only necessary to know that the tubes are symmetric.

Finally, we can have an asymmetric tube: closed at one end and open at the other. A common example is the pan pipes, z, which are closed at the bottom and open at the top. The standing wave with the lowest frequency is therefore one in which $1/4$ of a wavelength fits along the length of the tube, as shown in figure y/3.

Sometimes an instrument's physical appearance can be misleading. A concert flute, aa, is closed at the mouth end and open at the other, so we would expect it to behave like an asymmetric air column; in reality, it behaves like a symmetric air column open at both ends, because the embouchure hole (the hole the player blows over) acts like an open end. The clarinet and the saxophone look similar, having a mouthpiece and reed at one end and an open end at the other, but they act different. In fact the clarinet's air column has patterns of vibration that are asymmetric, the saxophone symmetric. The discrepancy comes from the difference between the conical tube of the sax and the cylindrical tube of the clarinet. The adjustment of the wave pattern from a plane wave to a spherical

wave is more gradual at the flaring bell of the saxophone.

self-check F

Draw a graph of density versus position for the first overtone of the air column in a tube open at one end and closed at the other. This will be the next-to-longest possible wavelength that allows for a point of maximum vibration at one end and a point of no vibration at the other. How many times shorter will its wavelength be compared to the wavelength of the lowest-frequency standing wave, shown in the figure? Based on this, how many times greater will its frequency be? ▷ Answer, p. 569

Discussion question

A Figure v on p. 537 shows the salamander in the salamander's frame of reference, so that the palm moves. In the palm's frame, would this be a traveling wave? Would the worm in example 1 on p. 497 execute a standing wave in its own frame? Is there anything qualitatively different about these two animals' patterns of motion, other than the fact that one wave is transverse and the other longitudinal?

20.5 ★ Some technical aspects of reflection

In this optional section we address some technical details of the treatment of reflection and transmission of waves. These gory details are likely to be of interest mainly to students majoring in the physical sciences.

Dependence of reflection on other variables besides velocity

In section 20.2 we derived the expressions for the transmitted and reflected amplitudes by demanding that two things match up on both sides of the boundary: the height of the wave and the slope of the wave. These requirements were stated purely in terms of kinematics (the description of how the wave moves) rather than dynamics (the explanation for the wave motion in terms of Newton's laws). For this reason, the results depended only on the purely kinematic quantity $\alpha = v_2/v_1$, as can be seen more clearly if we rewrite the expressions in the following form:

$$R = \frac{\alpha - 1}{\alpha + 1} \quad \text{and} \quad T = \frac{2\alpha}{\alpha + 1}.$$

But this purely kinematical treatment only worked because of a dynamical fact that we didn't emphasize. We assumed equality of the slopes, $s_1 = s_2$, because waves don't like to have kinks. The underlying dynamical reason for this, in the case of a wave on a string, is that a kink is pointlike, so the portion of the string at the kink is infinitesimal in size, and therefore has essentially zero mass. If the transverse forces acting on it differed by some finite amount, then its acceleration would be infinite, which is not possible. The difference between the two forces is $Ts_1 - Ts_2$, so $s_1 = s_2$. But this relies on the assumption that T is the same on both sides of the



z / A pan pipe is an asymmetric air column, open at the top and closed at the bottom.



aa / A concert flute looks like an asymmetric air column, open at the mouth end and closed at the other. However, its patterns of vibration are symmetric, because the embouchure hole acts like an open end.

boundary. Now this is true, because we can't put different amounts of tension on two ropes that are tied together end to end. Any excess tension applied to one rope is distributed equally to the other. For other types of waves, however, we cannot make a similar argument, and therefore it need not be true that $s_1 = s_2$.

