Chapter 1

Oscillations

David Morin, morin@physics.harvard.edu

A wave is a correlated collection of oscillations. For example, in a transverse wave traveling along a string, each point in the string oscillates back and forth in the transverse direction (not along the direction of the string). In sound waves, each air molecule oscillates back and forth in the longitudinal direction (the direction in which the sound is traveling). The molecules don't have any *net* motion in the direction of the sound propagation. In water waves, each water molecule also undergoes oscillatory motion, and again, there is no overall net motion.[1] So needless to say, an understanding of oscillations is required for an understanding of waves.

The outline of this chapter is as follows. In Section 1.1 we discuss simple harmonic motion, that is, motioned governed by a *Hooke's law* force, where the restoring force is proportional to the (negative of the) displacement. We discuss various ways to solve for the position $x(t)$, and we give a number of examples of such motion. In Section 1.2 we discuss damped harmonic motion, where the damping force is proportional to the velocity, which is a realistic damping force for a body moving through a fluid. We will find that there are three basic types of damped harmonic motion. In Section 1.3 we discuss damped and driven harmonic motion, where the driving force takes a sinusoidal form. (When we get to Fourier analysis, we will see why this is actually a very general type of force to consider.) We present three different methods of solving for the position $x(t)$. In the special case where the driving frequency equals the natural frequency of the spring, the amplitude becomes large. This is called *resonance*, and we will discuss various examples.

1.1 Simple harmonic motion

1.1.1 Hooke's law and small oscillations

Consider a Hooke's-law force, $F(x) = -kx$. Or equivalently, consider the potential energy, $V(x) = (1/2)kx^2$. An ideal spring satisfies this force law, although any spring will deviate significantly from this law if it is stretched enough. We study this $F(x) = -kx$ force because:

[1] The ironic thing about water waves is that although they might be the first kind of wave that comes to mind, they're much more complicated than most other kinds. In particular, the oscillations of the molecules are two dimensional instead of the normal one dimensional linear oscillations. Also, when waves "break" near a shore, everything goes haywire (the approximations that we repeatedly use throughout this book break down) and there ends up being some net forward motion. We'll talk about water waves in Chapter 12.

- We *can* study it. That it, we can solve for the motion exactly. There are many problems in physics that are extremely difficult or impossible to solve, so we might as well take advantage of a problem we can actually get a handle on.

- It is ubiquitous in nature (at least approximately). It holds in an exact sense for an idealized spring, and it holds in an approximate sense for a real-live spring, a small-angle pendulum, a torsion oscillator, certain electrical circuits, sound vibrations, molecular vibrations, and countless other setups. The reason why it applies to so many situations is the following.

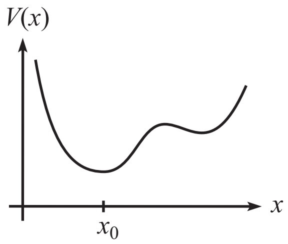


Figure 1

Let's consider an arbitrary potential, and let's see what it looks like near a local minimum. This is a reasonable place to look, because particles generally hang out near a minimum of whatever potential they're in. An example of a potential $V(x)$ is shown in Fig. 1. The best tool for seeing what a function looks like in the vicinity of a given point is the Taylor series, so let's expand $V(x)$ in a Taylor series around x_0 (the location of the minimum). We have

$$V(x) = V(x_0) + V'(x_0)(x - x_0) + \frac{1}{2!}V''(x_0)(x - x_0)^2 + \frac{1}{3!}V'''(x_0)(x - x_0)^3 + \cdots \quad (1)$$

On the righthand side, the first term is irrelevant because shifting a potential by a constant amount doesn't change the physics. (Equivalently, the force is the derivative of the potential, and the derivative of a constant is zero.) And the second term is zero due to the fact that we're looking at a minimum of the potential, so the slope $V'(x_0)$ is zero at x_0. Furthermore, the $(x - x_0)^3$ term (and all higher order terms) is negligible compared with the $(x - x_0)^2$ term if x is sufficiently close to x_0, which we will assume is the case.[2] So we are left with

$$\boxed{V(x) \approx \frac{1}{2}V''(x_0)(x - x_0)^2} \quad (2)$$

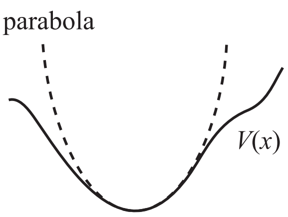


Figure 2

In other words, we have a potential of the form $(1/2)kx^2$, where $k \equiv V''(x_0)$, and where we have shifted the origin of x so that it is located at x_0. Equivalently, we are just measuring x relative to x_0.

We see that *any* potential looks basically like a Hooke's-law spring, as long as we're close enough to a local minimum. In other words, the curve can be approximated by a parabola, as shown in Fig. 2. This is why the harmonic oscillator is so important in physics.

We will find below in Eqs. (7) and (11) that the (angular) frequency of the motion in a Hooke's-law potential is $\omega = \sqrt{k/m}$. So for a general potential $V(x)$, the $k \equiv V''(x_0)$ equivalence implies that the frequency is

$$\omega = \sqrt{\frac{V''(x_0)}{m}}\,. \quad (3)$$

1.1.2 Solving for $x(t)$

The long way

The usual goal in a physics setup is to solve for $x(t)$. There are (at least) two ways to do this for the force $F(x) = -kx$. The straightforward but messy way is to solve the $F = ma$ differential equation. One way to write $F = ma$ for a harmonic oscillator is $-kx = m \cdot dv/dt$. However, this isn't so useful, because it contains three variables, x, v, and t. We therefore

[2] The one exception occurs when $V''(x)$ equals zero. However, there is essentially zero probability that $V''(x_0) = 0$ for any actual potential. And even if it does, the result in Eq. (3) below is still technically true; they frequency is simply zero.

can't use the standard strategy of separating variables on the two sides of the equation and then integrating. Equation have only two sides, after all. So let's instead write the acceleration as $a = v \cdot dv/dx$.[3] This gives

$$F = ma \implies -kx = m\left(v\frac{dv}{dx}\right) \implies -\int kx\,dx = \int mv\,dv. \tag{4}$$

Integration then gives (with E being the integration constant, which happens to be the energy)

$$E - \frac{1}{2}kx^2 = \frac{1}{2}mv^2 \implies v = \pm\sqrt{\frac{2}{m}}\sqrt{E - \frac{1}{2}kx^2}. \tag{5}$$

Writing v as dx/dt here and separating variables one more time gives

$$\frac{dx}{\sqrt{E}\sqrt{1 - \frac{kx^2}{2E}}} = \pm\sqrt{\frac{2}{m}}\int dt. \tag{6}$$

A trig substitution turns the lefthand side into an arccos (or arcsin) function. The result is (see Problem [to be added] for the details)

$$\boxed{x(t) = A\cos(\omega t + \phi)} \qquad \text{where} \quad \boxed{\omega = \sqrt{\frac{k}{m}}} \tag{7}$$

and where A and ϕ are arbitrary constants that are determined by the two initial conditions (position and velocity); see the subsection below on initial conditions. A happens to be $\sqrt{2E/k}$, where E is the above constant of integration. The solution in Eq. (7) describes *simple harmonic motion*, where $x(t)$ is a simple sinusoidal function of time. When we discuss damping in Section 1.2, we will find that the motion is somewhat sinusoidal, but with an important modification.

The short way

$F = ma$ gives

$$-kx = m\frac{d^2x}{dt^2}\,. \tag{8}$$

This equation tells us that we want to find a function whose second derivative is proportional to the negative of itself. But we already know some functions with this property, namely sines, cosines, and exponentials. So let's be fairly general and try a solution of the form,

$$x(t) = A\cos(\omega t + \phi). \tag{9}$$

A sine or an exponential function would work just as well. But a sine function is simply a shifted cosine function, so it doesn't really generate anything new; it just changes the phase. We'll talk about exponential solutions in the subsection below. Note that a phase ϕ (which shifts the curve on the t axis), a scale factor of ω in front of the t (which expands or contracts the curve on the t axis), and an overall constant A (which expands or contracts the curve on the x axis) are the only ways to modify a cosine function if we want it to stay a cosine. (Well, we could also add on a constant and shift the curve in the x direction, but we want the motion to be centered around $x = 0$.)

[3] This does indeed equal a, because $v \cdot dv/dx = dx/dt \cdot dv/dx = dv/dt = a$. And yes, it's legal to cancel the dx's here (just imagine them to be small but not infinitesimal quantities, and then take a limit).

If we plug Eq. (9) into Eq. (8), we obtain

$$\begin{aligned} -k\Big(A\cos(\omega t+\phi)\Big) &= m\Big(-\omega^2 A\cos(\omega t+\phi)\Big) \\ \Longrightarrow \quad (-k+m\omega^2)\Big(A\cos(\omega t+\phi)\Big) &= 0. \end{aligned} \tag{10}$$

Since this must be true for *all* t, we must have

$$k - m\omega^2 = 0 \quad \Longrightarrow \quad \omega = \sqrt{\frac{k}{m}}\,, \tag{11}$$

in agreement with Eq. (7). The constants ϕ and A don't appear in Eq. (11), so they can be anything and the solution in Eq. (9) will still work, provided that $\omega = \sqrt{k/m}$. They are determined by the initial conditions (position and velocity).

We have found one solution in Eq. (9), but how do we know that we haven't missed any other solutions to the $F = ma$ equation? From the trig sum formula, we can write our one solution as

$$A\cos(\omega t+\phi) = A\cos\phi\cos(\omega t) - A\sin\phi\sin(\omega t), \tag{12}$$

So we have actually found *two* solutions: a sin and a cosine, with arbitrary coefficients in front of each (because ϕ can be anything). The solution in Eq. (9) is simply the sum of these two individual solutions. The fact that the sum of two solutions is again a solution is a consequence of the *linearity* our $F = ma$ equation. By linear, we mean that x appears only through its first power; the number of derivatives doesn't matter.

We will now invoke the fact that an nth-order linear differential equation has n independent solutions (see Section 1.1.4 below for some justification of this). Our $F = ma$ equation in Eq. (8) involves the second derivative of x, so it is a second-order equation. So we'll accept the fact that it has two independent solutions. Therefore, since we've found two, we know that we've found them all.

The parameters

A few words on the various quantities that appear in the $x(t)$ in Eq. (9).

- ω is the *angular frequency.*[4] Note that

$$\begin{aligned} x\left(t+\frac{2\pi}{\omega}\right) = A\cos\big(\omega(t+2\pi/\omega)+\phi\big) &= A\cos(\omega t+\phi+2\pi) \\ &= A\cos(\omega t+\phi) \\ &= x(t). \end{aligned} \tag{13}$$

  Also, using $v(t) = dx/dt = -\omega A\sin(\omega t+\phi)$, we find that $v(t+2\pi/\omega) = v(t)$. So after a time of $T \equiv 2\pi/\omega$, both the position and velocity are back to where they were (and the force too, since it's proportional to x). This time T is therefore the *period.* The motion repeats after every time interval of T. Using $\omega = \sqrt{k/m}$, we can write $T = 2\pi\sqrt{m/k}$.

[4]It is sometimes also called the angular speed or angular velocity. Although there are technically differences between these terms, we'll generally be sloppy and use them interchangeably. Also, it gets to be a pain to keep saying the word "angular," so we'll usually call ω simply the "frequency." This causes some ambiguity with the frequency, ν, as measured in Hertz (cycles per second); see Eq. (14). But since ω is a much more natural quantity to use than ν, we will invariably work with ω. So "frequency" is understood to mean ω in this book.

The frequency in Hertz (cycles per second) is given by $\nu = 1/T$. For example, if $T = 0.1\,\text{s}$, then $\nu = 1/T = 10\,\text{s}^{-1}$, which means that the system undergoes 10 oscillations per second. So we have

$$\nu = \frac{1}{T} = \frac{\omega}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{k}{m}}\,. \tag{14}$$

To remember where the "2π" in $\nu = \omega/2\pi$ goes, note that ω is larger than ν by a factor of 2π, because one revolution has 2π radians in it, and ν is concerned with revolutions whereas ω is concerned with radians.

Note the extremely important point that the frequency is independent of the amplitude. You might think that the frequency should be smaller if the amplitude is larger, because the mass has farther to travel. But on the other hand, you might think that the frequency should be larger if the amplitude is larger, because the force on the mass is larger which means that it is moving faster at certain points. It isn't intuitively obvious which of these effects wins, although it does follow from dimensional analysis (see Problem [to be added]). It turns out that the effects happen to exactly cancel, making the frequency independent of the amplitude. Of course, in any real-life situation, the $F(x) = -kx$ form of the force will break down if the amplitude is large enough. But in the regime where $F(x) = -kx$ is a valid approximation, the frequency is independent of the amplitude.

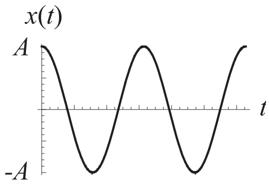


Figure 3

- A is the *amplitude*. The position ranges from A to $-A$, as shown in Fig. 3
- ϕ is the *phase*. It gives a measure of what the position is a $t = 0$. ϕ is dependent on when you pick the $t = 0$ time to be. Two people who start their clocks at different times will have different phases in their expressions for $x(t)$. (But they will have the same ω and A.) Two phases that differ by 2π are effectively the same phase.

  Be careful with the sign of the phase. Fig. 4 shows plots of $A\cos(\omega t + \phi)$, for $\phi = 0$, $\pm\pi/2$, and π. Note that the plot for $\phi = +\pi/2$ is shifted to the *left* of the plot for $\phi = 0$, whereas the plot for $\phi = -\pi/2$ is shifted to the *right* of the plot for $\phi = 0$. These are due to the fact that, for example, the $\phi = -\pi/2$ case requires a larger time to achieve the same position as the $\phi = 0$ case. So a given value of x occurs *later* in the $\phi = -\pi/2$ plot, which means that it is shifted to the *right*.

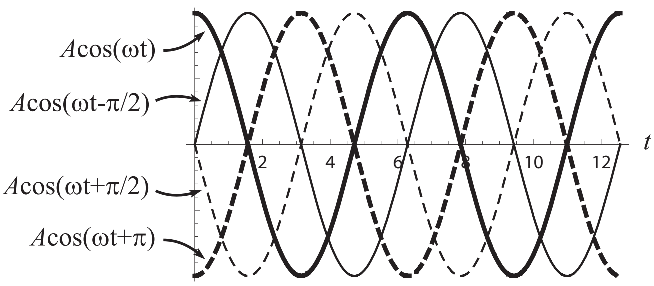


Figure 4

Various ways to write $x(t)$

We found above that $x(t)$ can be expressed as $x(t) = A\cos(\omega t + \phi)$. However, this isn't the only way to write $x(t)$. The following is a list of equivalent expressions.

$$x(t) = A\cos(\omega t + \phi)$$

$$
\begin{aligned}
&= A\sin(\omega t + \phi') \\
&= B_{\rm c}\cos\omega t + B_{\rm s}\sin\omega t \\
&= Ce^{i\omega t} + C^* e^{-i\omega t} \\
&= \mathrm{Re}\left(De^{i\omega t}\right). \qquad (15)
\end{aligned}
$$

A, $B_{\rm c}$, and $B_{\rm s}$ are real quantities here, but C and D are (possibly) complex. C^* denotes the complex conjugate of C. See Section 1.1.5 below for a discussion of matters involving complex quantities. Each of the above expressions for $x(t)$ involves two parameters – for example, A and ϕ, or the real and imaginary parts of C. This is consistent with the fact that there are two initial conditions (position and velocity) that must be satisfied.

