Chapter 7

2D waves and other topics

David Morin, morin@physics.harvard.edu

This chapter is fairly short. In Section 7.1 we derive the wave equation for two-dimensional waves, and we discuss the patterns that arise with vibrating membranes and plates. In Section 7.2 we discuss the Doppler effect, which is relevant when the source of the wave and/or the observer are/is moving through the medium in which the wave is traveling.

7.1 2D waves on a membrane

We studied transverse waves on a one-dimensional string in Chapter 4. Let's now look at transverse waves on a two-dimensional membrane, for example a soap film with a wire boundary. Let the equilibrium position of the membrane be the x-y plane. So z is the transverse direction (we'll use z here instead of our customary ψ). Consider a little rectangle in the x-y plane with sides Δx and Δy. During the wave motion, the patch of the membrane corresponding to this rectangle will be displaced in the z direction. But since we are assuming (as always) that the transverse displacement is small (more precisely, that the slope of the membrane is small), this patch is still approximately a rectangle. But it is slightly curved, and it is this curvature that causes there to be a net transverse force, just as was the case for the 1-D string. The only difference, as we'll shortly see, is that we have "double" the effect because the membrane is two dimensional.

As with the 1-D string, the smallness of the slope implies that all points in the membrane move essentially only in the transverse direction; there is no motion parallel to the x-y plane. This implies that the mass of the slightly-tilted patch is always essentially equal to $\sigma\Delta x\Delta y$, where σ is the mass density per unit area.

Let the surface tension be S. The units of surface tension are force/length. If you draw a line segment of length $d\ell$ on the membrane, then the force that the membrane on one side of the line exerts on the membrane on the other side is $S\,d\ell$. So the forces on the sides of the little patch are $S\,\Delta x$ and $S\,\Delta y$. A view of the patch, looking along the y direction, is shown in Fig. 1. This profile looks exactly like the picture we had in the 1-D case (see Fig. 4.2). So just as in that case, the difference in the slope at the two ends is what causes there to be a net force in the z direction (at least as far as the sides at x and $x+\Delta x$ go; there is a similar net force from the other two sides).[1] The net force in the z direction due to the

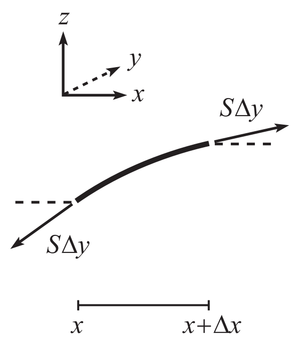

Figure 1

[1] As you look along the y axis, the patch won't look exactly like the 1-D curved segment shown in Fig. 1, because in general there is curvature in the y direction too. But this fact won't change our results.

forces shown in Fig. 1 is

$$\begin{aligned} F &= S\,\Delta y\big(z'(x+\Delta x) - z'(x)\big) \\ &= S\,\Delta y \Delta x \frac{z'(x+\Delta x) - z'(x)}{\Delta x} \\ &\approx S\,\Delta y \Delta x \frac{\partial^2 z}{\partial x^2}\,. \end{aligned} \tag{1}$$

We haven't bothered writing the arguments of the function $z(x, y, t)$.

We can do the same thing by looking at the profile along the x direction, and we find that the net force from the two sides at y and $y + \Delta y$ is $S\,\Delta x \Delta y(\partial^2 z/\partial x^2)$. The total transverse force is the sum of these two results. And since the mass of the patch is $\sigma \Delta x \Delta y$, the transverse $F = ma$ equation (or rather the $ma = F$ equation) for the patch is

$$\begin{aligned} (\sigma \Delta x \Delta y)\frac{\partial^2 z}{\partial t^2} &= S\,\Delta x \Delta y \left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right) \\ \Longrightarrow \quad \boxed{\frac{\partial^2 z}{\partial t^2} = \frac{S}{\sigma}\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right)} & \qquad \textit{(wave equation)} \end{aligned} \tag{2}$$

This looks quite similar to our old 1-D wave equation in Eq. (4.4), except that we now have partial derivatives with respect to two spatial coordinates. How do we solve this equation for the function $z(x, y, t)$? We know that any function can be written in terms of its Fourier components. Since we have three independent variables, the Fourier decomposition of $z(x, y, t)$ consists of the triple integral,

$$z(x, y, t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(k_x, k_y, \omega) e^{i(k_x x + k_y y + \omega t)} dk_x\, dk_y\, d\omega. \tag{3}$$