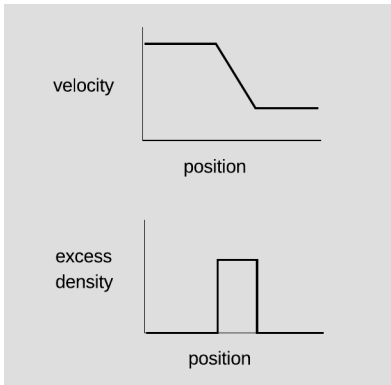
A more detailed analysis shows that in general we have not $\alpha = v_2/v_1$ but $\alpha = z_2/z_1$, where z is a quantity called impedance which is defined for this purpose. In a great many examples, as for the waves on a string, it is true that $v_2/v_1 = z_2/z_1$, but this is not a universal fact. Most of the exceptions are rather specialized and technical, such as the reflection of light waves when the media have magnetic properties, but one fairly common and important example is the case of sound waves, for which $z = \rho v$ depends not just on the wave velocity v but also on the density ρ . A practical example occurs in medical ultrasound scans, where the contrast of the image is made possible because of the very large differences in impedance between different types of tissue. The speed of sound in various tissues such as bone and muscle varies by about a factor of 2, which is not a particularly huge factor, but there are also large variations in density. The lung, for example, is basically a sponge or sack filled with air. For this reason, the acoustic impedances of the tissues show a huge amount of variation, with, e.g., $z_{\text{bone}}/z_{\text{lung}} \approx 40$.

Inverted and uninverted reflections in general

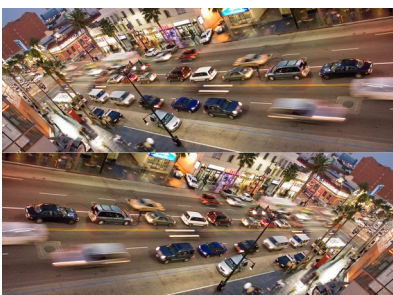
For waves on a string, reflections back into a faster medium are inverted, while those back into a slower medium are uninverted. Is this true for all types of waves? The rather subtle answer is that it depends on what property of the wave you are discussing.

Let's start by considering wave disturbances of freeway traffic. Anyone who has driven frequently on crowded freeways has observed the phenomenon in which one driver taps the brakes, starting a chain reaction that travels backward down the freeway as each person in turn exercises caution in order to avoid rear-ending anyone. The reason why this type of wave is relevant is that it gives a simple, easily visualized example of how our description of a wave depends on which aspect of the wave we have in mind. In steadily flowing freeway traffic, both the density of cars and their velocity are constant all along the road. Since there is no disturbance in this pattern of constant velocity and density, we say that there is no wave. Now if a wave is touched off by a person tapping the brakes, we can either describe it as a region of high density or as a region of decreasing velocity.

The freeway traffic wave is in fact a good model of a sound wave, and a sound wave can likewise be described either by the density (or pressure) of the air or by its speed. Likewise many other types of waves can be described by either of two functions, one of which is often the derivative of the other with respect to position.



ab / A disturbance in freeway traffic.



ac / In the mirror image, the areas of positive excess traffic density are still positive, but the velocities of the cars have all been reversed, so areas of positive excess velocity have been turned into negative ones.

Now let's consider reflections. If we observe the freeway wave in a mirror, the high-density area will still appear high in density, but velocity in the opposite direction will now be described by a negative number. A person observing the mirror image will draw the same density graph, but the velocity graph will be flipped across the x axis, and its original region of negative slope will now have positive slope. Although I don't know any physical situation that would correspond to the reflection of a traffic wave, we can immediately apply the same reasoning to sound waves, which often do get reflected, and determine that a reflection can either be density-inverting and velocity-noninverting or density-noninverting and velocity-inverting.

This same type of situation will occur over and over as one encounters new types of waves, and to apply the analogy we need only determine which quantities, like velocity, become negated in a mirror image and which, like density, stay the same.

A light wave, for instance, consists of a traveling pattern of electric and magnetic fields. All you need to know in order to analyze the reflection of light waves is how electric and magnetic fields behave under reflection; you don't need to know any of the detailed physics of electricity and magnetism. An electric field can be detected, for example, by the way one's hair stands on end. The direction of the hair indicates the direction of the electric field. In a mirror image, the hair points the other way, so the electric field is apparently reversed in a mirror image. The behavior of magnetic fields, however, is a little tricky. The magnetic properties of a bar magnet, for instance, are caused by the aligned rotation of the outermost orbiting electrons of the atoms. In a mirror image, the direction of rotation is reversed, say from clockwise to counterclockwise, and so the magnetic field is reversed twice: once simply because the whole picture is flipped and once because of the reversed rotation of the electrons. In other words, magnetic fields do not reverse themselves in a mirror image. We can thus predict that there will be two possible types of reflection of light waves. In one, the electric field is inverted and the magnetic field uninverted (example 8, p. 712). In the other, the electric field is uninverted and the magnetic field inverted.