The two parameters in a given expression are related to the two parameters in each of the other expressions. For example, $\phi' = \phi + \pi/2$, and the various relations among the other parameters can be summed up by

$$
\begin{aligned}
B_{\rm c} &= A\cos\phi = 2\mathrm{Re}(C) = \mathrm{Re}(D), \\
B_{\rm s} &= -A\sin\phi = -2\mathrm{Im}(C) = -\mathrm{Im}(D), \qquad (16)
\end{aligned}
$$

and a quick corollary is that $D = 2C$. The task of Problem [to be added] is to verify these relations. Depending on the situation at hand, one of the expressions in Eq. (15) might work better than the others, as we'll see in Section 1.1.7 below.

1.1.3 Linearity

As we mentioned right after Eq. (12), linear differential equations have the property that the sum (or any linear combination) of two solutions is again a solution. For example, if $\cos\omega t$ and $\sin\omega t$ are solutions, then $A\cos\omega t + B\sin\omega t$ is also a solution, for any constants A and B. This is consistent with the fact that the $x(t)$ in Eq. (12) is a solution to our Hooke's-law $m\ddot{x} = -kx$ equation.

This property of linear differential equations is easy to verify. Consider, for example, the second order (although the property holds for any order) linear differential equation,

$$A\ddot{x} + B\dot{x} + Cx = 0. \qquad (17)$$

Let's say that we've found two solutions, $x_1(t)$ and $x_2(t)$. Then we have

$$
\begin{aligned}
A\ddot{x}_1 + B\dot{x}_1 + Cx_1 &= 0, \\
A\ddot{x}_2 + B\dot{x}_2 + Cx_2 &= 0. \qquad (18)
\end{aligned}
$$

If we add these two equations, and switch from the dot notation to the d/dt notation, then we have (using the fact that the sum of the derivatives is the derivative of the sum)

$$A\frac{d^2(x_1+x_2)}{dt^2} + B\frac{d(x_1+x_2)}{dt} + C(x_1+x_2) = 0. \qquad (19)$$

But this is just the statement that $x_1 + x_2$ is a solution to our differential equation, as we wanted to show.

What if we have an equation that isn't linear? For example, we might have

$$A\ddot{x} + B\dot{x}^2 + Cx = 0. \qquad (20)$$

If x_1 and x_2 are solutions to this equation, then if we add the differential equations applied to each of them, we obtain

$$A\frac{d^2(x_1+x_2)}{dt^2} + B\left[\left(\frac{dx_1}{dt}\right)^2 + \left(\frac{dx_1}{dt}\right)^2\right] + C(x_1+x_2) = 0. \qquad (21)$$

This is *not* the statement that $x_1 + x_2$ is a solution, which is instead

$$A\frac{d^2(x_1+x_2)}{dt^2} + B\left(\frac{d(x_1+x_2)}{dt}\right)^2 + C(x_1+x_2) = 0. \tag{22}$$

The two preceding equations differ by the cross term in the square in the latter, namely $2B(dx_1/dt)(dx_2/dt)$. This is in general nonzero, so we conclude that x_1+x_2 is not a solution. No matter what the order if the differential equation is, we see that these cross terms will arise if and only if the equation isn't linear.

This property of linear differential equations – that the sum of two solutions is again a solution – is extremely useful. It means that we can build up solutions from other solutions. Systems that are governed by linear equations are *much* easier to deal with than systems that are governed by nonlinear equations. In the latter, the various solutions aren't related in an obvious way. Each one sits in isolation, in a sense. General Relativity is an example of a theory that is governed by nonlinear equations, and solutions are indeed very hard to come by.

1.1.4 Solving nth-order linear differential equations

The "fundamental theorem of algebra" states that any nth-order polynomial,

$$a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0, \tag{23}$$

can be factored into

$$a_n(z-r_1)(z-r_2)\cdots(z-r_n). \tag{24}$$

This is believable, but by no means obvious. The proof is a bit involved, so we'll just accept it here.

Now consider the nth-order linear differential equation,

$$a_n\frac{d^n x}{dt^n} + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \cdots + a_1\frac{dx}{dt} + a_0 = 0. \tag{25}$$

Because differentiation by t commutes with multiplication by a constant, we can invoke the equality of the expressions in Eqs. (23) and (24) to say that Eq. (25) can be rewritten as

$$a_n\left(\frac{d}{dt} - r_1\right)\left(\frac{d}{dt} - r_2\right)\cdots\left(\frac{d}{dt} - r_n\right)x = 0. \tag{26}$$

In short, we can treat the d/dt derivatives here like the z's in Eq. (24), so the relation between Eqs. (26) and (25) is the same as the relation between Eqs. (24) and (23). And because all the factors in Eq. (26) commute with each other, we can imagine making any of the factors be the rightmost one. Therefore, any solution to the equation,

$$\left(\frac{d}{dt} - r_i\right)x = 0 \quad\Longleftrightarrow\quad \frac{dx}{dt} = r_i x, \tag{27}$$

is a solution to the original equation, Eq. (25). The solutions to these n first-order equations are simply the exponential functions, $x(t) = Ae^{r_i t}$. We have therefore found n solutions, so we're done. (We'll accept the fact that there are only n solutions.) So this is why our strategy for solving differential equations is to always guess exponential solutions (or trig solutions, as we'll see in the following section).

1.1.5 Taking the real part

In the second (short) derivation of $x(t)$ we presented above, we guessed a solution of the form, $x(t) = A\cos(\omega t+\phi)$. However, anything that can be written in terms of trig functions can also be written in terms of exponentials. This fact follows from one of the nicest formulas in mathematics:

$$\boxed{e^{i\theta} = \cos\theta + i\sin\theta} \tag{28}$$

This can be proved in various ways, the quickest of which is to write down the Taylor series for both sides and then note that they are equal. If we replace θ with $-\theta$ in this relation, we obtain $e^{-i\theta} = \cos\theta - i\sin\theta$, because $\cos\theta$ and $\sin\theta$ are even and odd functions of θ, respectively. Adding and subtracting this equation from Eq. (28) allows us to solve for the trig functions in terms of the exponentials:

$$\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2}, \qquad \text{and} \qquad \sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}. \tag{29}$$

So as we claimed above, anything that can be written in terms of trig functions can also be written in terns of exponentials (and vice versa). We can therefore alternatively guess an exponential solution to our $-kx = m\ddot{x}$ differential equation. Plugging in $x(t) = Ce^{\alpha t}$ gives

$$-kCe^{\alpha t} = m\alpha^2 Ce^{\alpha t} \quad \Longrightarrow \quad \alpha^2 = -\frac{k}{m} \quad \Longrightarrow \quad \alpha = \pm i\omega, \quad \text{where } \omega = \sqrt{\frac{k}{m}}. \tag{30}$$

We have therefore found *two* solutions, $x_1(t) = C_1 e^{i\omega t}$, and $x_2(t) = C_2 e^{-i\omega t}$. The C_1 coefficient here need not have anything to do with the C_2 coefficient. Due to linearity, the most general solution is the sum of these two solutions,

$$\boxed{x(t) = C_1 e^{i\omega t} + C_2 e^{-i\omega t}} \tag{31}$$

This expression for $x(t)$ satisfies the $-kx = m\ddot{x}$ equation for *any* (possibly complex) values of C_1 and C_2. However, $x(t)$ must of course be real, because an object can't be at a position of, say, $3+7i$ meters (at least in this world). This implies that the two terms in Eq. (31) must be complex conjugates of each other, which in turn implies that C_2 must be the complex conjugate of C_1. This is the reasoning that leads to the fourth expression in Eq. (15).

There are two ways to write any complex number: either as the sum of a real and imaginary part, or as the product of a magnitude and a phase $e^{i\phi}$. The equivalence of these is a consequence of Eq. (28). Basically, if we plot the complex number in the complex plane, we can write it in either Cartesian or polar coordinates. If we choose the magnitude-phase way and write C_1 as $C_0 e^{i\phi}$, then the complex conjugate is $C_2 = C_0 e^{-i\phi}$. Eq. (31) then becomes

$$\begin{aligned} x(t) &= C_0 e^{i\phi} e^{i\omega t} + C_0 e^{-i\phi} e^{-i\omega t} \\ &= 2C_0\cos(\omega t+\phi), \end{aligned} \tag{32}$$

where we have used Eq. (29). We therefore end up with the trig solution that we had originally obtained by guessing, so everything is consistent.

Note that by adding the two complex conjugate solutions together in Eq. (32), we basically just took the real part of the $C_0 e^{i\phi} e^{i\omega t}$ solution (and then multiplied by 2, but that can be absorbed in a redefinition of the coefficient). So we will often simply work with the exponential solution, with the understanding that we must *take the real part* in the end to get the actual physical solution.

If this strategy of working with an exponential solution and then taking the real part seems suspect or mysterious to you, then for the first few problems you encounter, you should do things the formal way. That is, you should add on the second solution and then demand that $x(t)$ (or whatever the relevant variable is in a given setup) is real. This will result in the sum of two complex conjugates. After doing this a few of times, you will realize that you always end up with (twice) the real part of the exponential solutions (either of them). Once you're comfortable with this fact, you can take a shortcut and forget about adding on the second solution and the other intermediate steps. But that's what you're really doing.

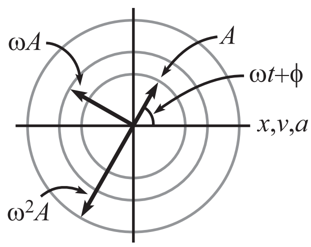
REMARK: The original general solution for $x(t)$ in Eq. (31) contains *four* parameters, namely the real and imaginary parts of C_1, and likewise for C_2 (ω is determined by k and m). Or equivalently, the four parameters are the magnitude and phase of C_1, and likewise for C_2. These four parameters are all independent, because we haven't yet invoked the fact that $x(t)$ must be real. If we do invoke this, it cuts down the number of parameters from four to two. These two parameters are then determined by the two initial conditions (position and velocity).

However, although there's no arguing with the "$x(t)$ must be real" reasoning, it does come a little out of the blue. It would be nice to work entirely in terms of initial conditions. But how can we solve for four parameters with only two initial conditions? Well, we can't. But the point is that there are actually *four* initial conditions, namely the real and imaginary parts of the initial position, and the real and imaginary parts of the initial velocity. That is, $x(0) = x_0 + 0 \cdot i$, and $v(0) = v_0 + 0 \cdot i$. It takes four quantities (x_0, 0, v_0, and 0 here) to specify these two (possibly complex) quantities. (Once we start introducing complex numbers into $x(t)$, we of course have to allow the initial position and velocity to be complex.) These four given quantities allow us to solve for the four parameters in $x(t)$. And in the end, this process gives (see Problem [to be added]) the same result as simply demanding that $x(t)$ is real. ♣

1.1.6 Phase relations and phasor diagrams

Let's now derive the phase relation between $x(t)$, $v(t)$, and $a(t)$. We have

$$\begin{aligned} x(t) &= A\cos(\omega t + \phi), \\ \Longrightarrow \quad v(t) &= \frac{dx}{dt} = -\omega A \sin(\omega t + \phi) = \omega A \cos\left(\omega t + \phi + \frac{\pi}{2}\right), \\ \Longrightarrow \quad a(t) &= \frac{dv}{dt} = -\omega^2 A\cos(\omega t + \phi) = \omega^2 A \cos\left(\omega t + \phi + \pi\right). \end{aligned} \tag{33}$$



horizontal projections give x,v, and a

Figure 5

We see that a leads v by $\pi/2$, and v leads x by $\pi/2$. These phase relations can be conveniently expressed in the *phasor* diagram in Fig. 5. The values of x, v, and a are represented by vectors, where it is understood that to get their actual values, we must take the projection of the vectors onto the horizontal axis. The whole set of three vectors swings around counterclockwise with angular speed ω as time goes on. The initial angle between the x phasor and the horizontal axis is picked to be ϕ. So the angle of the x phasor as a function of time is $\omega t + \phi$, the angle of the v phasor is $\omega t + \phi + \pi/2$, and the angle of the a phasor is $\omega t + \phi + \pi$. Since taking the horizontal projection simply brings in a factor of the cosine of the angle, we do indeed obtain the expressions in Eq. (33), as desired.

The units of x, v, and a are different, so technically we shouldn't be drawing all three phasors in the same diagram, but it helps in visualizing what's going on. Since the phasors swing around counterclockwise, the diagram makes it clear that $a(t)$ is $\pi/2$ *ahead* of $v(t)$, which is $\pi/2$ *ahead* of $x(t)$. So the actual cosine forms of x, v, and a look like the plots shown in Fig. 6 (we've chosen $\phi = 0$ here).

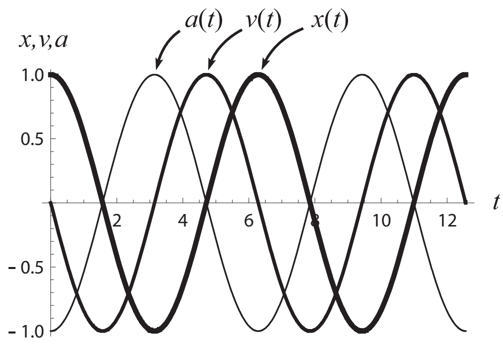


Figure 6

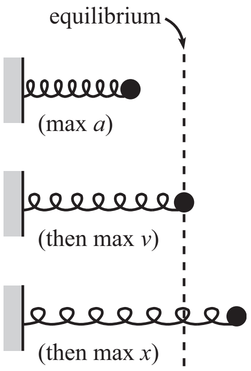


Figure 7

$a(t)$ reaches its maximum *before* $v(t)$ (that is, $a(t)$ is *ahead* of $v(t)$). And $v(t)$ reaches its maximum *before* $x(t)$ (that is, $v(t)$ is *ahead* of $x(t)$). So the plot of $a(t)$ is shifted to the *left* from $v(t)$, which is shifted to the *left* from $x(t)$. If we look at what an actual spring-mass system is doing, we have the three successive pictures shown in Fig. 7. Figures 5, 6, and 7 are three different ways of saying the same thing about the relative phases.

1.1.7 Initial conditions

As we mentioned above, each of the expressions for $x(t)$ in Eq. (15) contains two parameters, and these two parameters are determined from the initial conditions. These initial conditions are invariably stated as the initial position and initial velocity. In solving for the subsequent motion, any of the forms in Eq. (15) will work, but the

$$x(t) = B_{\text{c}} \cos \omega t + B_{\text{s}} \sin \omega t \tag{34}$$

form is the easiest one to work with when given $x(0)$ and $v(0)$. Using

$$v(t) = \frac{dx}{dt} = -\omega B_{\text{c}} \sin \omega t + \omega B_{\text{s}} \cos \omega t, \tag{35}$$

the conditions $x(0) = x_0$ and $v_0 = v_0$ yield

$$x_0 = x(0) = B_{\text{c}}, \qquad \text{and} \qquad v_0 = v(0) = \omega B_{\text{s}} \implies B_{\text{s}} = \frac{v_0}{\omega}\,. \tag{36}$$

Therefore,

$$\boxed{x(t) = x_0 \cos \omega t + \frac{v_0}{\omega} \sin \omega t} \tag{37}$$

If you wanted to use the $x(t) = A\cos(\omega t + \phi)$ form instead, then $v(t) = -\omega A \sin(\omega t + \phi)$. The initial conditions now give $x_0 = x(0) = A\cos\phi$ and $v_0 = -\omega A \sin\phi$. Dividing gives $\tan\phi = -v_0/\omega x_0$. Squaring and adding (after dividing by ω) gives $A = \sqrt{x_0^2 + (v_0/\omega)^2}$. We have chosen the positive root for A; the negative root would simply add π on to ϕ. So we have

$$x(t) = \sqrt{x_0^2 + \left(\frac{v_0}{\omega}\right)^2} \cos\left(\omega t + \arctan\left(\frac{-v_0}{\omega x_0}\right)\right). \tag{38}$$

The correct choice from the two possibilities for the arctan angle is determined by either $\cos\phi = x_0/A$ or $\sin\phi = -v_0/\omega A$. The result in Eq. (38) is much messier than the result in Eq. (37), so it is clearly advantageous to use the form of $x(t)$ given in Eq. (34).