This comes about in the same way that the double integral came about in Eq. (4.16). And since the wave equation in Eq. (2) is linear, it suffices to guess solutions of the form $e^{i(k_x x + k_y y + \omega t)}$. (Both Fourier analysis and linearity are necessary for this conclusion to hold). Plugging this guess into Eq. (2) and canceling a minus sign gives

$$\boxed{\omega^2 = \frac{S}{\sigma}\big(k_x^2 + k_y^2\big)} \qquad \textit{(dispersion relation)} \tag{4}$$

This looks basically the same as the 1-D dispersion relation for transverse waves on a string, $\omega^2 = c^2 k^2$, but with the simple addition of a second k^2 term. However, this seemingly minor modification has a huge consequence: In the 1-D case, only one k value corresponded to a given ω value. But in the 2-D case, an infinite number of k_x and k_y values correspond to a given ω value, namely all the (k_x, k_y) points on a circle of radius $\omega\sqrt{\sigma/S}$.

Let's now look at some boundary conditions. Things get very complicated with arbitrarily-shaped boundaries, so let's consider the case of a rectangular boundary. We can imagine having a soap film stretched across a rectangular wire boundary. Let the sides be parallel to the coordinate axes and have lengths L_x and L_y, and let one corner be located at the origin. The boundary condition for the membrane is that $z = 0$ on the boundary, because the membrane must be in contact with the wire. Let's switch from exponential solutions to trig solutions, which work much better here. We can write the trig solutions in many ways, but we'll choose the basis where $z(x, y, t)$ takes the form,

$$z(x, y, t) = A\,\text{trig}(k_x x)\,\text{trig}(k_y y)\,\text{trig}(\omega t), \tag{5}$$

where "trig" means either sine or cosine. Similar to the 1-D case, the $x = 0$ and $y = 0$ boundaries tell us that we can't have any cosine functions of x and y. So the solution must take the form,

$$z(x, y, t) = A \sin(k_x x) \sin(k_y y) \cos(\omega t + \phi). \tag{6}$$

And again similar to the 1-D case, the boundary conditions at $x = L_x$ and $y = L_y$ restrict k_x and k_y to satisfy

$$k_x x = n\pi, \quad k_y y = m\pi \quad \Longrightarrow \quad k_x = \frac{n\pi}{L_x}, \quad k_x = \frac{n\pi}{L_x}. \tag{7}$$

The most general solution for z is an arbitrary sum of these basis solutions, so we have

$$\begin{aligned} z(x, y, t) &= \sum_{n,m} A_{n,m} \sin\left(\frac{n\pi x}{L_x}\right) \sin\left(\frac{m\pi y}{L_y}\right) \cos(\omega_{n,m} t + \phi_{n,m}), \\ \text{where} \qquad \omega_{n,m}^2 &= \frac{S}{\sigma}\left[\left(\frac{n\pi}{L_x}\right)^2 + \left(\frac{m\pi}{L_y}\right)^2\right]. \end{aligned} \tag{8}$$

Each basis solution (that is, each normal mode) in this sum is a standing wave. The constants $A_{n,m}$ and $\phi_{n,m}$ are determined by the initial conditions. If n or m is zero, the z is identically zero, so n and m each effectively start at 1. Note that if we have a square with $L_x = L_y \equiv L$, then pairs of integers (n, m) yield identical frequencies if $n_1^2 + m_1^2 = n_2^2 + m_2^2$. A trivial case is where we simply switch the numbers, such as $(1, 3)$ and $(3, 1)$. But we can also have, for example, $(1, 7)$, $(7, 1)$, and $(5, 5)$.

What do these modes look like? In the case of a transverse wave on a 1-D string, it was easy to draw a snapshot on a piece of paper. But it's harder to do that in the present case, because the wave takes up three dimensions. We could take a photograph of an actual 3-D wave and then put the photograph on this page, or we could draw the wave with the aid of a computer or with fantastic artistic skills. But let's go a little more low-tech and low-talent. We'll draw the membrane in a simple binary sense, indicating only whether the z value is positive or negative. The nodes (where z is always zero) will be indicted by dotted lines. If we pick $L_x \neq L_y$ to be general, then the lowest few values of n and m yield the diagrams shown in Fig. 2. n signifies the number of (equal) regions the x direction is broken up into. And likewise for m and the y direction.