Summary

Selected vocabulary

reflection	the bouncing back of part of a wave from a boundary
transmission	the continuation of part of a wave through a boundary
absorption	the gradual conversion of wave energy into heating of the medium
standing wave	a wave pattern that stays in one place

Notation

λ	wavelength (Greek letter lambda)
---------------------	----------------------------------

Summary

Whenever a wave encounters the boundary between two media in which its speeds are different, part of the wave is reflected and part is transmitted. The reflection is always reversed front-to-back, but may also be inverted in amplitude. Whether the reflection is inverted depends on how the wave speeds in the two media compare, e.g., a wave on a string is uninverted when it is reflected back into a segment of string where its speed is lower. The greater the difference in wave speed between the two media, the greater the fraction of the wave energy that is reflected. Surprisingly, a wave in a dense material like wood will be strongly reflected back into the wood at a wood-air boundary.

A one-dimensional wave confined by highly reflective boundaries on two sides will display motion which is periodic. For example, if both reflections are inverting, the wave will have a period equal to twice the time required to traverse the region, or to that time divided by an integer. An important special case is a sinusoidal wave; in this case, the wave forms a stationary pattern composed of a superposition of sine waves moving in opposite direction.

Problems

Key

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

1 Light travels faster in warmer air. On a sunny day, the sun can heat a road and create a layer of hot air above it. Let's model this layer as a uniform one with a sharp boundary separating it from the cooler air above. Use this model to explain the formation of a mirage appearing like the shiny surface of a pool of water.

2 (a) Compute the amplitude of light that is reflected back into air at an air-water interface, relative to the amplitude of the incident wave. The speeds of light in air and water are 3.0×10^8 and 2.2×10^8 m/s, respectively.

(b) Find the energy of the reflected wave as a fraction of the incident energy. [Hint: The answers to the two parts are not the same.]

✓

3 A concert flute produces its lowest note, at about 262 Hz, when half of a wavelength fits inside its tube. Compute the length of the flute.

▷ Answer, p. 569

4 (a) A good tenor saxophone player can play all of the following notes without changing her fingering, simply by altering the tightness of her lips: E♭ (150 Hz), E♭ (300 Hz), B♭ (450 Hz), and E♭ (600 Hz). How is this possible? (I'm not asking you to analyze the coupling between the lips, the reed, the mouthpiece, and the air column, which is very complicated.)

(b) Some saxophone players are known for their ability to use this technique to play "freak notes," i.e., notes above the normal range of the instrument. Why isn't it possible to play notes below the normal range using this technique?

C	261.6 Hz
D	293.7
E	329.6
F	349.2
G	392.0
A	440.0
B♭	466.2

5 The table gives the frequencies of the notes that make up the key of F major, starting from middle C and going up through all seven notes.

(a) Calculate the first four or five harmonics of C and G, and determine whether these two notes will be consonant or dissonant. (Recall that harmonics that differ by about 1-10% cause dissonance.)

(b) Do the same for C and B♭.

Problem 5.

6 Brass and wind instruments go up in pitch as the musician warms up. As a typical empirical example, a trumpet's frequency might go up by about 1%. Let's consider possible physical reasons for the change in pitch. (a) Solids generally expand with increasing temperature, because the stronger random motion of the atoms tends to bump them apart. Brass expands by 1.88×10^{-5} per degree C. Would this tend to raise the pitch, or lower it? Estimate the size of the effect in comparison with the observed change in frequency. (b) The speed of sound in a gas is proportional to the square root of the absolute temperature, where zero absolute temperature is -273 degrees C. As in part a, analyze the size and direction of the effect.