All of the expressions for $x(t)$ in Eq. (15) contain two parameters. If someone proposed a solution with only one parameter, then there is no way that it could be a general solution, because we don't have the freedom to satisfy the *two* initial conditions. Likewise, if someone proposed a solution with three parameters, then there is no way we could ever determine all three parameters, because we have only two initial conditions. So it is good that the expressions in Eq. (15) contain two parameters. And this is no accident. It follows from the fact that our $F = ma$ equation is a *second*-order differential equation; the general solution to such an equation always contains two free parameters.

We therefore see that the fact that *two* initial conditions completely specify the motion of the system is intricately related to the fact that the $F = ma$ equation is a *second*-order differential equation. If instead of $F = m\ddot{x}$, Newton's second law was the first-order equation, $F = m\dot{x}$, then we wouldn't have the freedom of throwing a ball with an initial velocity of our choosing; everything would be determined by the initial position only. This is clearly not how things work. On the other hand, if Newton's second law was the third-order equation, $F = md^3x/dt^3$, then the motion of a ball wouldn't be determined by an initial position and velocity (along with the forces in the setup at hand); we would also have to state the initial acceleration. And again, this is not how things work (in this world, at least).

1.1.8 Energy

$F(x) = -kx$ is a conservative force. That is, the work done by the spring is path-independent. Or equivalently, the work done depends only on the initial position x_i and the final position x_f. You can quickly show that that work is $\int(-kx)\,dx = kx_\text{i}^2/2 - kx_\text{f}^2/2$. Therefore, since the force is conservative, the energy is conserved. Let's check that this is indeed the case. We have

$$\begin{aligned} E &= \frac{1}{2}kx^2 + \frac{1}{2}mv^2 \\ &= \frac{1}{2}k\big(A\cos(\omega t+\phi)\big)^2 + \frac{1}{2}m\big(-\omega A\sin(\omega t+\phi)\big)^2 \\ &= \frac{1}{2}A^2\Big(k\cos^2(\omega t+\phi) + m\omega^2\sin^2(\omega t+\phi)\Big) \\ &= \frac{1}{2}kA^2\Big(\cos^2(\omega t+\phi) + \sin^2(\omega t+\phi)\Big) \qquad (\text{using } \omega^2 \equiv k/m) \\ &= \frac{1}{2}kA^2. \end{aligned} \tag{39}$$

This makes sense, because $kA^2/2$ is the potential energy of the spring when it is stretched the maximum amount (and so the mass is instantaneously at rest). Fig. 8 shows how the energy breaks up into kinetic and potential, as a function of time. We have arbitrarily chosen ϕ to be zero. The energy sloshes back and forth between kinetic and potential. It is all potential at the points of maximum stretching, and it is all kinetic when the mass passes through the origin.

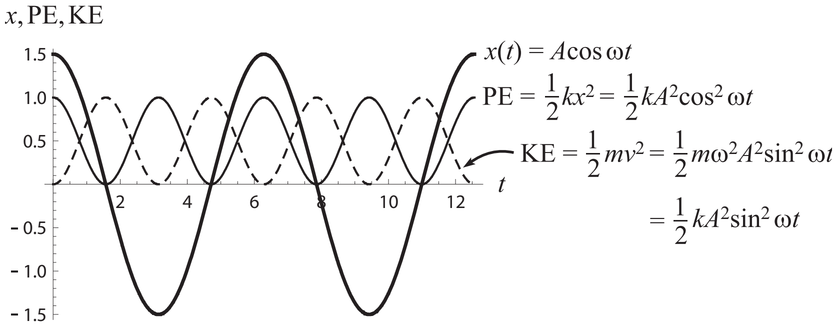


Figure 8

1.1.9 Examples

Let's now look at some examples of simple harmonic motion. Countless examples exist in the real word, due to the Taylor-series argument in Section 1.1. Technically, all of these examples are just approximations, because a force never takes *exactly* the $F(x) = -kx$ form; there are always slight modifications. But these modifications are negligible if the amplitude is small enough. So it is understood that we will always work in the approximation of small amplitudes. Of course, the word "small" here is meaningless by itself. The correct statement is that the amplitude must be small compared with some other quantity that is inherent to the system and that has the same units as the amplitude. We generally won't worry about exactly what this quantity is; we'll just assume that we've picked an amplitude that is sufficiently small.

The moral of the examples below is that whenever you arrive at an equation of the form $\ddot{z} + (\text{something})z = 0$, you know that z undergoes simple harmonic motion with $\omega = \sqrt{\text{something}}$.

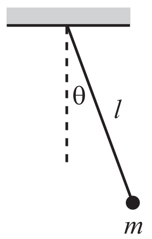


Figure 9

Simple pendulum

Consider the simple pendulum shown in Fig. 9. (The word "simple" refers to the fact that the mass is a point mass, as opposed to an extended mass in the "physical " pendulum below.) The mass hangs on a massless string and swings in a vertical plane. Let ℓ be the length of the string, and let $\theta(t)$ be the angle the string makes with the vertical. The gravitational force on the mass in the tangential direction is $-mg\sin\theta$. So $F = ma$ in the tangential direction gives

$$-mg\sin\theta = m(\ell\ddot{\theta}) \tag{40}$$

The tension in the string combines with the radial component of gravity to produce the radial acceleration, so the radial $F = ma$ equation serves only to tell us the tension, which we won't need here.

Eq. (40) isn't solvable in closed form. But for small oscillations, we can use the $\sin\theta \approx \theta$ approximation to obtain

$$\ddot{\theta} + \omega^2\theta = 0, \qquad \text{where } \omega \equiv \sqrt{\frac{g}{\ell}}\,. \tag{41}$$

This looks exactly like the $\ddot{x}+\omega^2 x$ equation for the Hooke's-law spring, so all of our previous results carry over. The only difference is that ω is now $\sqrt{g/\ell}$ instead of $\sqrt{k/m}$. Therefore, we have

$$\theta(t) = A\cos(\omega t + \phi), \tag{42}$$

where A and ϕ are determined by the initial conditions. So the pendulum undergoes simple harmonic motion with a frequency of $\sqrt{g/\ell}$. The period is therefore $T = 2\pi/\omega = 2\pi\sqrt{\ell/g}$. The true motion is arbitrarily close to this, for sufficiently small amplitudes. Problem [to be added] deals with the higher-order corrections to the motion in the case where the amplitude is not small.

Note that the angle θ bounces back and forth between $\pm A$, where A is small. But the phase angle $\omega t + \phi$ increases without bound, or equivalently keeps running from 0 to 2π repeatedly.

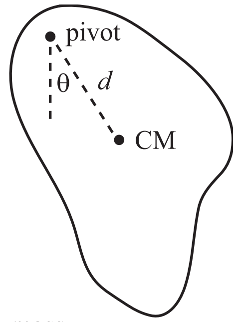


Figure 10

Physical pendulum

Consider the "physical" pendulum shown in Fig. 10. The planar object swings back and forth in the vertical plane that it lies in. The pivot is inside the object in this case, but it need not be. You could imagine a stick (massive or massless) that is glued to the object and attached to the pivot at its upper end.

We can solve for the motion by looking at the torque on the object around the pivot. If the moment of inertia of the object around the pivot is I, and if the object's CM is a distance d from the pivot, then $\tau = I\alpha$ gives (using the approximation $\sin\theta \approx \theta$)

$$-mgd\sin\theta = I\ddot{\theta} \quad \Longrightarrow \quad \ddot{\theta} + \omega^2\theta = 0, \qquad \text{where } \omega \equiv \sqrt{\frac{mgd}{I}}. \tag{43}$$

So we again have simple harmonic motion. Note that if the object is actually a point mass, then we have $I = md^2$, and the frequency becomes $\omega = \sqrt{g/d}$. This agrees with the simple-pendulum result in Eq. (41) with $\ell \to d$.

If you wanted, you could also solve this problem by using $\tau = I\alpha$ around the CM, but then you would need to also use the $F_x = ma_x$ and $F_y = ma_y$ equations. All of these equations involve messy forces at the pivot. The point is that by using $\tau = I\alpha$ around the pivot, you can sweep these forces under the rug.

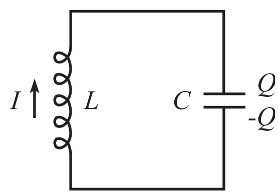


Figure 11

LC circuit

Consider the LC circuit shown in Fig. 11. Kirchhoff's rule (which says that the net voltage drop around a closed loop must be zero) applied counterclockwise yields

$$-L\frac{dI}{dt} - \frac{Q}{C} = 0. \tag{44}$$

But $I = dQ/dt$, so we have

$$-L\ddot{Q} - \frac{Q}{C} = 0 \quad \Longrightarrow \quad \ddot{Q} + \omega^2 Q = 0, \qquad \text{where } \omega \equiv \sqrt{\frac{1}{LC}}. \tag{45}$$

So we again have simple harmonic motion. In comparing this $L\ddot{Q} + (1/C)Q$ equation with the simple-harmonic $m\ddot{x} + kx = 0$ equation, we see that L is the analog of m, and $1/C$ is the analog of k. L gives a measure of the inertia of the system; the larger L is, the more the inductor resists changes in the current (just as a large m makes it hard to change the velocity). And $1/C$ gives a measure of the restoring force; the smaller C is, the smaller the charge is that the capacitor wants to hold, so the larger the restoring force is that tries to keep Q from getting larger (just as a large k makes it hard for x to become large).

1.2 Damped oscillations

1.2.1 Solving the differential equation

Let's now consider damped harmonic motion, where in addition to the spring force we also have a damping force. We'll assume that this damping force takes the form,

$$F_{\text{damping}} = -b\dot{x}. \tag{46}$$

Note that this is *not* the force from sliding friction on a table. That would be a force with constant magnitude $\mu_{\text{k}}N$. The $-b\dot{x}$ force here pertains to a body moving through a fluid, provided that the velocity isn't too large. So it is in fact a realistic force. The $F = ma$ equation for the mass is

$$\begin{aligned} F_{\text{spring}} + F_{\text{damping}} &= m\ddot{x} \\ \Longrightarrow \quad -kx - b\dot{x} &= m\ddot{x} \\ \Longrightarrow \quad \ddot{x} + \gamma\dot{x} + \omega_0^2 x &= 0, \qquad \text{where} \qquad \omega_0 \equiv \sqrt{\frac{k}{m}}, \qquad \gamma \equiv \frac{b}{m}. \end{aligned} \tag{47}$$

We'll use ω_0 to denote $\sqrt{k/m}$ here instead of the ω we used in Section 1.1, because there will be a couple frequencies floating around in this section, so it will be helpful to distinguish them.

In view of the discussion in Section 1.1.4, the method we will always use to solve a linear differential equation like this is to try an exponential solution,

$$x(t) = Ce^{\alpha t}. \tag{48}$$

Plugging this into Eq. (47) gives

$$\begin{aligned} & \alpha^2 Ce^{\alpha t} + \gamma\alpha Ce^{\alpha t} + \omega_0^2 Ce^{\alpha t} = 0 \\ \Longrightarrow \quad & \alpha^2 + \gamma\alpha + \omega_0^2 = 0 \\ \Longrightarrow \quad & \alpha = \frac{-\gamma \pm \sqrt{\gamma^2 - 4\omega_0^2}}{2}. \end{aligned} \tag{49}$$

We now have three cases to consider, depending on the sign of $\gamma^2 - 4\omega_0^2$. These cases are called *underdamping*, *overdamping*, and *critical damping*.

1.2.2 Underdamping ($\gamma < 2\omega_0$)

The first of the three possible cases is the case of light damping, which holds if $\gamma < 2\omega_0$. In this case, the discriminant in Eq. (49) is negative, so α has a complex part.[5] Let's define the real quantity ω_{u} (where the "u" stands for underdamped) by

$$\omega_{\text{u}} \equiv \frac{1}{2}\sqrt{4\omega_0^2 - \gamma^2} \quad \Longrightarrow \quad \boxed{\omega_{\text{u}} = \omega_0\sqrt{1 - \left(\frac{\gamma}{2\omega_0}\right)^2}} \tag{50}$$

Then the α in Eq. (49) becomes $\alpha = -\gamma/2 \pm i\omega_{\text{u}}$. So we have the two solutions,

$$x_1(t) = C_1 e^{(-\gamma/2 + i\omega_{\text{u}})t} \qquad \text{and} \qquad x_2(t) = C_2 e^{(-\gamma/2 - i\omega_{\text{u}})t}. \tag{51}$$

[5]This reminds me of a joke: The reason why life is complex is because it has both a real part and an imaginary part.

We'll accept here the fact that a second-order differential equation (which is what Eq. (47) is) has at most two linearly independent solutions. Therefore, the most general solution is the sum of the above two solutions, which is

$$x(t) = e^{-\gamma t/2}\left(C_1 e^{i\omega_u t} + C_2 e^{-i\omega_u t}\right). \tag{52}$$

However, as we discussed in Section 1.1.5, $x(t)$ must of course be real. So the two terms here must be complex conjugates of each other, to make the imaginary parts cancel. This implies that $C_2 = C_1^*$, where the star denotes complex conjugation. If we let $C_1 = Ce^{i\phi}$ then $C_2 = C_1^* = Ce^{-i\phi}$, and so $x(t)$ becomes

$$\begin{aligned} x_{\text{underdamped}}(t) &= e^{-\gamma t/2} C\left(e^{i(\omega_u t+\phi)} + e^{-i(\omega_u t+\phi)}\right) \\ &= e^{-\gamma t/2} C \cdot 2\cos(\omega_u t + \phi) \\ &\equiv \boxed{Ae^{-\gamma t/2}\cos(\omega_u t+\phi)} \qquad \text{(where } A \equiv 2C). \end{aligned} \tag{53}$$

As we mentioned in Section 1.1.5, we've basically just taken the real part of either of the two solutions that appear in Eq. (52).

We see that we have sinusoidal motion that slowly decreases in amplitude due to the $e^{-\gamma t/2}$ factor. In other words, the curves $\pm Ae^{-\gamma t/2}$ form the envelope of the sinusoidal motion. The constants A and ϕ are determined by the initial conditions, whereas the constants γ and ω_u (which is given by Eq. (50)) are determined by the setup. The task of Problem [to be added] is to find A and ϕ for the initial conditions, $x(0) = x_0$ and $v(0) = 0$. Fig. 12 shows a plot of $x(t)$ for $\gamma = 0.2$ and $\omega_0 = 1\,\text{s}^{-1}$, with the initial conditions of $x(0) = 1$ and $v(0) = 0$.

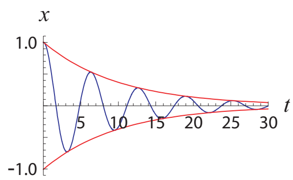


Figure 12

Note that the frequency $\omega_u = \omega_0\sqrt{1-(\gamma/2\omega_0)^2}$ is smaller than the natural frequency, ω_0. However, this distinction is generally irrelevant, because if γ is large enough to make ω_u differ appreciably from ω_0, then the motion becomes negligible after a few cycles anyway. For example, if ω_u differs from ω_0 by even just 20% (so that $\omega_u = (0.8)\omega_0$), then you can show that this implies that $\gamma = (1.2)\omega_0$. So after just two cycles (that is, when $\omega_u t = 4\pi$), the damping factor equals

$$e^{-(\gamma/2)t} = e^{-(0.6)\omega_0 t} = e^{-(0.6/0.8)\omega_u t} = e^{-(3/4)(4\pi)} = e^{-3\pi} \approx 1 \cdot 10^{-4}, \tag{54}$$

which is quite small.