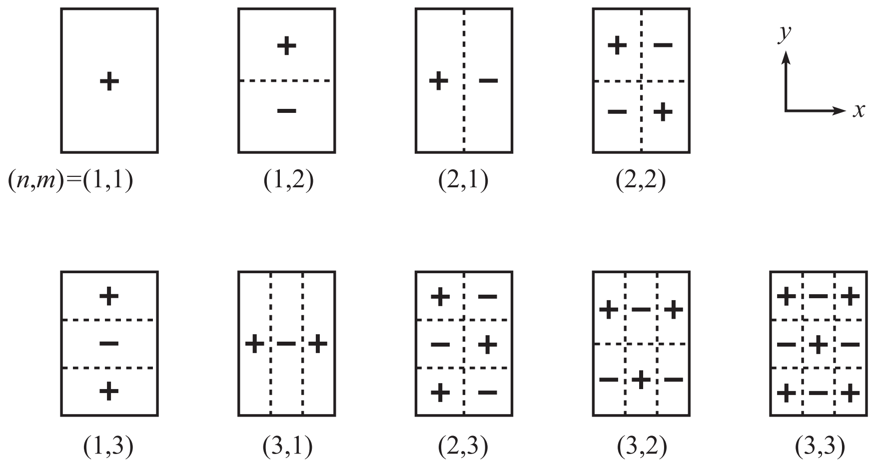

Figure 2

For each of these snapshots, a little while later when $\cos(\omega_{n,m}t+\phi_{n,m}) = 0$, the transverse displacement z will be zero everywhere, so the entire membrane will lie in the x-y plane. For the next half cycle after this time, all the $+$'s and $-$'s in each figure will be reversed. They will flip flop back and forth after each half cycle. Observe that the signs are opposite for any two regions on either side of a dotted line, consistent with the fact that z is always zero on the dotted-line nodes.

The solution for, say, the $(3, 2)$ mode is

$$z(x, y, t) = A_{3,2} \sin\left(\frac{3\pi x}{L_x}\right) \sin\left(\frac{2\pi y}{L_y}\right) \cos(\omega_{3,2}t + \phi_{3,2}), \tag{9}$$

where $\omega_{3,2}$ is given by Eq. (8). The first sine factor here is zero for $x = 0,\ L_x/3,\ 2L_x/3$, and L_x. And the second sine factor is zero for $y = 0,\ L_y/2$, and L_y. These agree with the dotted nodes in the $(3, 2)$ picture in Fig. 2. In each direction, the dotted lines are equally spaced.

Note that the various $A_{n,m}$ frequencies are *not* simple multiples of each other, as they are for a vibrating string with two fixed ends (see Section 4.5.2). For example, if $L_x = L_y \equiv L$, then the frequencies in Eq. (8) take the form,

$$\frac{\pi}{L}\sqrt{\frac{S}{\sigma}}\sqrt{n^2 + m^2}. \tag{10}$$

So the first few frequencies are $\omega_{1,1} \propto \sqrt{2}$, $\omega_{2,1} \propto \sqrt{5}$, $\omega_{2,2} \propto \sqrt{8}$, $\omega_{3,1} \propto \sqrt{10}$, and so on. Some of these are simple multiples of each other, such as $\omega_{2,2} = 2\omega_{1,1}$, but in general the ratios are irrational. So there are lots of messy harmonics. That's why musical instruments are usually one-dimensional objects. The frequencies of their modes form a nice linear progression (or are in rational multiples of you include the effects of pressing down keys or valves).

The other soap-film boundary that is reasonably easy to deal with is a circle. In this case, it is advantageous to write the partial derivative in terms of polar coordinates. It can be shown that (see Problem [to be added])

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\,. \tag{11}$$

When this is substituted into Eq. (2), the solutions $z(r, \theta, t)$ aren't as simple as in the Cartesian case, but it's still possible to get a handle on them. They involve the so-called *Bessel functions*. The pictures analogous to the ones in Fig. 2 are shown in Fig. 3. These again can be described in terms of two numbers. In the rectangular case, the nodal lines divided each direction evenly. But here the nodal lines are equally spaced in the θ direction, but *not* in the r direction.