7 Your exhaled breath contains about 4.5% carbon dioxide, and is therefore more dense than fresh air by about 2.3%. By analogy with the treatment of waves on a string in section 19.2, we expect that the speed of sound will be inversely proportional to the square root of the density of the gas. Calculate the effect on the frequency produced by a wind instrument.

8 The expressions for the amplitudes of reflected and transmitted waves depend on the unitless ratio v_2/v_1 (or, more generally, on the ratio of the impedances). Call this ratio α . (a) Show that changing α to $1/\alpha$ (e.g., by interchanging the roles of the two media) has an effect on the reflected amplitude that can be expressed in a simple way, and discuss what this means in terms of inversion and energy. (b) Find the two values of α for which $|R| = 1/2$.

Three essential mathematical skills

More often than not when a search-and-rescue team finds a hiker dead in the wilderness, it turns out that the person was not carrying some item from a short list of essentials, such as water and a map. There are three mathematical essentials in this course.

1. Converting units

basic technique: section 0.9, p. 29; conversion of area, volume, etc.: section 1.1, p. 41

Examples:

$$0.7 \cancel{\text{kg}} \times \frac{10^3 \text{ g}}{1 \cancel{\text{kg}}} = 700 \text{ g}.$$

To check that we have the conversion factor the right way up (10^3 rather than $1/10^3$), we note that the smaller unit of grams has been *compensated* for by making the number larger.

For units like m^2 , kg/m^3 , etc., we have to raise the conversion factor to the appropriate power:

$$4 \text{ m}^3 \times \left(\frac{10^3 \text{ mm}}{1 \text{ m}} \right)^3 = 4 \times 10^9 \cancel{\text{m}^3} \times \frac{\text{mm}^3}{\cancel{\text{m}^3}} = 4 \times 10^9 \text{ mm}^3$$

Examples with solutions — p. 36, #6; p. 59, #10

Problems you can check at lightandmatter.com/area1checker.html — p. 36, #5; p. 36, #4; p. 36, #7; p. 59, #1; p. 60, #19

2. Reasoning about ratios and proportionalities

The technique is introduced in section 1.2, p. 43, in the context of area and volume, but it applies more generally to any relationship in which one variable depends on another raised to some power.

Example: When a car or truck travels over a road, there is wear and tear on the road surface, which incurs a cost. Studies show that the cost per kilometer of travel C is given by

$$C = kw^4,$$

where w is the weight per axle and k is a constant. The weight per axle is about 13 times higher for a semi-trailer than for my Honda Fit. How many times greater is the cost imposed on the federal government when the semi travels a given distance on an interstate freeway?

▷ First we convert the equation into a proportionality by throwing out k , which is the same for both vehicles:

$$C \propto w^4$$

Next we convert this proportionality to a statement about ratios:

$$\frac{C_1}{C_2} = \left(\frac{w_1}{w_2} \right)^4 \approx 29,000$$

Since the gas taxes paid by the trucker are nowhere near 29,000 times more than those I pay to drive my Fit the same distance, the federal government is effectively awarding a massive subsidy to the trucking company. Plus my Fit is cuter.

Examples with solutions — p. 59, #11; p. 59, #12; p. 60, #17; p. 119, #16; p. 120, #22; p. 256, #6; p. 282, #10; p. 283, #15; p. 285, #19; p. 311, #8; p. 311, #9

Problems you can check at lightandmatter.com/area1checker.html — p. 60, #16; p. 60, #18; p. 61, #23; p. 62, #24; p. 62, #25; p. 201, #7; p. 256, #8; p. 281, #5; p. 282, #8; p. 285, #21; p. 310, #4; p. 424, #2

3. Vector addition

section 7.3, p. 210

Example: The $\Delta \mathbf{r}$ vector from San Diego to Los Angeles has magnitude 190 km and direction 129° counterclockwise from east. The one from LA to Las Vegas is 370 km at 38° counterclockwise from east. Find the distance and direction from San Diego to Las Vegas.