Very light damping ($\gamma \ll \omega_0$)

In the limit of very light damping (that is, $\gamma \ll \omega_0$), we can use the Taylor-series approximation $\sqrt{1+\epsilon} \approx 1 + \epsilon/2$ in the expression for ω_u in Eq. (50) to write

$$\omega_u = \omega_0\sqrt{1-\left(\frac{\gamma}{2\omega_0}\right)^2} \approx \omega_0\left(1 - \frac{1}{2}\left(\frac{\gamma}{2\omega_0}\right)^2\right) = \omega_0 - \frac{\gamma^2}{8\omega_0^2}. \tag{55}$$

So ω_u essentially equals ω_0, at least to first order in γ.

Energy

Let's find the energy of an underdamped oscillator, $E = m\dot{x}^2/2 + kx^2/2$, as a function of time. To keep things from getting too messy, we'll let the phase ϕ in the expression for $x(t)$ in Eq. (53) be zero. The associated velocity is then

$$v = \frac{dx}{dt} = Ae^{-\gamma t/2}\left(-\frac{\gamma}{2}\cos\omega_u t - \omega_u \sin\omega_u t\right). \tag{56}$$

The energy is therefore

$$\begin{aligned} E &= \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2 \\ &= \frac{1}{2}mA^2e^{-\gamma t}\left(-\frac{\gamma}{2}\cos\omega_{\rm u}t - \omega_{\rm u}\sin\omega_{\rm u}t\right)^2 + \frac{1}{2}kA^2e^{-\gamma t}\cos^2\omega_{\rm u}t. \end{aligned} \tag{57}$$

Using the definition of $\omega_{\rm u}$ from Eq. (50), and using $k = m\omega_0^2$, this becomes

$$\begin{aligned} E &= \frac{1}{2}mA^2e^{-\gamma t}\left(\frac{\gamma^2}{4}\cos^2\omega_{\rm u}t + \gamma\omega_{\rm u}\cos\omega_{\rm u}t\sin\omega_{\rm u}t + \left(\omega_0^2 - \frac{\gamma^2}{4}\right)\sin^2\omega_{\rm u}t + \omega_0^2\cos^2\omega_{\rm u}t\right) \\ &= \frac{1}{2}mA^2e^{-\gamma t}\left(\frac{\gamma^2}{4}\Big(\cos^2\omega_{\rm u}t - \sin^2\omega_{\rm u}t\Big) + \gamma\omega_{\rm u}\cos\omega_{\rm u}t\sin\omega_{\rm u}t + \omega_0^2\Big(\cos^2\omega_{\rm u}t + \sin^2\omega_{\rm u}t\Big)\right) \\ &= \frac{1}{2}mA^2e^{-\gamma t}\left(\frac{\gamma^2}{4}\cos 2\omega_{\rm u}t + \frac{\gamma\omega_{\rm u}}{2}\sin 2\omega_{\rm u}t + \omega_0^2\right). \end{aligned} \tag{58}$$

As a double check, when $\gamma = 0$ this reduces to the correct value of $E = m\omega_0^2A^2/2 = kA^2/2$. For nonzero γ, the energy decreases in time due to the $e^{-\gamma t}$ factor. The lost energy goes into heat that arises from the damping force.

Note that the oscillating parts of E have frequency $2\omega_{\rm u}$. This is due to the fact that the forward and backward motions in each cycle are equivalent as far as the energy loss due to damping goes.

Eq. (58) is an exact result, but let's now work in the approximation where γ is very small, so that the $e^{-\gamma t}$ factor decays very slowly. In this approximation, the motion looks essentially sinusoidal with a roughly constant amplitude over a few oscillations. So if we take the average of the energy over a few cycles (or even just exactly one cycle), the oscillatory terms in Eq. (58) average to zero, so we're left with

$$\langle E\rangle = \frac{1}{2}m\omega_0^2A^2e^{-\gamma t} = \frac{1}{2}kA^2e^{-\gamma t}, \tag{59}$$

where the brackets denote the average over a few cycles. In retrospect, we could have obtained this small-γ result without going through the calculation in Eq. (58). If γ is very small, then Eq. (53) tells us that at any given time we essentially have a simple harmonic oscillator with amplitude $Ae^{-\gamma t/2}$, which is roughly constant. The energy of this oscillator is the usual $(k/2)(\text{amplitude})^2$, which gives Eq. (59).

Energy decay

What is the rate of change of the average energy in Eq. (59)? Taking the derivative gives

$$\boxed{\frac{d\langle E\rangle}{dt} = -\gamma\langle E\rangle} \tag{60}$$

This tells us that the *fractional* rate of change of $\langle E\rangle$ is γ. That is, in one unit of time, $\langle E\rangle$ loses a fraction γ of its value. However, this result holds only for small γ, and furthermore it holds only in an average sense. Let's now look at the exact rate of change of the energy as a function of time, for any value of γ, not just small γ.

The energy of the oscillator is $E = m\dot{x}^2/2 + kx^2/2$. So the rate of change is

$$\frac{dE}{dt} = m\dot{x}\ddot{x} + kx\dot{x} = (m\ddot{x} + kx)\dot{x}. \tag{61}$$

Since the original $F = ma$ equation was $m\ddot{x} + b\dot{x} + kx = 0$, we have $m\ddot{x} + kx = -b\dot{x}$, Therefore,

$$\frac{dE}{dt} = (-b\dot{x})\dot{x} \quad \Longrightarrow \quad \boxed{\frac{dE}{dt} = -b\dot{x}^2} \tag{62}$$

This is always negative, which makes sense because energy is always being lost to the damping force. In the limit where $b = 0$, we have $dE/dt = 0$, which makes sense because we simply have undamped simple harmonic motion, for which we already know that the energy is conserved.

We could actually have obtained this $-b\dot{x}^2$ result with the following quicker reasoning. The damping force is $F_{\text{damping}} = -b\dot{x}$, so the power (which is dE/dt) it produces is $F_{\text{damping}}v = (-b\dot{x})\dot{x} = -b\dot{x}^2$.

Having derived the exact $dE/dt = -b\dot{x}^2$ result, we can give another derivation of the result for $\langle E\rangle$ in Eq. (60). If we take the average of $dE/dt = -b\dot{x}^2$ over a few cycles, we obtain (using the fact that the average of the rate of change equals the rate of change of the average)

$$\frac{d\langle E\rangle}{dt} = -b\langle \dot{x}^2\rangle. \tag{63}$$

We now note that the average energy over a few cycles equals twice the average kinetic energy (and also twice the average potential energy), because the averages of the kinetic and potential energies are equal (see Fig. 8). Therefore,

$$\langle E\rangle = m\langle \dot{x}^2\rangle. \tag{64}$$

Comparing Eqs. (63) and (64) yields

$$\frac{d\langle E\rangle}{dt} = -\frac{b}{m}\langle E\rangle \equiv -\gamma\langle E\rangle, \tag{65}$$

in agreement with Eq. (60). Basically, the averages of both the damping power and the kinetic energy are proportional to $\langle \dot{x}^2\rangle$. And the ratio of the proportionality constants is $-b/m \equiv -\gamma$.

Q value

The ratio γ/ω_0 (or its inverse, ω_0/γ) comes up often (for example, in Eq. (50)), so let's define

$$\boxed{Q \equiv \frac{\omega_0}{\gamma}} \tag{66}$$

Q is dimensionless, so it is simply a number. Small damping means large Q. The Q stands for "quality," so an oscillator with small damping has a high quality, which is a reasonable word to use. A given damped-oscillator system has particular values of γ and ω_0 (see Eq. (47)), so it therefore has a particular value of Q. Since Q is simply a number, a reasonable question to ask is: By what factor has the amplitude decreased after Q cycles? If we consider the case of very small damping (which is reasonable, because if the damping isn't small, the oscillations die out quickly, so there's not much to look at), it turns out that the answer is independent of Q. This can be seen from the following reasoning.

The time to complete Q cycles is given by $\omega_{\text{u}}t = Q(2\pi) \implies t = 2\pi Q/\omega_{\text{u}}$. In the case of very light damping ($\gamma \ll \omega_0$), Eq. (50) gives $\omega_{\text{u}} \approx \omega_0$, so we have $t \approx 2\pi Q/\omega_0$. But since we defined $Q \equiv \omega_0/\gamma$, this time equals $t \approx 2\pi(\omega_0/\gamma)/\omega_0 = 2\pi/\gamma$. Eq. (53) then tells us that at this time, the amplitude has decreased by a factor of

$$e^{-\gamma t/2} = e^{-(\gamma/2)(2\pi/\gamma)} = e^{-\pi} \approx 0.043, \tag{67}$$

which is a nice answer if there ever was one! This result provides an easy way to determine Q, and hence γ. Just count the number of cycles until the amplitude is about 4.3% of the original value. This number equals Q, and then Eq. (66) yields γ, assuming that we know the value of ω_0.

1.2.3 Overdamping ($\gamma > 2\omega_0$)

If $\gamma > 2\omega_0$, then the two solutions for α in Eq. (49) are both real. Let's define μ_1 and μ_2 by

$$\mu_1 \equiv \frac{\gamma}{2} + \sqrt{\frac{\gamma^2}{4} - \omega_0^2}\,, \qquad \text{and} \qquad \mu_2 \equiv \frac{\gamma}{2} - \sqrt{\frac{\gamma^2}{4} - \omega_0^2}\,. \tag{68}$$

The most general solution for $x(t)$ can then be written as

$$x_{\text{overdamped}}(t) = \boxed{C_1 e^{-\mu_1 t} + C_2 e^{-\mu_2 t}} \tag{69}$$

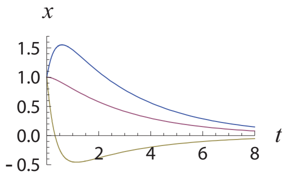


Figure 13

where C_1 and C_2 are determined by the initial conditions. Note that both μ_1 and μ_2 are positive, so both of these terms undergo exponential decay, and not exponential growth (which would be a problem physically). Since $\mu_1 > \mu_2$, the first term decays more quickly than the second, so for large t we are essentially left with only the $C_2 e^{-\mu_2 t}$ term.

The plot of $x(t)$ might look like any of the plots in Fig. 13, depending on whether you throw the mass away from the origin, release it from rest, or throw it back (fairly quickly) toward the origin. In any event, the curve can cross the origin at most once, because if we set $x(t) = 0$, we find

$$C_1 e^{-\mu_1 t} + C_2 e^{-\mu_2 t} = 0 \implies -\frac{C_1}{C_2} = e^{(\mu_1 - \mu_2)t} \implies t = \frac{1}{\mu_1 - \mu_2} \ln\left(\frac{-C_1}{C_2}\right). \tag{70}$$

We have found at most one solution for t, as we wanted to show. In a little more detail, the various cases are: There is one positive solution for t if $-C_1/C_2 > 1$; one zero solution if $-C_1/C_2 = 1$; one negative solution if $0 < -C_1/C_2 < 1$; and no (real) solution if $-C_1/C_2 < 0$. Only in the first of these four cases does that mass cross the origin at some later time after you release/throw it (assuming that this moment corresponds to $t = 0$).

Very heavy damping ($\gamma \gg \omega_0$)

Consider the limit where $\gamma \gg \omega_0$. This corresponds to a very weak spring (small ω_0) immersed in a very thick fluid (large γ), such a molasses. If $\gamma \gg \omega_0$, then Eq. (68) gives $\mu_1 \approx \gamma$. And if we use a Taylor series for μ_2 we find

$$\mu_2 = \frac{\gamma}{2} - \frac{\gamma}{2}\sqrt{1 - \frac{4\omega_0^2}{\gamma^2}} \approx \frac{\gamma}{2} - \frac{\gamma}{2}\left(1 - \frac{1}{2}\frac{4\omega_0^2}{\gamma^2}\right) = \frac{\omega_0^2}{\gamma} \ll \gamma. \tag{71}$$

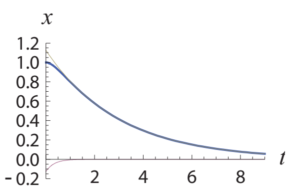


Figure 14

We therefore see that $\mu_1 \gg \mu_2$, which means that the $e^{-\mu_1 t}$ term goes to zero much faster than the $e^{-\mu_2 t}$ term. So if we ignore the quickly-decaying $e^{-\mu_1 t}$ term, Eq. (69) becomes

$$x(t) \approx C_2 e^{-\mu_2 t} \approx C_2 e^{-(\omega_0^2/\gamma)t} \equiv C_2 e^{-t/T}, \qquad \text{where} \qquad T \equiv \frac{\gamma}{\omega_0^2}\,. \tag{72}$$

A plot of a very heavily damped oscillator is shown in Fig. 14. We have chosen $\omega_0 = 1\,\text{s}^{-1}$ and $\gamma = 3\,\text{s}^{-1}$. The initial conditions are $x_0 = 1$ and $v_0 = 0$. The two separate exponential decays are shown, along with their sum. This figure makes it clear that γ doesn't have to be much larger than ω_0 for the heavy-damping approximation to hold. Even with γ/ω_0 only equal to 3, the fast-decay term dies out on a time scale of $1/\mu_1 \approx 1/\gamma = (1/3)\,\text{s}$, and the slow-decay term dies out on a time scale of $1/\mu_2 \approx \gamma/\omega_0^2 = 3\,\text{s}$.

$T = \gamma/\omega_0^2$ is called the "relaxation time." The displacement decreases by a factor of $1/e$ for every increase of T in the time. If $\gamma \gg \omega_0$, we have $T \equiv \gamma/\omega_0^2 \gg 1/\omega_0$. In other words, T is much larger than the natural period of the spring, $2\pi/\omega_0$. The mass slowly creeps back toward the origin, as you would expect for a weak spring in molasses.

Note that $T \equiv \gamma/\omega_0^2 \equiv (b/m)(k/m) = b/k$. So Eq. (72) becomes $x(t) \approx C_2 e^{-(k/b)t}$. What is the damping force is associated with this $x(t)$? We have

$$F_{\text{damping}} = -b\dot{x} = -b\left(-\frac{k}{b}C_2 e^{-(k/b)t}\right) = k\left(C_2 e^{-(k/b)t}\right) = k \cdot x(t) = -F_{\text{spring}}. \tag{73}$$

This makes sense. If we have a weak spring immersed in a thick fluid, the mass is hardly moving (or more relevantly, hardly accelerating). So the drag force and the spring force must essentially cancel each other. This also makes it clear why the relaxation time, $T = b/k$, is independent of the mass. Since the mass is hardly moving, its inertia (that is, its mass) is irrelevant. The only things that matter are the (essentially) equal and opposite spring and damping forces. So the only quantities that matter are b and k.

1.2.4 Critical damping ($\gamma = 2\omega_0$)

If $\gamma = 2\omega_0$, then we have a problem with our method of solving for $x(t)$, because the two α's in Eq. (49) are equal, since the discriminant is zero. Both solutions are equal to $-\gamma/2$, which equals ω_0 because we're assuming $\gamma = 2\omega_0$. So we haven't actually found two independent solutions, which means that we won't be able to satisfy arbitrary initial conditions for $x(0)$ and $v(0)$. This isn't good. We must somehow find another solution, in addition to the $e^{-\omega_0 t}$ one.