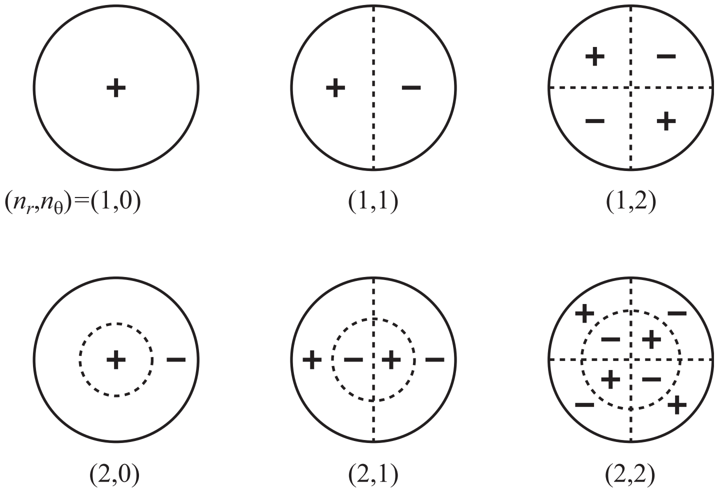

Figure 3

Chladni plates

Consider a metal plate that is caused to vibrate by, say, running a violin bow across its edge. If this is done properly (it takes some practice), then it is possible to excite a single mode with a particular frequency. Another method, which is more high-tech and failsafe (but less refined), is to blast the plate with a sound wave of a given frequency. If the frequency matches the frequency of one of the modes, then we will have a resonance effect, in the same manner that we obtained resonances in the coupled oscillator in Section 2.1.5. If the plate is sprinkled with some fine sand, the sand will settle at the nodal lines (or curves), because the plate is moving at all the other non-node points, and this motion kicks the sand off those locations. Since the sand isn't kicked off the nodes, that's where it settles. This is basically the same reason why sand collects at the side of a road and not on it. Wind from the cars pushes the sand off the road, and there's no force pushing it back on. The sand is on a one-way dead-end street, so to speak.

The nodal curves (which are different for the different modes) generally take on very interesting shapes, so we get all sorts of cool figures with the sand. Ernst Chladni (1756-1827) studied these figures in great detail. They depend on the shape of the metal plate and the mode that the plate is in. They also depend on the boundary conditions you choose. For example, you can hold the plate somewhere in the interior, or on the edge. And furthermore you can choose to hold it at any number of places. And furthermore still, there are different *ways* to hold it; you can have a clamp or a hinge. Or you can even support the plate with a string. These all give different boundary conditions. We won't get into the details, but note that if you grab the plate at a given point with a clamp or a hinge, you create a node there. (See Problem [to be added] if you do want to get into the details.) [Pictures will be added.]

7.2 Doppler effect

7.2.1 Derivation

When we talk about the frequency of a wave, we normally mean the frequency as measured in the frame in which the air (or whatever medium is relevant) is at rest. And we also normally assume that the source is at rest in the medium. But what if the source or the observer is/are moving with respect to the air? (We'll work in terms of sound waves here.) What frequency does the observer hear then? We'll find that it is modified, and this effect is known as the *Doppler effect*. In everyday experience, the Doppler effect is most widely observed with sound waves. However, it is relevant to any wave, and in particular there are important applications with electromagnetic waves (light).

Let's look at the two basic cases of a moving source and a moving observer. In both of these cases, we'll do all of the calculations in the frame of the ground, or more precisely, the frame in which the air is at rest (on average; the molecules are of course oscillating back and forth longitudinally).

Moving source

Assume that you are standing at rest on a windless day, and a car with a sound source (say, a siren) on it is heading straight toward you with speed v_s ("s" for source). The source emits sound, that is, pressure waves. Let the frequency (in the source's frame) be f cycles per second (Hertz). Let's look at two successive maxima of the pressure (actually any two points whose phases differ by 2π would suffice). In the time between the instants when the source is producing these maximum pressures, the source will travel a distance $v_s t$, where $t = 1/f$. Also during this time t, the first of the pressure maxima will travel a distance ct, where c is the wave speed. When the second pressure maximum is produced, it is therefore a distance of only $d = ct - v_s t$ behind the first maximum, instead of the ct distance if the source were at rest. The wavelength is therefore smaller. The situation is summarized in Fig. 4.