▷ Graphical addition is discussed on p. 210. Here we concentrate on analytic addition, which involves adding the x components to find the total x component, and similarly for y . The trig needed in order to find the components of the second leg (LA to Vegas) is laid out in figure e on p. 207 and explained in detail in example 3 on p. 207:

$$\Delta x_2 = (370 \text{ km}) \cos 38^\circ = 292 \text{ km}$$

$$\Delta y_2 = (370 \text{ km}) \sin 38^\circ = 228 \text{ km}$$

(Since these are intermediate results, we keep an extra sig fig to avoid accumulating too much rounding error.) Once we understand the trig for one example, we don't need to reinvent the wheel every time. The pattern is completely universal, provided that we first make sure to get the angle expressed according to the usual trig convention, counterclockwise from the x axis. Applying the pattern to the first leg, we have:

$$\Delta x_1 = (190 \text{ km}) \cos 129^\circ = -120 \text{ km}$$

$$\Delta y_1 = (190 \text{ km}) \sin 129^\circ = 148 \text{ km}$$

For the vector directly from San Diego to Las Vegas, we have

$$\Delta x = \Delta x_1 + \Delta x_2 = 172 \text{ km}$$

$$\Delta y = \Delta y_1 + \Delta y_2 = 376 \text{ km.}$$

The distance from San Diego to Las Vegas is found using the Pythagorean theorem,

$$\sqrt{(172 \text{ km})^2 + (376 \text{ km})^2} = 410 \text{ km}$$

(rounded to two sig figs because it's one of our final results). The direction is one of the two possible values of the inverse tangent

$$\tan^{-1}(\Delta y/\Delta x) = \{65^\circ, 245^\circ\}.$$

Consulting a sketch shows that the first of these values is the correct one.

Examples with solutions — p. 234, #8; p. 234, #9; p. 389, #8

Problems you can check at lightandmatter.com/area1checker.html — p. 216, #3; p. 216, #4; p. 233, #1; p. 233, #3; p. 236, #16; p. 281, #3; p. 286, #23; p. 388, #3

Hints for volume 1

Hints for chapter 10

Page 284, problem 17:

If you try to calculate the two forces and subtract, your calculator will probably give a result of zero due to rounding. Instead, reason about the fractional amount by which the quantity $1/r^2$ will change. As a warm-up, you may wish to observe the percentage change in $1/r^2$ that results from changing r from 1 to 1.01.

Hints for chapter 15

Page 427, problem 18:

The first two parts can be done more easily by setting $a = 1$, since the value of a only changes the distance scale. One way to do part b is by graphing.

Solutions to selected problems for volume 1

Solutions for chapter 0

Page 36, problem 6:

$$134 \text{ mg} \times \frac{10^{-3} \text{ g}}{1 \text{ mg}} \times \frac{10^{-3} \text{ kg}}{1 \text{ g}} = 1.34 \times 10^{-4} \text{ kg}$$

Page 37, problem 8:

(a) Let's do 10.0 g and 1000 g. The arithmetic mean is 505 grams. It comes out to be 0.505 kg, which is consistent. (b) The geometric mean comes out to be 100 g or 0.1 kg, which is consistent. (c) If we multiply meters by meters, we get square meters. Multiplying grams by grams should give square grams! This sounds strange, but it makes sense. Taking the square root of square grams (g^2) gives grams again. (d) No. The superduper mean of two quantities with units of grams wouldn't even be something with units of grams! Related to this shortcoming is the fact that the superduper mean would fail the kind of consistency test carried out in the first two parts of the problem.

Page 37, problem 12:

(a) They're all defined in terms of the ratio of side of a triangle to another. For instance, the tangent is the length of the opposite side over the length of the adjacent side. Dividing meters by meters gives a unitless result, so the tangent, as well as the other trig functions, is unitless. (b) The tangent function gives a unitless result, so the units on the right-hand side had better cancel out. They do, because the top of the fraction has units of meters squared, and so does the bottom.

Solutions for chapter 1

Page 59, problem 10:

$$1 \text{ mm}^2 \times \left(\frac{1 \text{ cm}}{10 \text{ mm}} \right)^2 = 10^{-2} \text{ cm}^2$$

Page 59, problem 11:

The bigger scope has a diameter that's ten times greater. Area scales as the square of the linear

dimensions, so $A \propto d^2$, or in the language of ratios $A_1/A_2 = (d_1/d_2)^2 = 100$. Its light-gathering power is a hundred times greater.