It turns out that $te^{-\omega_0 t}$ is also a solution to the $F = ma$ equation, $\ddot{x} + 2\omega_0\dot{x} + \omega_0^2 x = 0$ (we have used $\gamma = 2\omega_0$ here). Let's verify this. First note that

$$\begin{aligned} \dot{x} &= \frac{d}{dt}\left(te^{-\omega_0 t}\right) = e^{-\omega_0 t}(1 - \omega_0 t), \\ \Longrightarrow \quad \ddot{x} &= \frac{d}{dt}\left(e^{-\omega_0 t}(1 - \omega_0 t)\right) = e^{-\omega_0 t}\big(-\omega_0 - \omega_0(1 - \omega_0 t)\big). \end{aligned} \tag{74}$$

Therefore,

$$\ddot{x} + 2\omega_0\dot{x} + \omega_0^2 x = e^{-\omega_0 t}\Big((-2\omega_0 + \omega_0^2 t) + 2\omega_0(1 - \omega_0 t) + \omega_0^2 t\Big) = 0, \tag{75}$$

as desired. Why did we consider the function $te^{-\omega_0 t}$ as a possible solution? Well, it's a general result from the theory of differential equations that if a root α of the characteristic equation is repeated k times, then

$$e^{\alpha t}, \quad te^{\alpha t}, \quad t^2 e^{\alpha t}, \quad \cdots, \quad t^{k-1}e^{\alpha t} \tag{76}$$

are all solutions to the original differential equation. But you actually don't need to invoke this result. You can just take the limit, as $\gamma \to 2\omega_0$, of either of the underdamped or overdamped solutions. You will find that you end up with a $e^{-\omega_0 t}$ and a $te^{-\omega_0 t}$ solution (see Problem [to be added]). So we have

$$x_{\text{critical}}(t) = \boxed{(A + Bt)e^{-\omega_0 t}} \tag{77}$$

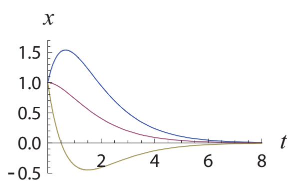


Figure 15

A plot of this is shown in Fig. 15. It looks basically the same as the overdamped plot in Fig. 13. But there is an important difference. The critically damped motion has the property that it converges to the origin in the quickest manner, that is, quicker than both the overdamped or underdamped motions. This can be seen as follows.

- QUICKER THAN OVERDAMPED: From Eq. (77), the critically damped motion goes to zero like $e^{-\omega_0 t}$ (the Bt term is inconsequential compared with the exponential term). And from Eq. (69), the overdamped motion goes to zero like $e^{-\mu_2 t}$ (since $\mu_1 > \mu_2$). But from the definition of μ_2 in Eq. (68), you can show that $\mu_2 < \omega_0$ (see Problem [to be added]). Therefore, $e^{-\omega_0 t} < e^{-\mu_2 t}$, so $x_{\text{critical}}(t) < x_{\text{overdamped}}(t)$ for large t, as desired.

- QUICKER THAN UNDERDAMPED: As above, the critically damped motion goes to zero like $e^{-\omega_0 t}$. And from Eq. (53), the envelope of the underdamped motion goes to zero like $e^{-(\gamma/2)t}$. But the assumption of underdamping is that $\gamma < 2\omega_0$, which means that $\gamma/2 < \omega_0$. Therefore, $e^{-\omega_0 t} < e^{-(\gamma/2)t}$, so $x_{\text{critical}}(t) < x^{(\text{envelope})}_{\text{underdamped}}(t)$ for large t, as desired. The underdamped motion reaches the origin first, of course, but it doesn't stay there. It overshoots and oscillates back and forth. The critically damped oscillator has the property that it converges to zero quickest without overshooting. This is very relevant when dealing with, for example, screen doors or car shock absorbers. After going over a bump in a car, you want the car to settle down to equilibrium as quickly as possible without bouncing around.

1.3 Driven and damped oscillations

1.3.1 Solving for $x(t)$

Having looked at damped oscillators, let's now look at damped and *driven* oscillators. We'll take the driving force to be $F_{\text{driving}}(t) = F_{\text{d}} \cos\omega t$. The driving frequency ω is in general equal to neither the natural frequency of the oscillator, $\omega_0 = \sqrt{k/m}$, nor the frequency of the underdamped oscillator, ω_{u}. However, we'll find that interesting things happen when $\omega_0 = \omega$. To avoid any confusion, let's explicitly list the various frequencies:

- ω_0: the natural frequency of a simple oscillator, $\sqrt{k/m}$.
- ω_{u}: the frequency of an underdamped oscillator, $\sqrt{\omega_0^2 - \gamma^2/4}$.
- ω: the frequency of the driving force, which you are free to pick.

There are two reasons why we choose to consider a force of the form $\cos\omega t$ (a $\sin\omega t$ form would work just as well). The first is due to the form of our $F = ma$ equation:

$$\begin{aligned} F_{\text{spring}} + F_{\text{damping}} + F_{\text{driving}} &= ma \\ \Longrightarrow\ -kx - b\dot{x} + F_{\text{d}}\cos\omega t &= m\ddot{x}. \end{aligned} \tag{78}$$

This is simply Eq. (47) with the additional driving force tacked on. The crucial property of Eq. (78) is that it is *linear* in x. So if we solve the equation and produce the function $x_1(t)$ for one driving force $F_1(t)$, and then solve it again and produce the function $x_2(t)$ for another driving force $F_2(t)$, then the sum of the x's is the solution to the situation where both forces are present. To see this, simply write down Eq. (78) for $x_1(t)$, and then again for $x_2(t)$, and then add the equations. The result is

$$-k(x_1 + x_2) - b(\dot{x}_1 + \dot{x}_2) + (F_1 + F_2) = m(\ddot{x}_1 + \ddot{x}_2). \tag{79}$$

In other words, $x_1(t) + x_2(t)$ is the solution for the force $F_1(t) + F_2(t)$. It's easy to see that this procedure works for any number of functions, not just two. It even works for an infinite number of functions.

The reason why this "superposition" result is so important is that when we get to *Fourier analysis* in Chapter 3, we'll see that *any* general function (well, as long as it's reasonably well behaved, which will be the case for any function we'll be interested in) can be written as the sum (or integral) of $\cos\omega t$ and $\sin\omega t$ terms with various values of ω. So if we use this fact to write an arbitrary force in terms of sines and cosines, then from the preceding paragraph, if we can solve the special case where the force is proportional to $\cos\omega t$ (or $\sin\omega t$), then we can add up the solutions (with appropriate coefficients that depend on the details of Fourier analysis) to obtain the solution for the original arbitrary force. To sum up, the combination of *linearity* and *Fourier analysis* tells us that it is sufficient to figure out how to solve Eq. (78) in the specific case where the force takes the form of $F_{\rm d}\cos\omega t$.

The second reason why we choose to consider a force of the form $\cos\omega t$ is that $F(t) = F_{\rm d}\cos\omega t$ is in fact a very realistic force. Consider a spring that has one end attached to a mass and the other end attached to a support. If the support is vibrating with position $x_{\rm end}(t) = A_{\rm end}\cos\omega t$ (which often happens in real life), then the spring force is

$$F_{\rm spring}(x) = -k(x - x_{\rm end}) = -kx + kA_{\rm end}\cos\omega t. \tag{80}$$

This is exactly the same as a non-vibrating support, with the addition of someone exerting a force $F_{\rm d}\cos\omega t$ directly on the mass, with $F_{\rm d} = kA_{\rm end}$.

We'll now solve for $x(t)$ in the case of damped and driven motion. That is, we'll solve Eq. (78), which we'll write in the form,

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = F\cos\omega t, \qquad \text{where } \gamma \equiv \frac{b}{m}, \quad \omega_0 \equiv \sqrt{\frac{k}{m}}, \quad F \equiv \frac{F_{\rm d}}{m}. \tag{81}$$

There are (at least) three ways to solve this, all of which involve guessing a sinusoidal or exponential solution.

Method 1

Let's try a solution of the form,

$$x(t) = A\cos(\omega t + \phi), \tag{82}$$

where the ω here is the *same* as the driving frequency. If we tried a different frequency, then the lefthand side of Eq. (81) would have this different frequency (the derivatives don't affect the frequency), so it would have no chance of equaling the $F\cos\omega t$ term on the righthand side.

Note how the strategy of guessing Eq. (82) differs from the strategy of guessing Eq. (48) in the damped case. The goal there was to *find* the frequency of the motion, whereas in the present case we're assuming that it equals the driving frequency ω. It might very well be the case that there doesn't exist a solution with this frequency, but we have nothing to lose by trying. Another difference between the present case and the damped case is that we will actually be able to solve for A and ϕ, whereas in the damped case these parameters could take on any values, until the initial conditions are specified.

If we plug $x(t) = A\cos(\omega t + \phi)$ into Eq. (81), we obtain

$$-\omega^2 A\cos(\omega t + \phi) - \gamma\omega A\sin(\omega t + \phi) + \omega_0^2 A\cos(\omega t + \phi) = F\cos\omega t. \tag{83}$$

We can cleverly rewrite this as

$$\omega^2 A\cos(\omega t + \phi + \pi) + \gamma\omega A\cos(\omega t + \phi + \pi/2) + \omega_0^2 A\cos(\omega t + \phi) = F\cos\omega t, \tag{84}$$

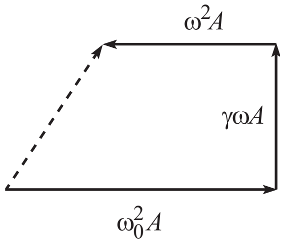


Figure 16

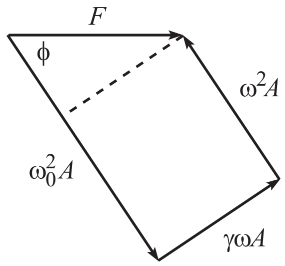


Figure 17

which just says that a is 90° ahead of v, which itself is 90° ahead of x. There happens to be a slick geometrical interpretation of this equation that allows us to quickly solve for A and ϕ. Consider the diagram in Fig. 16. The quantities ω_0, γ, ω, and F are given. We have picked an arbitrary value of A and formed a vector with length $\omega_0^2 A$ pointing to the right. Then we've added on a vector with length $\gamma\omega A$ pointing up, and then another vector with length $\omega^2 A$ point to the left. The sum is the dotted-line vector shown. We can make the magnitude of this vector be as small or as large as we want by scaling the diagram down or up with an arbitrary choice of A.

If we pick A so that the magnitude of the vector sum equals F, and if we rotate the whole diagram through the angle ϕ that makes the sum horizontal (ϕ is negative here), then we end up with the diagram in Fig. 17. The benefit of forming this diagram is the following. Consider the horizontal projections of all the vectors. The fact that the sum of the three tilted vectors equals the horizontal vector implies that the sum of their three horizontal components equals F. That is (remember that ϕ is negative),

$$\omega_0^2 A\cos\phi + \gamma\omega A\cos(\phi+\pi/2) + \omega^2 A\cos(\phi+\pi) = F\cos(0). \tag{85}$$

This is just the statement that Eq. (84) holds when $t = 0$. However, if A and ϕ are chosen so that it holds at $t = 0$, then it holds at any other time too, because we can imagine rotating the entire figure counterclockwise through the angle ωt. This simply increases the arguments of *all* the cosines by ωt. The statement that the x components of the rotated vectors add up correctly (which they do, because the figure keeps the same shape as it is rotated, so the sum of the three originally-tilted vectors still equals the originally-horizontal vector) is then

$$\omega_0^2 A\cos(\omega t+\phi) + \gamma\omega A\cos(\omega t+\phi+\pi/2) + \omega^2 A\cos(\omega t+\phi+\pi) = F\cos\omega t, \tag{86}$$

which is the same as Eq. (84), with the terms on the left in reverse order. Our task therefore reduces to determining the values of A and ϕ that generate the quadrilateral in Fig. 17.

The phase ϕ

If we look at the right triangle formed by drawing the dotted line shown, we can quickly read off

$$\tan\phi = \frac{-\gamma\omega A}{(\omega_0^2-\omega^2)A} \quad\Longrightarrow\quad \boxed{\tan\phi = \frac{-\gamma\omega}{\omega_0^2-\omega^2}} \tag{87}$$

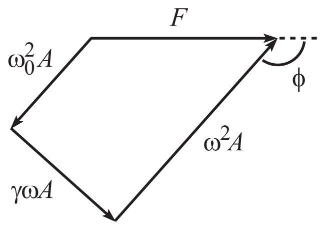


Figure 18

We've put the minus sign in by hand here because ϕ is negative. This follows from the fact that we made a "left turn" in going from the $\omega_0^2 A$ vector to the $\gamma\omega A$ vector. And then another left turn in going from the $\gamma\omega A$ vector to the $\omega^2 A$ vector. So no matter what the value of ω is, ϕ must lie somewhere in the range $-\pi \le \phi \le 0$. Angles less than $-90°$ arise when $\omega > \omega_0$, as shown in Fig. 18. $\phi \approx 0$ is obtained when $\omega \approx 0$, and $\phi \approx -\pi$ is obtained when $\omega \approx \infty$. Plots of $\phi(\omega)$ for a few different values of γ are shown in Fig. 19.

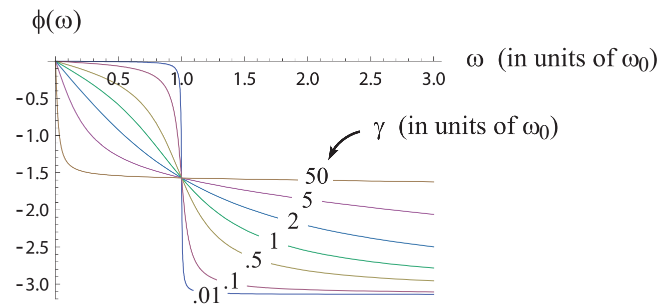


Figure 19

If γ is small, then the ϕ curve starts out with the roughly constant value of zero, and then jumps quickly to the roughly constant value of $-\pi$. The jump takes place in a small range of ω near ω_0. Problem [to be added] addresses this phenomenon. If γ is large, then the ϕ curve quickly drops to $-\pi/2$, and then very slowly decreases to π. Nothing interesting happens at $\omega = \omega_0$ in this case. See Problem [to be added].

Note that for small γ, the ϕ curve has an inflection point (which is point where the second derivative is zero), but for large γ it doesn't. The value of γ that is the cutoff between these two regimes is $\gamma = \sqrt{3}\omega_0$ (see Problem [to be added]). This is just a fact of curiosity; I don't think it has any useful consequence.

The amplitude A

To find A, we can apply the Pythagorean theorem to the right triangle in Fig. 17. This gives

$$\Big((\omega_0^2 - \omega^2)A\Big)^2 + (\gamma\omega A)^2 = F^2 \quad \Longrightarrow \quad \boxed{A = \frac{F}{\sqrt{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}} \tag{88}$$

Our solution for $x(t)$ is therefore

$$\boxed{x(t) = A\cos(\omega t + \phi)} \tag{89}$$

where ϕ and A are given by Eqs. (87) and (88). Plots of the amplitude $A(\omega)$ for a few different values of γ are shown in Fig. 20.

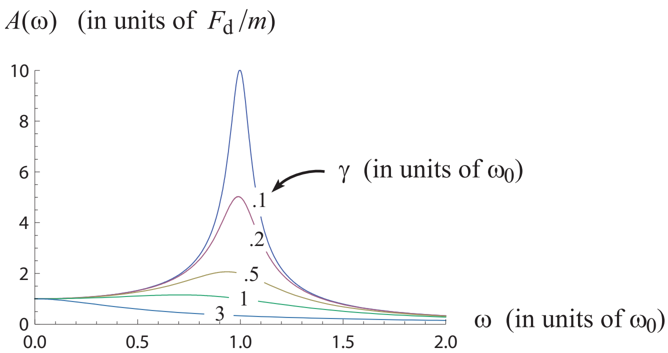


Figure 20

At what value of ω does the maximum of $A(\omega)$ occur? $A(\omega)$ is maximum when the denominator of the expression in Eq. (88) is minimum, so setting the derivative (with respect to ω) of the quantity under the square root equal to zero gives

$$2(\omega_0^2 - \omega^2)(-2\omega) + \gamma^2(2\omega) = 0 \quad \Longrightarrow \quad \omega = \sqrt{\omega_0^2 - \gamma^2/2}. \tag{90}$$

For small γ (more precisely, for $\gamma \ll \omega_0$), this yields $\omega \approx \omega_0$. If $\gamma = \sqrt{2}\omega_0$, then the maximum occurs at $\omega = 0$. If γ is larger than $\sqrt{2}\omega_0$, then the maximum occurs at $\omega = 0$, and the curve monotonically decreases as ω increases. These facts are consistent with Fig. 20. For small γ, which is the case we'll usually be concerned with, the maximum value of $A(\omega)$ is essentially equal to the value at ω_0, which is $A(\omega_0) = F/\gamma\omega_0$.