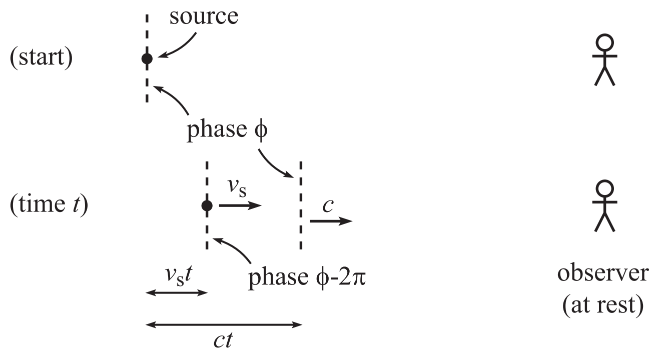

Figure 4

The movement of the source doesn't affect the wave speed, because the speed is a function of only the quantities γ, p_0, and ρ (see Eq. (5.14)); the derivation in Section 5.2 assumed nothing about the movement of the source. So the time between the arrivals at your ear of the two successive pressure maxima is $T = d/c = (c - v_s)t/c$. The frequency that you observe is therefore (the subscript "ms" is for moving source)

$$f_{ms} = \frac{1}{T} = \frac{c}{c - v_s} \cdot \frac{1}{t} = \boxed{\frac{c}{c - v_s} f} \tag{12}$$

This result is valid for $v_s < c$. We'll talk about the $v_s \geq c$ case in Section 7.2.3.

If $v_s = 0$, then Eq. (12) yields $f_{ms} = f$, of course. And if $v_s \to c$, then f_{ms} approaches infinity. This makes sense, because the pressure maxima are separated by essentially zero distance in this case (the wavelength is very small), so they pile up and a large number hit your ear in a given time interval.[2] Eq. (12) is also valid for negative v_s, and we see that $f_{ms} \to 0$ as $v_s \to -\infty$. This also makes sense, because the pressure maxima are very far apart.

Moving observer

Let's now have the source be stationary but the observer (you) be moving directly toward the source with speed v_o ("o" for observer). Consider two successive meetings between you and the pressure maxima. As shown in Fig. 5, the distance between successive maxima (in the ground frame) is simply ct, where $t = 1/f$.

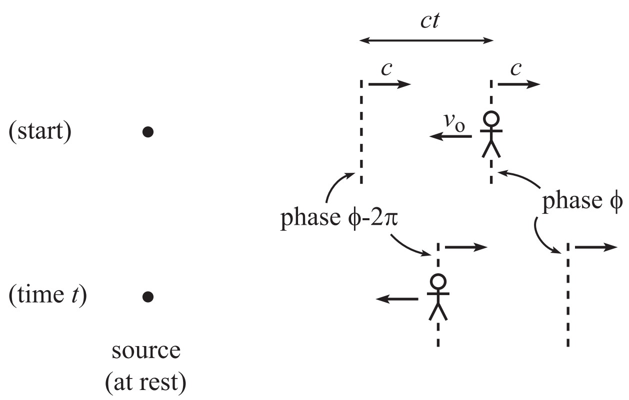

Figure 5

This gap is closed at a rate of $c + v_o$, because you are heading to the left with speed v_o and the pressure maxima are heading to the right with speed c. So the time between successive meetings is $T = ct/(c + v_o)$. The frequency that you observe is therefore (the subscript "mo" is for moving observer)

$$f_{mo} = \frac{1}{T} = \frac{c + v_o}{c} \cdot \frac{1}{t} = \boxed{\frac{c + v_o}{c} f} \tag{13}$$

This result is valid for $v_o > -c$ (where negative velocities correspond to moving to the right). If $v_o < -c$, then you are moving to the right faster than the pressure maxima, so they can never catch up to you as you recede away. In the cutoff case where $v_o = -c$, we have $f_{mo} = 0$. This makes sense, because you are receding away from the pressure maxima as fast as they are moving.

If $v_o = 0$, then Eq. (13) yields $f_{mo} = f$, of course. And if $v_o = c$, then f_{mo} equals $2f$, so it doesn't diverge as in the moving-source case. The pressure maxima are the "normal" distance of ct apart, but the gap is being closed at twice the normal rate. If $v_o \to \infty$, then Eq. (13) yields $f_{mo} \to \infty$, which makes sense. You encounter a large number of pressure maxima in a given time simply because you are moving so fast.