Page 59, problem 12:

Since they differ by two steps on the Richter scale, the energy of the bigger quake is 10^4 times greater. The wave forms a hemisphere, and the surface area of the hemisphere over which the energy is spread is proportional to the square of its radius, $A \propto r^2$, or $r \propto \sqrt{A}$, which means $r_1/r_2 = \sqrt{A_1/A_2}$. If the amount of vibration was the same, then the surface areas must be in the ratio $A_1/A_2 = 10^4$, which means that the ratio of the radii is 10^2 .

Page 60, problem 17:

The cone of mixed gin and vermouth is the same shape as the cone of vermouth, but its linear dimensions are doubled. Translating the proportionality $V \propto L^3$ into an equation about ratios, we have $V_1/V_2 = (L_1/L_2)^3 = 8$. Since the ratio of the whole thing to the vermouth is 8, the ratio of gin to vermouth is 7.

Page 60, problem 19:

The proportionality $V \propto L^3$ can be restated in terms of ratios as $V_1/V_2 = (L_1/L_2)^3 = (1/10)^3 = 1/1000$, so scaling down the linear dimensions by a factor of $1/10$ reduces the volume by $1/1000$, to a milliliter.

Page 61, problem 21:

Let's estimate the Great Wall's volume, and then figure out how many bricks that would represent. The wall is famous because it covers pretty much all of China's northern border, so let's say it's 1000 km long. From pictures, it looks like it's about 10 m high and 10 m wide, so the total volume would be $10^6 \text{ m} \times 10 \text{ m} \times 10 \text{ m} = 10^8 \text{ m}^3$. If a single brick has a volume of 1 liter, or 10^{-3} m^3 , then this represents about 10^{11} bricks. If one person can lay 10 bricks in an hour (taking into account all the preparation, etc.), then this would be 10^{10} man-hours.

Page 62, problem 26:

Directly guessing the number of jelly beans would be like guessing volume directly. That would be a mistake. Instead, we start by estimating the linear dimensions, in units of beans. The contents of the jar look like they're about 10 beans deep. Although the jar is a cylinder, its exact geometrical shape doesn't really matter for the purposes of our order-of-magnitude estimate. Let's pretend it's a rectangular jar. The horizontal dimensions are also something like 10 beans, so it looks like the jar has about $10 \times 10 \times 10$ or $\sim 10^3$ beans inside.

Solutions for chapter 2

Page 89, problem 4:

$$\begin{aligned} 1 \text{ light-year} &= v\Delta t \\ &= (3 \times 10^8 \text{ m/s}) (1 \text{ year}) \\ &= (3 \times 10^8 \text{ m/s}) \left[(1 \text{ year}) \times \left(\frac{365 \text{ days}}{1 \text{ year}} \right) \times \left(\frac{24 \text{ hours}}{1 \text{ day}} \right) \times \left(\frac{3600 \text{ s}}{1 \text{ hour}} \right) \right] \\ &= 9.5 \times 10^{15} \text{ m} \end{aligned}$$

Page 89, problem 5:

Velocity is relative, so having to lean tells you nothing about the current value of the train's velocity. Fullerton is moving at a huge speed relative to Beijing, but that doesn't produce any

noticeable effect in either city. The fact that you have to lean tells you that the train is *changing* its speed, but it doesn't tell you what the train's current speed is.

Page 89, problem 7:

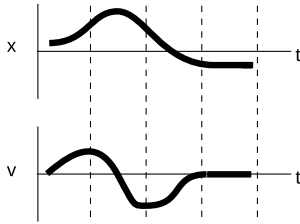
To the person riding the moving bike, bug A is simply going in circles. The only difference between the motions of the two wheels is that one is traveling through space, but motion is relative, so this doesn't have any effect on the bugs. It's equally hard for each of them.