$A(\omega)$ goes to zero as $\omega \to \infty$. The value of $A(\omega)$ at $\omega = 0$ is

$$A(0) = \frac{F}{\omega_0^2} \equiv \frac{F_\mathrm{d}/m}{k/m} = \frac{F_\mathrm{d}}{k}. \tag{91}$$

This is independent of γ, consistent with Fig. 20. The reason why $A(0)$ takes the very simple form of F_d/k will become clear in Section 1.3.2 below.

Using the same techniques that we'll use below to obtain in Eq. (128) the width of the power curve, you can show (see Problem [to be added]) that the width of the $A(\omega)$ curve (defined to be the width at half max) is

$$\boxed{\text{width} = \sqrt{3}\,\gamma} \tag{92}$$

So the curves get narrower (at half height) as γ decreases.

However, the curves don't get narrower in an absolute sense. By this we mean that for a given value of ω, say $\omega = (0.9)\omega_0$, the value of A in Fig. 20 increases as γ decreases. Equivalently, for a given value of A, the width of the curve at this value increases as γ decreases. These facts follow from the fact that as $\gamma \to 0$, the A in Eq. (88) approaches the function $F/|\omega_0^2 - \omega^2|$. This function is the envelope of all the different $A(\omega)$ functions for different values of γ. If we factor the denominator in $F/|\omega_0^2 - \omega^2|$, we see that near ω_0 (but not right at ω_0), A behaves like $(F/2\omega_0)/|\omega_0 - \omega|$.

Remember that both A and ϕ are completely determined by the quantities ω_0, γ, ω, and F. The initial conditions have nothing to do with A and ϕ. How, then, can we satisfy arbitrary initial conditions with no free parameters at our disposal? We'll address this question after we discuss the other two methods for solving for $x(t)$.

Method 2

Let's try a solution of the form,

$$x(t) = A\cos\omega t + B\sin\omega t. \tag{93}$$

(This A isn't the same as the A in Method 1.) As above, the frequency here must be the same as the driving frequency if this solution is to have any chance of working. If we plug this expression into the $F = ma$ equation in Eq. (81), we get a fairly large mess. But if we group the terms according to whether they involve a $\cos\omega t$ or $\sin\omega t$, we obtain (you should verify this)

$$(-\omega^2 B - \gamma\omega A + \omega_0^2 B)\sin\omega t + (-\omega^2 A + \gamma\omega B + \omega_0^2 A)\cos\omega t = F\cos\omega t. \tag{94}$$

We have two unknowns here, A and B, but only one equation. However, this equation is actually two equations. The point is that we want it to hold for *all* values of t. But $\sin\omega t$ and $\cos\omega t$ are linearly independent functions, so the only way this equation can hold for all t is if the coefficients of $\sin\omega t$ and $\cos\omega t$ on each side of the equation match up independently. That is,

$$\begin{aligned} -\omega^2 B - \gamma\omega A + \omega_0^2 B &= 0, \\ -\omega^2 A + \gamma\omega B + \omega_0^2 A &= F. \end{aligned} \tag{95}$$

We now have two unknowns and two equations. Solving for either A or B in the first equation and plugging the result into the second one gives

$$A = \frac{(\omega_0^2 - \omega^2)F}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \qquad \text{and} \qquad B = \frac{\gamma\omega F}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}. \tag{96}$$

The solution for $x(t)$ is therefore

$$\boxed{x(t) = \frac{(\omega_0^2 - \omega^2)F}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}\cos\omega t + \frac{\gamma\omega F}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}\sin\omega t} \quad (97)$$

We'll see below in Method 3 that this solution is equivalent to the $x(t) = A\cos(\omega t + \phi)$ solution from Method 1, given that A and ϕ take on the particular values we found.

Method 3

First, consider the equation,

$$\ddot{y} + \gamma\dot{y} + \omega_0^2 y = Fe^{i\omega t}. \quad (98)$$

This equation isn't actually physical, because the driving "force" is the complex quantity $Fe^{i\omega t}$. Forces need to be real, of course. And likewise the solution for $y(t)$ will complex, so it can't actually represent an actual position. But as we'll shortly see, we'll be able to extract a physical result by taking the real part of Eq. (98).

Let's guess a solution to Eq. (98) of the form $y(t) = Ce^{i\omega t}$. When we get to Fourier analysis in Chapter 3, we'll see that this is the only function that has any possible chance of working. Plugging in $y(t) = Ce^{i\omega t}$ gives

$$-\omega^2 Ce^{i\omega t} + i\gamma\omega Ce^{i\omega t} + \omega_0^2 Ce^{i\omega t} = F \cdot Ce^{i\omega t} \quad \Longrightarrow \quad C = \frac{F}{\omega_0^2 - \omega^2 + i\gamma\omega}. \quad (99)$$

What does this solution have to do with our original scenario involving a driving force proportional to $\cos\omega t$? Well, consider what happens when we take the real part of Eq. (98). Using the fact that differentiation commutes with the act of taking the real part, which is true because

$$\text{Re}\left(\frac{d}{dt}(a + ib)\right) = \frac{da}{dt} = \frac{d}{dt}\Big(\text{Re}(a + ib)\Big), \quad (100)$$

we obtain

$$\begin{aligned} \text{Re}(\ddot{y}) + \text{Re}(\gamma\dot{y}) + \text{Re}(\omega_0^2 y) &= \text{Re}(Fe^{i\omega t}) \\ \frac{d^2}{dt^2}\Big(\text{Re}(y)\Big) + \gamma\frac{d}{dt}\Big(\text{Re}(y)\Big) + \omega_0^2\Big(\text{Re}(y)\Big) &= F\cos\omega t. \end{aligned} \quad (101)$$

In other words, if y is a solution to Eq. (98), then the real part of y is a solution to our original (physical) equation, Eq. (81), with the $F\cos\omega t$ driving force. So we just need to take the real part of the solution for y that we found, and the result will be the desired position x. That is,

$$x(t) = \text{Re}\Big(y(t)\Big) = \text{Re}\Big(Ce^{i\omega t}\Big) = \text{Re}\left(\frac{F}{\omega_0^2 - \omega^2 + i\gamma\omega}e^{i\omega t}\right). \quad (102)$$

Note that the quantity $\text{Re}(Ce^{i\omega t})$ is what matters here, and *not* $\text{Re}(C)\text{Re}(e^{i\omega t})$. The equivalence of this solution for $x(t)$ with the previous ones in Eqs. (89) and (97) can be seen as follows. Let's consider Eq. (97) first.

- Agreement with Eq. (97):

  Any complex number can be written in either the Cartesian $a + bi$ way, or the polar (magnitude and phase) $Ae^{i\phi}$ way. If we choose to write the C in Eq. (99) in the

Cartesian way, we need to get the i out of the denominator. If we "rationalize" the denominator of C and expand the $e^{i\omega t}$ term in Eq. (102), then $x(t)$ becomes

$$x(t) = \text{Re}\left(\frac{F\big((\omega_0^2 - \omega^2) - i\gamma\omega\big)}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}(\cos\omega t + i\sin\omega t)\right). \tag{103}$$

The real part comes from the product of the real parts and the product of the imaginary parts, so we obtain

$$x(t) = \frac{(\omega_0^2 - \omega^2)F}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}\cos\omega t + \frac{\gamma\omega F}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}\sin\omega t, \tag{104}$$

in agreement with Eq. (97) in Method 2.

- Agreement with Eq. (89):

  If we choose to write the C in Eq. (99) in the polar $Ae^{i\phi}$ way, we have

$$A = \sqrt{C \cdot C^*} = \frac{F}{\sqrt{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}, \tag{105}$$

  and

$$\tan\phi = \frac{\text{Im}(C)}{\text{Re}(C)} = \frac{-\gamma\omega}{\omega_0^2 - \omega^2}, \tag{106}$$

  where we have obtained the real and imaginary parts from the expression for C that we used in Eq. (103). (This expression for $\tan\phi$ comes from the fact that the ratio of the imaginary and real parts of $e^{i\phi} = \cos\phi + i\sin\phi$ equals $\tan\phi$.) So the $x(t)$ in Eq. (102) becomes

$$x(t) = \text{Re}\big(y(t)\big) = \text{Re}\big(Ce^{i\omega t}\big) = \text{Re}\big(Ae^{i\phi}e^{i\omega t}\big) = A\cos(\omega t + \phi). \tag{107}$$

  This agrees with the result obtained in Eq. (89) in Method 1, because the A and ϕ in Eqs. (105) and (106) agree with the A and ϕ in Eqs. (88) and (87).

The complete solution

Having derived the solution for $x(t)$ in three different ways, let's look at what we've found. We'll use the $x(t) = A\cos(\omega t + \phi)$ form of the solution in the following discussion.

As noted above, A and ϕ have definite values, given the values of $\omega_0 \equiv \sqrt{k/m}$, $\gamma \equiv b/m$, $F \equiv F_{\text{d}}/m$, and ω. There is no freedom to impose initial conditions. The solution in Eq. (107) therefore cannot be the most general solution, because the most general solution must allow us to be able to satisfy arbitrary initial conditions. So what is the most general solution? It is the sum of the $A\cos(\omega t + \phi)$ solution we just found (this is called the "particular" solution) and the solution from Section 1.2.1 (the "homogeneous" solution) that arose when there was no driving force, and thus a zero on the righthand side of the $F = ma$ equation in Eq. (47). This sum is indeed a solution to the $F = ma$ equation in Eq. (81) because this equation is *linear* in x. The homogeneous solution simply produces a zero on the righthand side, which doesn't mess up the equality generated by the particular solution. In equations, we have (with the sum of the particular and homogeneous solutions written as $x = x_{\text{p}} + x_{\text{h}}$)

$$\begin{aligned}
\ddot{x} + \gamma\dot{x} + \omega_0^2 x &= (\ddot{x}_{\text{p}} + \ddot{x}_{\text{h}}) + \gamma(\dot{x}_{\text{p}} + \dot{x}_{\text{h}}) + \omega_0^2(x_{\text{p}} + x_{\text{h}}) \\
&= \big(\ddot{x}_{\text{p}} + \gamma\dot{x}_{\text{p}} + \omega_0^2 x_{\text{p}}\big) + \big(\ddot{x}_{\text{h}} + \gamma\dot{x}_{\text{h}} + \omega_0^2 x_{\text{h}}\big) \\
&= F\cos\omega t + 0,
\end{aligned} \tag{108}$$

which means that $x = x_p + x_h$ is a solution to the $F = ma$ equation, as desired. The two unknown constants in the homogeneous solution yield the freedom to impose arbitrary initial conditions. For concreteness, if we assume that we have an underdamped driven oscillator, which has the homogeneous solution given by Eq. (53), then the complete solution, $x = x_p + x_h$, is

$$\boxed{x(t) = A_p \cos(\omega t + \phi) + A_h e^{-\gamma t/2} \cos(\omega_h t + \theta)} \qquad (\textit{underdamped}) \qquad (109)$$

A word on the various parameters in this result:

- ω is the driving frequency, which can be chosen arbitrarily.
- A_p and ϕ are functions of $\omega_0 \equiv \sqrt{k/m}$, $\gamma \equiv b/m$, $F \equiv F_d/m$, and ω. They are given in Eqs. (88) and (87).
- ω_h is a function of ω_0 and γ. It is given in Eq. (50).
- A_h and θ are determined by the initial conditions.

However, having said all this, we should note the following very important point. For large t (more precisely, for $t \gg 1/\gamma$), the homogeneous solution goes to zero due to the $e^{-\gamma t/2}$ term. So no matter what the initial conditions are, we're left with essentially the *same* particular solution for large t. In other words, if two different oscillators are subject to exactly the same driving force, then even if they start with wildly different initial conditions, the motions will essentially be the same for large t. All memory of the initial conditions is lost.[6] Therefore, since the particular solution is the one that survives, let's examine it more closely and discuss some special cases.

1.3.2 Special cases for ω

Slow driving ($\omega \ll \omega_0$)

If ω is very small compared with ω_0, then we can simplify the expressions for ϕ and A in Eqs. (87) and (88). Assuming that γ isn't excessively large (more precisely, assuming that $\gamma\omega \ll \omega_0^2$), we find

$$\phi \approx 0, \qquad \text{and} \qquad A \approx \frac{F}{\omega_0^2} \equiv \frac{F_d/m}{k/m} = \frac{F_d}{k}\,. \qquad (110)$$

Therefore, the position takes the form (again, we're just looking at the particular solution here),

$$x(t) = A\cos(\omega t + \phi) = \frac{F_d}{k}\cos\omega t. \qquad (111)$$

Note that the spring force is then $F_{\text{spring}} = -kx = -F_d \cos\omega t$, which is simply the negative of the driving force. In other words, the driving force essentially balances the spring force. This makes sense: The very small frequency, ω, of the motion implies that the mass is hardly moving (or more relevantly, hardly accelerating), which in turn implies that the net force must be essentially zero. The damping force is irrelevant here because the small velocity (due to the small ω) makes it negligible. So the spring force must balance the driving force. The mass and the damping force play no role in this small-frequency motion. So the effect of the driving force is to simply balance the spring force.

[6] The one exception is the case where there is no damping whatsoever, so that γ is exactly zero. But all mechanical systems have at least a tiny bit of damping (let's not worry about superfluids and such), so we'll ignore the $\gamma = 0$ case.

Mathematically, the point is that the first two terms on the lefthand side of the $F = ma$ equation in Eq. (81) are negligible:

$$\underline{\ddot{x}} + \underline{\gamma\dot{x}} + \omega_0^2 x = F\cos\omega t. \tag{112}$$

This follows from the fact that since $x(t) = A\cos(\omega t + \phi)$, the coefficients of each of the sinusoidal terms on the lefthand side are proportional to ω^2, $\gamma\omega$, and ω_0^2, respectively. And since we're assuming both $\omega \ll \omega_0$ and $\gamma\omega \ll \omega_0^2$, the first two terms are negligible compared with the third. The acceleration and velocity of the mass are negligible. The position is all that matters.

The $\phi \approx 0$ result in Eq. (110) can be seen as follows. We saw above that the driving force cancels the spring force. Another way of saying this is that the driving force is 180° out of phase with the $-kx = k(-x)$ spring force. This means that the driving force is *in* phase with the position x. Intuitively, the larger the force you apply, the larger the spring force and hence the larger the position x. The position just follows the force.

Fast driving ($\omega \gg \omega_0$)

If ω is very large compared with ω_0, then we can again simplify the expressions for ϕ and A in Eqs. (87) and (88). Assuming that γ isn't excessively large (which now means that $\gamma \ll \omega$), we find

$$\phi \approx -\pi, \qquad \text{and} \qquad A \approx \frac{F}{\omega^2} = \frac{F_\text{d}}{m\omega^2}\,. \tag{113}$$

Therefore, the position takes the form,

$$x(t) = A\cos(\omega t + \phi) = \frac{F_\text{d}}{m\omega^2}\cos(\omega t - \pi) = -\frac{F_\text{d}}{m\omega^2}\cos\omega t. \tag{114}$$

Note that the mass times the acceleration is then $m\ddot{x} = F_\text{d}\cos\omega t$, which is the driving force. In other words, the driving force is essentially solely responsible for the acceleration. This makes sense: Since there are ω's in the denominator of $x(t)$, and since ω is assumed to be large, we see that $x(t)$ is very small. The mass hardly moves, so the spring and damping forces play no role in this high-frequency motion. The driving force provides essentially all of the force and therefore causes the acceleration.