[2]However, if v_s is too close to c, then the wavelength becomes short enough to make it roughly the same size as the amplitude of the displacement wave. Our assumption of small slope (which we used in Section 5.2) then breaks down, and we can't trust any of these results. Nonlinear effects become important, but we won't get into those here.

REMARKS:

1. If the observed frequency is less than f, then we say that the sound (or whatever wave) is *redshifted*. If it is greater than f, then we say it is *blueshifted*. This terminology comes from the fact that red light is at the low-frequency (long wavelength) end of the visible spectrum, and blue light is at the high-frequency (short wavelength) end. This terminology is carried over to other kinds of waves, even though there is of course nothing red or blue about, say, sound waves. Well, unless someone is yelling "red" or "blue," I suppose.
2. The results in Eqs. (12) and (13) don't reduce to the same thing when $v_s = v_o$, even though these two cases yield the same relative speed between the source and observer. This is because the situation is *not* symmetrical in v_s and v_o; there is a preferred frame, namely the frame in which the air is at rest. The speed with respect to this frame matters, and not just the relative speed between the observer and the source. It makes sense that things aren't symmetrical, because a speed of $v = c$ intuitively should make f_{ms} equal to infinity, but not f_{mo}.
3. What if *both* you and the source are moving toward each other with speeds v_s and v_o? Imagine a hypothetical stationary observer located somewhere between you and the source. This observer will hear the frequency f_{ms} given in Eq. (12). For all you know, you are listening to this stationary observer emit a sound with frequency f_{ms}, instead of the original moving source with frequency f. (The stationary observer can emit a wave exactly in phase with the one he hears. Or equivalently, he can duck and have the wave go right past him.) So the frequency you hear is obtained by letting the f in Eq. (13) equal f_{ms}. The frequency when both the source and observer are moving is therefore
$$F_{mso} = \left(\frac{c+v_o}{c}\right)\left(\frac{c}{c-v_s}\right) = \boxed{\frac{c+v_o}{c-v_s}f} \tag{14}$$
4. For small v_s, we can use $1/(1-\epsilon) \approx 1+\epsilon$ to write the result in Eq. (12) as
$$f_{ms} = \frac{1}{1-v_s/c}f \approx \left(1+\frac{v_s}{c}\right)f. \tag{15}$$
And the result in Eq. (13) can be written (exactly) as
$$f_{mo} = \left(1+\frac{v_o}{c}\right)f. \tag{16}$$
So the two results take approximately the same form for small speeds.
5. If the source isn't moving in a line directly toward or away from you (or vice versa), then things are a little more complicated, but not too bad (see Problem [to be added]). The frequency changes continuously from $fc/(c-v_s)$ at $t=-\infty$ to $fc/(c+v_s)$ at $t=+\infty$ (the same formula with $v_s \to -v_s$). So it slides from one value to the other. You've undoubtedly heard a siren doing this. If the source instead hypothetically moved in a line right through you, then it would abruptly drop from the higher to the lower of these frequencies. So, in the words of John Dobson, "The reason the siren slides is because it doesn't hit you."
6. Consider a wall (or a car, or whatever) moving with speed v toward a stationary sound source. If the source emits a frequency f, and if the wall reflects the sound back toward the source, what reflected frequency is observed by someone standing next to the source? The reflection is a two-step process. First the wall acts like an observer, so from Eq. (13) it receives a frequency of $f(c+v)/c$. But then it acts like a source and emits whatever frequency it receives (imagine balls bouncing off a wall). So from Eq. (12) the observer hears a frequency of $c/(c-v)\cdot f(c+v)/c = f(c+v)/(c-v)$. The task of Problem [to be added] is to find the observed frequency if the observer (and source) is additionally moving with speed u toward the wall (which is still moving with speed v). ♣

Some examples and applications of the Doppler effect are:

SIRENS ON AMBULANCES, POLICE CARS, ETC: As the sirens move past you, the pitch goes from high to low. However, with *fire trucks*, most of the change in pitch of the siren is due to a different effect. These sirens generally start at a given pitch and then gradually get lower, independent of the truck's movement. This is because most fire trucks use siren disks to generate the sound. A siren disk is a disk with holes in it that is spun quickly in front of a fast jet of air. The result is high pressure or low pressure, depending on whether the air goes through a hole, or gets blocked by the disk. If the frequency with which the holes pass in front of the jet is in the audible range, then a sound is heard. (The wave won't be an exact sine wave, but that doesn't matter.) The spinning disk, however, usually moves due to an initial kick and not a sustained motor, so it gradually slows down. Hence the gradual decrease in pitch.