Page 90, problem 10:

In one second, the ship moves v meters to the east, and the person moves v meters north relative to the deck. Relative to the water, he traces the diagonal of a triangle whose length is given by the Pythagorean theorem, $(v^2 + v^2)^{1/2} = \sqrt{2}v$. Relative to the water, he is moving at a 45-degree angle between north and east.

Solutions for chapter 3

Page 119, problem 14:



Page 119, problem 15:

Taking g to be 10 m/s^2 , the bullet loses 10 m/s of speed every second, so it will take 10 s to come to a stop, and then another 10 s to come back down, for a total of 20 s .

Page 119, problem 16:

$\Delta x = \frac{1}{2}at^2$, so for a fixed value of Δx , we have $t \propto 1/\sqrt{a}$. Translating this into the language of ratios gives $t_M/t_E = \sqrt{a_E/a_M} = \sqrt{3} = 1.7$.

Page 119, problem 17:

$$\begin{aligned} v &= \frac{dx}{dt} \\ &= 10 - 3t^2 \\ a &= \frac{dv}{dt} \\ &= -6t \\ &= -18 \text{ m/s}^2 \end{aligned}$$

Page 120, problem 18:

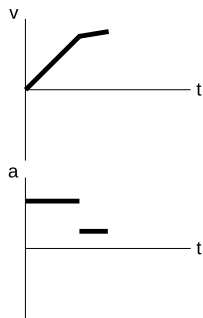
(a) Solving for $\Delta x = \frac{1}{2}at^2$ for a , we find $a = 2\Delta x/t^2 = 5.51 \text{ m/s}^2$. (b) $v = \sqrt{2a\Delta x} = 66.6 \text{ m/s}$. (c) The actual car's final velocity is less than that of the idealized constant-acceleration car. If the real car and the idealized car covered the quarter mile in the same time but the real car was moving more slowly at the end than the idealized one, the real car must have been going faster than the idealized car at the beginning of the race. The real car apparently has a greater acceleration at the beginning, and less acceleration at the end. This makes sense, because every car has some maximum speed, which is the speed beyond which it cannot accelerate.

Page 120, problem 19:

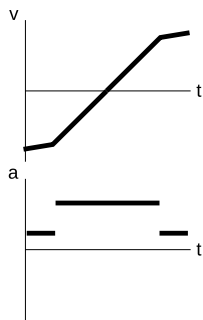
Since the lines are at intervals of one m/s and one second, each box represents one meter. From $t = 0$ to $t = 2$ s, the area under the curve represents a positive Δx of 6 m. (The triangle has half the area of the 2×6 rectangle it fits inside.) After $t = 2$ s, the area above the curve represents negative Δx . To get -6 m worth of area, we need to go out to $t = 6$ s, at which point the triangle under the axis has a width of 4 s and a height of 3 m/s, for an area of 6 m (half of 3×4).

Page 120, problem 20:

(a) We choose a coordinate system with positive pointing to the right. Some people might expect that the ball would slow down once it was on the more gentle ramp. This may be true if there is significant friction, but Galileo's experiments with inclined planes showed that when friction is negligible, a ball rolling on a ramp has constant acceleration, not constant speed. The speed stops increasing as quickly once the ball is on the more gentle slope, but it still keeps on increasing. The a-t graph can be drawn by inspecting the slope of the v-t graph.



(b) The ball will roll back down, so the second half of the motion is the same as in part a. In the first (rising) half of the motion, the velocity is negative, since the motion is in the opposite direction compared to the positive x axis. The acceleration is again found by inspecting the slope of the v-t graph.



Page 120, problem 21:

This is a case where it's probably easiest to draw the acceleration graph first. While the ball is in the air (bc, de, etc.), the only force acting on it is gravity, so it must have the same, constant acceleration during each hop. Choosing a coordinate system where the positive x axis points up, this becomes a negative acceleration (force in the opposite direction compared to the axis). During the short times between hops when the ball is in contact with the ground (cd, ef, etc.), it experiences a large acceleration, which turns around its velocity very rapidly. These short positive accelerations probably aren't constant, but it's hard to know how they'd really look. We just idealize them as constant accelerations. Similarly, the hand's force on the ball during the time ab is probably not constant, but we can draw it that way, since we don't know