Mathematically, the point is that the second two terms on the lefthand side of the $F = ma$ equation in Eq. (81) are negligible:

$$\ddot{x} + \underline{\gamma\dot{x}} + \underline{\omega_0^2 x} = F\cos\omega t. \tag{115}$$

As we noted after Eq. (112), the coefficients of each of the sinusoidal terms on the lefthand side are proportional to ω^2, $\gamma\omega$, and ω_0^2, respectively. And since we're assuming both $\omega_0 \ll \omega$ and $\gamma \ll \omega$, the second two terms are negligible compared with the first. The velocity and position of the mass are negligible. The acceleration is all that matters.

The $\phi \approx -\pi$ result in Eq. (113) can be seen as follows. Since the driving force provides essentially all of the force, it is therefore in phase with the acceleration. But the acceleration is always out of phase with $x(t)$ (at least for sinusoidal motion). So the driving force is out of phase with $x(t)$. Hence the $\phi \approx -\pi$ result and the minus sign in the expression for $x(t)$ in Eq. (114).

Resonance ($\omega = \omega_0$)

If ω equals ω_0, then we can again simplify the expressions for ϕ and A in Eqs. (87) and (88). We don't need to make any assumptions about γ in this case, except that it isn't exactly equal to zero. We find

$$\phi = -\frac{\pi}{2}, \qquad \text{and} \qquad A \approx \frac{F}{\gamma\omega} = \frac{F}{\gamma\omega_0} = \frac{F_\text{d}}{\gamma m\omega_0}. \tag{116}$$

Therefore, the position takes the form,

$$x(t) = A\cos(\omega t + \phi) = \frac{F_\text{d}}{\gamma m\omega_0}\cos(\omega t - \pi/2) = \frac{F_\text{d}}{\gamma m\omega_0}\sin\omega t. \tag{117}$$

Note that the damping force is then $F_\text{damping} = -(\gamma m)\dot{x} = -F_\text{d}\cos\omega t$, which is the negative of the driving force. In other words, the driving force essentially balances the damping force. This makes sense: If $\omega = \omega_0$, then the system is oscillating at ω_0, so the spring and the mass are doing just what they would be doing if the damping and driving forces weren't present. You can therefore consider the system to be divided into two separate systems: One is a simple harmonic oscillator, and the other is a driving force that drags a massless object with the same shape as the original mass (so that it recreates the damping force) back and forth in a fluid (or whatever was providing the original damping force). None of the things involved here (spring, mass, fluid, driver) can tell the difference between the original system and this split system. So the effect of the driving force is to effectively cancel the damping force, while the spring and the mass do their natural thing.

Mathematically, the point is that the first and third terms on the lefthand side of the $F = ma$ equation in Eq. (81) cancel each other:

$$\underline{\ddot{x}} + \gamma\dot{x} + \underline{\omega_0^2 x} = F\cos\omega t. \tag{118}$$

As above, the coefficients of each of the sinusoidal terms on the lefthand side are proportional (in magnitude) to ω^2, $\gamma\omega$, and ω_0^2, respectively. Since $\omega = \omega_0$, the first and third terms cancel (the second derivative yields a minus sign in the first term). The remaining parts of the equation then say that the driving force balances the damping force. Note that the amplitude must take on the special value of $F_\text{d}/\gamma m\omega_0$ for this to work.

If γ is small, then the amplitude A in Eq. (116) is large. Furthermore, for a given value of γ, the amplitude is largest when (roughly) $\omega = \omega_0$. Hence the name "resonance." We've added the word "roughly" here because it depends on what we're taking to be the given quantity, and what we're taking to be the variable. If ω is given, and if we want to find the maximum value of the A in Eq. (88) as a function of ω_0, then we want to pick ω_0 to equal ω, because this choice makes $\omega_0^2 - \omega^2$ equal to zero, and we can't do any better than that, with regard to making the denominator of A small. On the other hand, if ω_0 is given, and if we want to find the maximum value of A as a function of ω, then we need to take the derivative of A with respect to ω. We did this in Eq. (90) above, and the result was $\omega = \sqrt{\omega_0^2 - \gamma^2/2}$. So the peaks of the curves in Fig. 20 (where A was considered to be a function of ω) weren't located exactly at $\omega = \omega_0$. However, we will generally be concerned with the case of small γ (more precisely $\gamma \ll \omega_0$), in which case the peak occurs essentially at $\omega = \omega_0$, even when A is considered to be a function of ω.

Having noted that the amplitude is maximum when $\omega = \omega_0$, we can now see where the $\phi \approx -\pi/2$ result in Eq. (116) comes from. If we want to make the amplitude as large as possible, then we need to put a lot of energy into the system. Therefore, we need to do a lot of work. So we want the driving force to act over the largest possible distance. This means that we want the driving force to be large when the mass velocity is large. (Basically, power is force times velocity.) In other words, we want the driving force to be in phase with the velocity. And since x is always 90° behind v, x must also be 90° behind the force. This agrees with the $\phi = -\pi/2$ phase in x. In short, this $\phi = -\pi/2$ phase implies that the force always points to the right when the mass is moving to the right, and always points to the left when the mass is moving to the left. So we're always doing positive work. For any other phase, there are times when we're doing negative work.

In view of Eqs. (112), (115), and (118), we see that the above three special cases are differentiated by which one of the terms on the lefthand side of Eq. (81) survives. There is a slight difference, though: In the first two cases, two terms disappear because they are small. In the third case, they disappear because they are equal and opposite.

1.3.3 Power

In a driven and damped oscillator, the driving force feeds energy into the system during some parts of the motion and takes energy out during other parts (except in the special case of resonance where it always feeds energy in). The damping force always takes energy out, because the damping force always points antiparallel to the velocity. For the steady-state solution (the "particular" solution), the motion is periodic, so the energy should stay the same on average; the amplitude isn't changing. The average net power (work per time) from the driving force must therefore equal the negative of the average power from the damping force. Let's verify this. Power is the rate at which work is done, so we have

$$dW = F dx \quad \Longrightarrow \quad P \equiv \frac{dW}{dt} = F\frac{dx}{dt} = Fv. \tag{119}$$

The powers from the damping and driving forces are therefore:

POWER DISSIPATED BY THE DAMPING FORCE: This equals

$$\begin{aligned} P_{\text{damping}} = F_{\text{damping}} v &= (-b\dot{x})\dot{x} \\ &= -b\big(-\omega A \sin(\omega t + \phi)\big)^2 \\ &= -b(\omega A)^2 \sin^2(\omega t + \phi). \end{aligned} \tag{120}$$

Since the average value of $\sin^2\theta$ over a complete cycle is 1/2 (obtained by either doing an integral or noting that $\sin^2\theta$ has the same average as $\cos^2\theta$, and these two averages add up to 1), the average value of the power from the damping force is

$$\langle P_{\text{damping}} \rangle = \boxed{-\frac{1}{2}b(\omega A)^2} \tag{121}$$

POWER SUPPLIED BY THE DRIVING FORCE: This equals

$$\begin{aligned} P_{\text{driving}} = F_{\text{driving}} v &= (F_{\text{d}} \cos\omega t)\dot{x} \\ &= (F_{\text{d}} \cos\omega t)\big(-\omega A \sin(\omega t + \phi)\big) \\ &= -F_{\text{d}}\omega A \cos\omega t \,\big(\sin\omega t \cos\phi + \cos\omega t \sin\phi\big). \end{aligned} \tag{122}$$

The results in Eqs. (120) and (122) aren't the negatives of each other for all t, because the energy waxes and and wanes throughout a cycle. (The one exception is on resonance with $\phi = -\pi/2$, as you can verify.) But *on average* they must cancel each other, as we noted above. This is indeed the case, because in Eq. (122), the $\cos\omega t \sin\omega t$ term averages to zero, while the $\cos^2\omega t$ term averages to 1/2. So the average value of the power from the driving force is

$$\langle P_{\text{driving}} \rangle = -\frac{1}{2}F_{\text{d}}\omega A \sin\phi. \tag{123}$$

Now, what is $\sin\phi$? From Fig. 17, we have $\sin\phi = -\gamma\omega A/F \equiv -\gamma m\omega A/F_\text{d}$. So Eq. (123) gives

$$\langle P_\text{driving}\rangle = -\frac{1}{2}F_\text{d}\omega A\left(\frac{-\gamma m\omega A}{F_\text{d}}\right) = \frac{1}{2}\gamma m(\omega A)^2 = \boxed{\frac{1}{2}b(\omega A)^2} \tag{124}$$

Eqs. (121) and (124) therefore give $\langle P_\text{damping}\rangle + \langle P_\text{driving}\rangle = 0$, as desired.

What does $\langle P_\text{driving}\rangle$ look like as a function of ω? Using the expression for A in Eq. (88), along with $b \equiv \gamma m$, we have

$$\begin{aligned}\langle P_\text{driving}\rangle &= \frac{1}{2}b(\omega A)^2\\ &= \frac{(\gamma m)\omega^2}{2}\cdot\frac{(F_\text{d}/m)^2}{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2}\\ &= \frac{(\gamma m)F_\text{d}^2}{2\gamma^2 m^2}\cdot\frac{\gamma^2\omega^2}{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2}\\ &= \boxed{\frac{F_\text{d}^2}{2\gamma m}\cdot\frac{\gamma^2\omega^2}{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2}} \equiv \frac{F_\text{d}^2}{2\gamma m}\cdot f(\omega).\end{aligned} \tag{125}$$

We have chosen to write the result this way because the function $f(\omega)$ is a dimensionless function of ω. The F_d^2 out front tells us that for given ω, ω_0, and γ, the average power $\langle P_\text{driving}\rangle$ grows as the driving amplitude F_d becomes larger, which makes sense. Fig. 21 shows some plots of the dimensionless function $f(\omega)$ for a few values of γ. In other words, it shows plots of $\langle P_\text{driving}\rangle$ in units of $F_\text{d}^2/2\gamma m \equiv F_\text{d}^2/2b$.

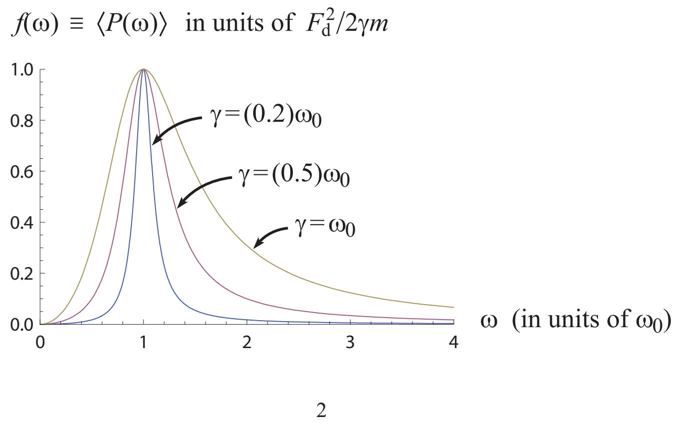


Figure 21

Fig. 22 shows plots of $f(\omega)/\gamma$ for the same values of γ. That is, it shows plots of the actual average power, $\langle P_\text{driving}\rangle$, in units of $F_\text{d}^2/2m$. These plots are simply $1/\gamma$ times the plots in Fig. 21.

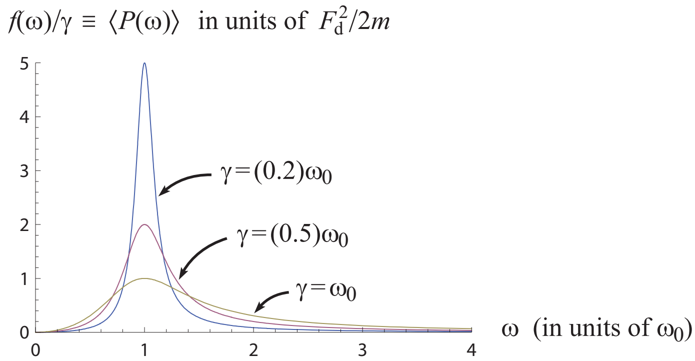


Figure 22

The curves in Fig. 22 get thinner and taller as γ gets smaller. How do the widths depend on γ? We'll define the "width" to be the width at half max. The maximum value of $\langle P_{\text{driving}}\rangle$ (or equivalently, of $f(\omega)$) is achieved where its derivative with respect to ω is zero. This happens to occur right at $\omega = \omega_0$ (see Problem [to be added]). The maximum value of $f(\omega)$ then 1, as indicated in Fig. 21 So the value at half max equals 1/2. This is obtained when the denominator of $f(\omega)$ equals $2\gamma^2\omega^2$, that is, when

$$(\omega_0^2 - \omega^2)^2 = \gamma^2\omega^2 \implies \omega_0^2 - \omega^2 = \pm\gamma\omega. \tag{126}$$

There are two quadratic equations in ω here, depending on the sign. The desired width at half max equals the difference between the positive roots, call them ω_1 and ω_2, of each of these two equations. If you want, you can use the quadratic formula to find these roots, and then take the difference. But a cleaner way is to write down the statements that ω_1 and ω_2 are solutions to the equations:

$$\begin{aligned} \omega_0^2 - \omega_1^2 &= \gamma\omega_1, \\ \omega_0^2 - \omega_2^2 &= -\gamma\omega_2, \end{aligned} \tag{127}$$

and then take the difference. The result is

$$\omega_2^2 - \omega_1^2 = \gamma(\omega_2 + \omega_1) \implies \omega_2 - \omega_1 = \gamma \implies \boxed{\text{width} = \gamma} \tag{128}$$

(We have ignored the $\omega_1 + \omega_2$ solution to this equation, since we are dealing with positive ω_1 and ω_2.) So we have the nice result that the width at half max is *exactly* equal to γ. This holds for *any* value of γ, even though the the plot of $\langle P_{\text{driving}}\rangle$ looks like a reasonably symmetric peak only if γ is small compared with ω_0. This can be see in Fig. 23, which shows plots of $f(\omega)$ for the reasonably small value of $\gamma = (0.4)\omega_0$ and the reasonably large value of $\gamma = 2\omega_0$.

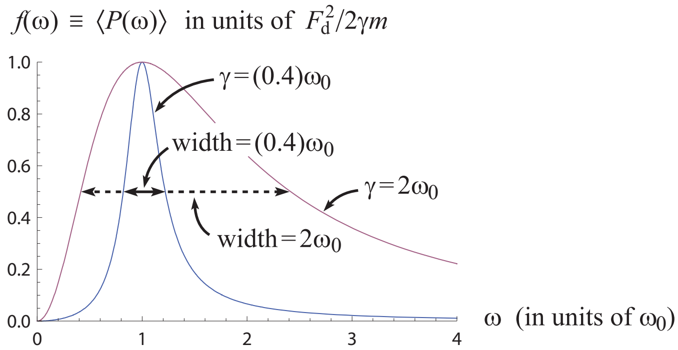


Figure 23

To sum up, the maximum height of the $\langle P_{\text{driving}}\rangle$ curve is proportional to $1/\gamma$ (it equals $F_d^2/2\gamma m$), and the width of the curve at half max is proportional to γ (it's just γ). So the curves get narrower as γ decreases.