Dopper radar: Light waves reflect off a moving object, and the change in frequency is observed (see Remark 6 above). This has applications in speed guns, weather, and medicine. Along the same lines is Doppler sonar, in particular underwater with submarines.

Astronomy: Applications include the speed of stars and galaxies, the expansion of the universe, and the determination of binary star systems. The spectral lines of atomic transitions are shifted due to the motion of the star or galaxy. These applications rely on the (quite reasonable) assumption that the frequencies associated with atomic transitions are independent of their location in the universe. That is, a hydrogen atom in a distant galaxy is identical to a hydrogen atom here on earth. Hard to prove, of course, but a reasonable thing to assume.

Temperature determination of stars, plasma: This makes use of the fact that not only do spectral lines shift, they also *broaden* due to the large range of velocities of the atoms in a star (the larger the temperature, the larger the range).

7.2.2 Relativity

The difference in the results in Eqs. (12) and (13) presents an issue in the context of relativity. If a source is moving toward you with speed v and emits a certain frequency of light, then the frequency you observe must be the same as it would be if instead you were moving toward the source with speed v. This is true because one of the postulates of relativity is that there is no preferred reference frame. All that matters is the relative speed.

It is critical that we're talking about a light wave here, because light requires no medium to propagate in. (Gravity waves would work too, since they can propagate in vacuum.) If we were talking about a sound wave, then the air would define a preferred reference frame, thereby allowing the two frequencies to be different, as is the case in Eqs. (12) and (13).

So which of the above results is correct for light waves? Well, actually they're both wrong. We derived them using nonrelativistic physics, so they work fine for everyday speeds. But they are both invalid for relativistic speeds. Let's see how we can correct each of them. Let's label the $v_{\rm s}$ and $v_{\rm o}$ in the above results as v.

In the "moving source" setup, the frequency of the source in your frame is now f/γ (where $\gamma = 1/\sqrt{1 - v^2/c^2}$), because the source's clock runs slow in your frame, due to time dilation. f/γ is the frequency in your frame with which the phase of the light wave passes through a given value, say zero, *as it leaves the source*. But as in the nonrelativistic case, this isn't the frequency that you observe, due to the fact that the "wavefronts" (locations of equal phase) end up closer together. This part of the calculation proceeds just as above, so the only difference is that the emission frequency f is changed to f/γ. From Eq. (12), you therefore observe a frequency of $(f/\gamma)c/(c - v)$.

In the "moving observer" setup, the frequency we calculated in the nonrelativistic case was the frequency as measured in the source's frame. But your clock runs slow in the source's frame, due to time dilation. The frequency that you observe is therefore *larger* by a factor of γ. (It is larger because more wavefronts will hit your eye during the time that one second elapses on your clock, because your clock is running slow.) From Eq. (13), you therefore observe a frequency of $\gamma \cdot f(c+v)/c$.

As we argued above, these two results must be equal. And indeed they are, because

$$\frac{f}{\gamma}\frac{c}{c-v} = \gamma f \frac{c+v}{c} \quad\Longleftrightarrow\quad \frac{c^2}{c^2-v^2} = \gamma^2 \quad\Longleftrightarrow\quad \frac{1}{1-v^2/c^2} = \gamma^2, \tag{17}$$

which is true, as we wanted to show.

7.2.3 Shock waves

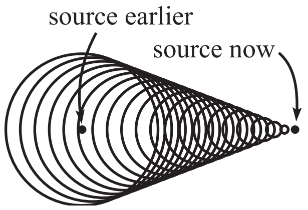

Figure 6

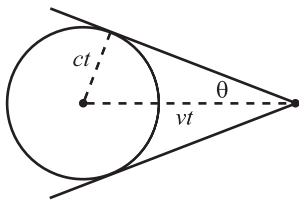

Figure 7

Let's return to the world of nonrelativistic physics. In the "moving source" setup above, we noted that the result isn't valid if $v_s > c$. So what happens in this case? Since the source is moving faster than the sound (or whatever) wave, the source gets to the observer *before* the previously-emitted wavefronts get there. If we draw a number of wavefronts (places with, say, maximal pressure) that were emitted at various times, we obtain the picture shown in Fig. 6.

The cone is a "shock wave" where the phases of waves emitted at different times are equal. This causes constructive interference. (In the case where $v_s < c$, the different wavefronts never interact with each other, so there is never any constructive interference; see Fig. 8 below.) The amplitude of the wave on the surface of the cone is therefore very large. So someone standing off to the side will hear a loud "sonic boom" when the surface of the cone passes by.

We can find the half angle of the cone in the following way. Fig. 7 shows the circular wavefront that was emitted at time t ago, along with the original and present locations of the source. The source travels a distance vt in this time, and the sound travels a distance ct. So the half angle of the cone satisfies

$$\sin\theta = \frac{c}{v}\,. \tag{18}$$

The total angle is therefore $2\theta = 2\sin^{-1}(c/v)$. This result is valid only for $v \geq c$. The larger v is, the narrower the cone. If $v \to \infty$ then $\theta \to 0$. And if $v = c$ then $2\theta = 180°$, so the "cone" is very wide, to the point of being just a straight line.

A summary of the various cases of the relative size of v and c is shown in Fig. 8. In the $v = c$ case, it is intuitively clear that the waves pile up at the location of the moving source, because the waves are never able to gain any ground on the source. In the $v > c$ case, the cone actually arises from this same effect (although to a lesser extent) for the following reason. If $v > c$, there is a particular moment in time when the distance between the source and the observer is decreasing at speed c. (This follows from continuity; the rate of decrease is v at infinity, and zero at closest approach.) The transverse component of the source's velocity isn't important for the present purposes, so at this moment the source is effectively moving directly toward the observer with speed c. The reasoning in the $v = c$ case then applies. The task of Problem [to be added] is to be quantitative about this.

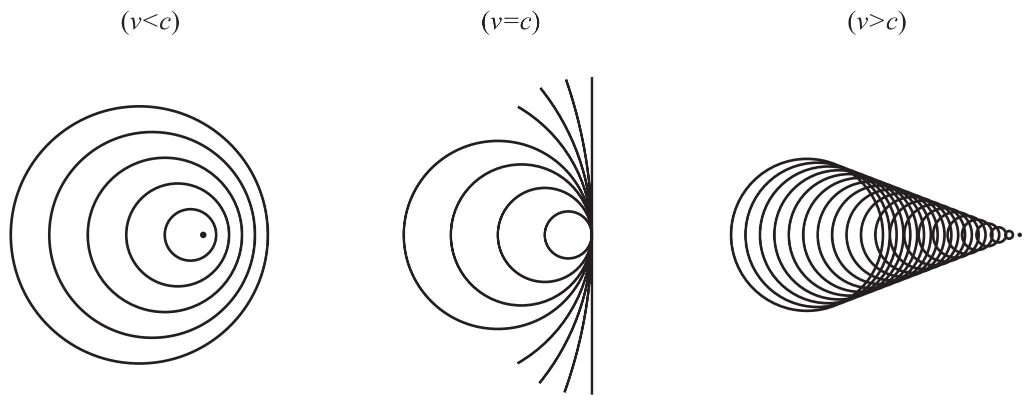

Figure 8

Shock waves exist whenever the speed of the source exceeds the speed of wave (whatever it may be) in the medium through which the source is moving. Examples include (1) planes exceeding the speed of sound, (2) boats exceeding the speed of water waves; however, this subject is more complicated due to the dispersive nature of water waves – we'll talk about this in Chapter 11, (3) charged particles moving through a material faster than the speed of light in that material (which equals c/n, where n is the index of refraction); this is called "Cherenkov radiation," and (4) the crack of a whip.

This last example is particularly interesting, because the thing that makes it possible for the tip of a whip to travel faster than the speed of sound is impedance matching; see the "Gradually changing string density" example in Section 4.3.2. Due to this impedance matching, a significant amount of the initial energy that you give to the whip ends up in the tip. And since the tip is very light, it must therefore be moving very fast. If the linear mass density of the whip changed abruptly, then not much of the initial energy would be transmitted across the boundary. The snap of a wet towel is also the same effect; see *The Physics Teacher*, pp. 376-377 (1993).