Furthermore, the curves get narrower in an absolute sense, unlike the $A(\omega)$ curves in Fig. 20 (see the discussion at the end of the "Method 1" part of Section 1.3.1). By this we mean that for a given value of ω (except ω_0), say $\omega = (0.9)\omega_0$, the value of P in Fig. 22 decreases as γ decreases. Equivalently, for a given value of P, the width of the curve at this value decreases as γ decreases. These facts follow from the fact that as $\gamma \to 0$, the P in Eq. (125) is proportional to the function $\gamma/\big((\omega_0^2-\omega^2)^2+\gamma^2\omega^2\big)$. For any given value of ω (except ω_0), this becomes very small if γ is sufficiently small.

Since the height and width of the power curve are proportional to $1/\gamma$ and γ, respectively, we might suspect that the area under the curve is independent of γ. This is indeed the case. The integral is very messy to calculate in closed form, but if you find the area by numerical integration for a few values of γ, that should convince you.

Let's now discuss intuitively why the $\langle P_{\text{driving}}\rangle$ curve in Fig. 22 goes to zero at $\omega \approx 0$ and $\omega \approx \infty$, and why it is large at $\omega = \omega_0$.

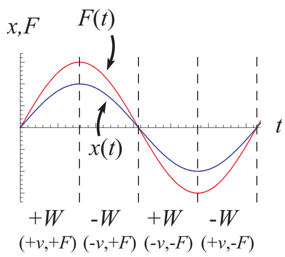


Figure 24

- $\omega \approx 0$: In this case, Eq. (110) gives the phase as $\phi \approx 0$, so the motion is in phase with the force. The consequence of this fact is that half the time your driving force is doing positive work, and half the time it is doing negative work. These cancel, and on average there is no work done. In more detail, let's look at each quarter cycle; see Fig. 24 (the graph is plotted with arbitrary units on the axes). As you (very slowly) drag the mass to the right from the origin to the maximum displacement, you are doing positive work, because your force is in the direction of the motion. But then as you (very slowly) let the spring pull the mass back toward to the origin, you are doing negative work, because your force is now in the direction opposite the motion. The same cancelation happens in the other half of the cycle.

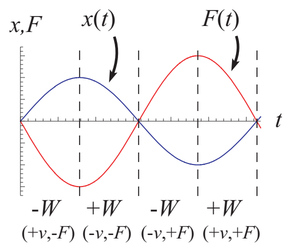


Figure 25

- $\omega \approx \infty$: In this case, Eq. (113) gives the phase as $\phi \approx -\pi$, so the motion is out of phase with the force. And again, this implies that half the time your driving force is doing positive work, and half the time it is doing negative work. So again there is cancelation. Let's look at each quarter cycle again; see Fig. 25. As the mass (very quickly) moves from the origin to the maximum displacement, you are doing negative work, because your force is in the direction opposite the motion (you are the thing that is slowing the mass down). But then as you (very quickly) drag the mass back toward to the origin, you are doing positive work, because your force is now in the direction of the motion (you are the thing that is speeding the mass up). The same cancelation happens in the other half of the cycle.

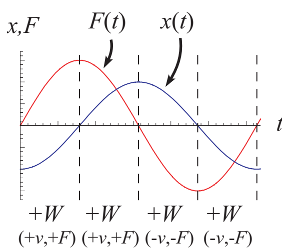


Figure 26

- $\omega = \omega_0$: In this case, Eq. (116) gives the phase as $\phi = -\pi/2$, so the motion is in "quadrature" with the force. The consequence of this fact is that your driving force is *always* doing positive work. Let's now look at each half cycle; see Fig. 26. Start with the moment when the mass has maximum negative displacement. For the next half cycle until it reaches maximum positive displacement, you are doing positive work, because both your force and the velocity point to the right. And for other half cycle where the mass move back to the maximum negative displacement, you are *also* doing positive work, because now both your force and the velocity point to the left. In short, the velocity, which is obtained by taking the derivative of the position in Eq. (117), is always in phase with the force. So you are always doing positive work, and there is no cancelation.

Q values

Recall the $Q \equiv \omega_0/\gamma$ definition in Eq. (66). Q has interpretations for both the transient ("homogeneous") solution and the steady-state ("particular") solution.

- For the transient solution, we found in Eq. (67) that Q is the number of oscillations it takes for the amplitude to decrease by a factor of $e^{-\pi} \approx 4\%$.

For the steady-state solution, there are actually two interpretations of Q.

- The first is that it equals the ratio of the amplitude at resonance to the amplitude at small ω. This can be seen from the expression for A in Eq. (88). For $\omega = \omega_0$ we have $A = F/\gamma\omega_0$, while for $\omega \approx 0$ we have $A \approx F/\omega_0^2$. Therefore,

$$\frac{A_{\text{resonance}}}{A_{\omega\approx 0}} = \frac{F/\gamma\omega_0}{F/\omega_0^2} = \frac{\omega_0}{\gamma} \equiv Q. \tag{129}$$

  So the larger the Q value, the larger the amplitude at resonance. The analogous statement in terms of power is that the larger the value of Q, the larger the $F_\text{d}^2/2\gamma m = F_\text{d}^2 Q/2\omega_0 m$ value of the power at resonance.

- The second steady-state interpretation of Q comes from the fact that the widths of both the amplitude and power curves are proportional to γ (see Eqs. (92) and (128)). Therefore, since $Q \equiv \omega_0/\gamma$, the widths of the peaks are proportional to $1/Q$. So the larger the Q value, the thinner the peaks.

  Putting these two facts together, a large Q value means that the amplitude curve is tall and thin. And likewise for the power curve.

Let's now look at some applications of these interpretations.

TUNING FORKS: The transient-solution interpretation allows for an easy calculation of Q, at least approximately. Consider a tuning fork, for example. A typical frequency is $\omega = 440\,\text{s}^{-1}$ (a concert A pitch). Let's say that it takes about 5 seconds to essentially die out (when exactly it reaches 4% of the initial amplitude is hard to tell, but we're just doing a rough calculation here). This corresponds to $5 \cdot 440 \approx 2000$ oscillations. So this is (approximately) the Q value of the tuning fork.

RADIOS: Both of the steady-state-solution interpretations (tall peak and thin peak) are highly relevant in any wireless device, such at a radio, cell phone, etc. The natural frequency of the RLC circuit in, say, a radio is "tuned" (usually by adjusting the capacitance) so that it has a certain resonant frequency; see Problem [to be added]. If this frequency corresponds to the frequency of the electromagnetic waves (see Chapter 8) that are emitted by a given radio station, then a large-amplitude oscillation will be created in the radio's circuit. This can then be amplified and sent to the speakers, creating the sound that you hear.

The taller the power peak, the stronger the signal that is obtained. If a radio station is very far away, then the electromagnetic wave has a very small amplitude by the time it gets to your radio. This means that the analog of F_d in Eq. (125) is very small. So the only chance of having a sizeable value (relative to the oscillations from the inevitable noise of other electromagnetic waves bombarding your radio) of the $F_\text{d}^2/2\gamma m$ term is to have γ be small, or equivalently Q be large. (The electrical analog of γ is the resistance of the circuit.)

However, we need *two* things to be true if we want to have a pleasant listening experience. We not only need a strong signal from the station we want to listen to, we also need a *weak* signal from every other station, otherwise we'll end up with a garbled mess. The thinness of the power curve saves the day here. If the power peak is thin enough, then a nearby radio-station frequency (even, say, $\omega = (0.99)\omega_0$) will contribute negligible power to the circuit. It's like this second station doesn't exist, which is exactly how we want things to look.

Atomic clocks: Another application where the second of the steady-state-solution interpretations is critical is atomic clocks. Atomic clocks involve oscillations between certain energy levels in atoms, but there's no need to get into the details here. Suffice it to say that there exists a damped oscillator with a particular natural frequency, and you can drive this oscillator. The basic goal in using an atomic clock is to measure with as much accuracy and precision as possible the value of the natural frequency of the atomic oscillations. You can do this by finding the driving frequency that produces the largest oscillation amplitude, or equivalently that requires the largest power input. The narrower the amplitude (or power) curve, the more confident you can be that your driving frequency ω equals the natural frequency ω_0. This is true for the following reason.

Consider a wide amplitude curve like the first one shown in Fig. 27. It's hard to tell, by looking at the size of the resulting amplitude, whether you're at, say ω_1 or ω_2, or ω_3 (all measurements have some inherent error, so you can never be sure *exactly* what amplitude you've measured). You might define the time unit of one second under the assumption that ω_1 is the natural frequency, whereas someone else (or perhaps you on a different day) might define a second by thinking that ω_3 is the natural frequency. This yields an inconsistent standard of time. Although the natural frequency of the atomic oscillations has the same value everywhere, the point is that people's opinions on what this value actually is will undoubtedly vary if the amplitude curve is wide. Just because there's a definite value out there doesn't mean that we know what it is.[7]

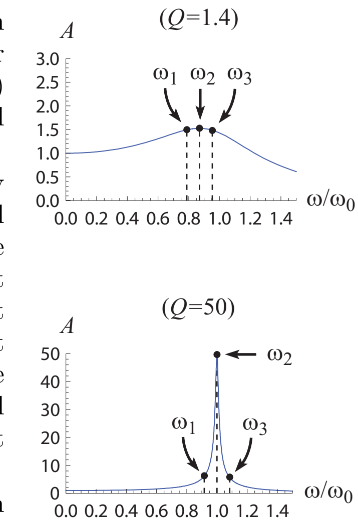


Figure 27

If, on the other hand, we have a narrow amplitude curve like the second one shown in Fig. 27, then a measurement of a large amplitude can quite easily tell you that you're somewhere around ω_1, versus ω_2 or ω_3. Basically, the uncertainty is on the order of the width of the curve, so the smaller the width, the smaller the uncertainty. Atomic clocks have very high Q values, on the order of 10^{17}. The largeness of this number implies a very small amplitude width, and hence very accurate clocks.

The tall-peak property of a large Q value isn't too important in atomic clocks. It was important in the case of a radio, because you might want to listen to a radio station that is far away. But with atomic clocks there isn't an issue with the oscillator having to pick up a weak driving signal. The driving mechanism is right next to the atoms that are housing the oscillations.

The transient property of large Q (that a large number of oscillations will occur before the amplitude dies out) also isn't too important in atomic clocks. You are continuing to drive the system, so there isn't any danger of the oscillations dying out.

1.3.4 Further discussion of resonance

Let's now talk a bit more about resonance. As we saw in Eq. (118), the $\ddot{x}$ and $\omega_0^2 x$ terms cancel for the steady-state solution, $x(t) = A\cos(\omega t + \phi)$, because $\omega = \omega_0$ at resonance. You can consider the driving force to be simply balancing the damping force, while the spring and the mass undergo their standard simple-harmonic motion. If $\omega = \omega_0$, and if $A < F_d/\gamma m\omega$ (which means that the system isn't in a steady state), then the driving force is larger than the damping force, so the motion grows. If, on the other hand, $A > F_d/\gamma m\omega$,

[7]This is the classical way of thinking about it. The (correct) quantum-mechanical description says that there actually isn't a definite natural frequency; the atoms themselves don't even know what it is. All that exists is a distribution of possible natural frequencies. But for the present purposes, it's fine to think about things classically.

then the driving force is less than the damping force, so the motion shrinks. This is why $A = F_d/\gamma m\omega$ at resonance.

As shown in Fig. 26, the force leads the motion by 90° at resonance, so the force is in phase with the velocity. This leads to the largest possible energy being fed into the system, because the work is always positive, so there is no cancelation with negative work. There are many examples of resonance in the real world, sometimes desirable, and sometimes undesirable. Let's take a look at a few.

Desirable resonance

- RLC CIRCUITS: As you will find if you do Problem [to be added], you can use Kirchhoff's rules in an RLC circuit to derive an equation exactly analogous to the damped/driven oscillator equation in Eq. (81). The quantities m, γ, k, and F_d in the mechanical system are related to the quantities L, R, $1/C$, and V_0 in the electrical system, respectively. Resonance allows you to pick out a certain frequency and ignore all the others. This is how radios, cell phones, etc. work, as we discussed in the "Q values" section above.

  If you have a radio sitting on your desk, then it is being bombarded by radio waves with all sorts of frequencies. If you want to pick out a certain frequency, then you can "tune" your radio to that frequency by changing the radio's natural frequency ω_0 (normally done by changing the capacitance C in the internal circuit). Assuming that the damping in the circuit is small (this is determined by R), then from the plot of A in Fig. 20, there will be a large oscillation in the circuit at the radio station's frequency, but a negligible oscillation at all the other frequencies that are bombarding the radio.

- MUSICAL INSTRUMENTS: The "pipe" of, say, a flute has various natural frequencies (depending on which keys are pressed), and these are the ones that survive when you blow air across the opening. We'll talk much more about musical instruments in Chapter 5. There is a subtlety about whether some musical instruments function because of resonance or because of "positive feedback," but we won't worry about that here.

- THE EAR: The hair-like nerves in the cochlea have a range of resonant frequencies which depend on the position in the cochlea. Depending on which ones vibrate, a signal is (somehow) sent to the brain telling it what the pitch is. It is quite remarkable how this works.

Undesirable resonance

- VEHICLE VIBRATIONS: This is particularly relevant in aircraft. Even the slightest driving force (in particular from the engine) can create havoc if its frequency matches up with *any* of the resonant frequencies of the plane. There is no way to theoretically predict every single one of the resonant frequencies, so the car/plane/whatever has to be tested at all frequencies by sweeping through them and looking for large amplitudes. This is difficult, because you need the final product. You can't do it with a prototype in early stages of development.

- TACOMA NARROWS BRIDGE FAILURE: There is a famous video of this bridge oscillating wildly and then breaking apart. As with some musical instruments, this technically shouldn't be called "resonance." But it's the same basic point – a natural frequency of the object was excited, in one way or another.

- Millennium Bridge in London: This pedestrian bridge happened to have a lateral resonant frequency on the order of 1 Hz. So when it started to sway (for whatever reason), people began to walk in phase with it (which is the natural thing to do). This had the effect of driving it more and further increasing the amplitude. Dampers where added, which fixed the problem.

- Tall buildings: A tall building has a resonant frequency of swaying (or actually a couple, depending on the direction; and there can be twisting, too). If the effects from the wind or earthquakes happen to unfortuitously drive the building at this frequency, then the sway can become noticeable. "Tuned mass dampers" (large masses connected to damping mechanisms, in the upper floors) help alleviate this problem.

- Space station: In early 2009, a booster engine on the space station changed its firing direction at a frequency that happened to match one of the station's resonant frequencies (about 0.5 Hz). The station began to sway back and forth, made noticeable by the fact that free objects in the air were moving back and forth. Left unchecked, a larger and larger amplitude would of course be very bad for the structure. It was fortunately stopped in time.

Chapter 2

Normal modes

David Morin, morin@physics.harvard.edu

In Chapter 1 we dealt with the oscillations of one mass. We saw that there were various possible motions, depending on what was influencing the mass (spring, damping, driving forces). In this chapter we'll look at oscillations (generally without damping or driving) involving more than one object. Roughly speaking, our counting of the number of masses will proceed as: two, then three, then infinity. The infinite case is relevant to a continuous system, because such a system contains (ignoring the atomic nature of matter) an infinite number of infinitesimally small pieces. This is therefore the chapter in which we will make the transition from the oscillations of one particle to the oscillations of a continuous object, that is, to waves.

The outline of this chapter is as follows. In Section 2.1 we solve the problem of two masses connected by springs to each other and to two walls. We will solve this in two ways – a quick way and then a longer but more fail-safe way. We encounter the important concepts of *normal modes* and *normal coordinates*. We then add on driving and damping forces and apply some results from Chapter 1. In Section 2.2 we move up a step and solve the analogous problem involving three masses. In Section 2.3 we solve the general problem involving N masses and show that the results reduce properly to the ones we already obtained in the $N = 2$ and $N = 3$ cases. In Section 2.4 we take the $N \to \infty$ limit (which corresponds to a continuous stretchable material) and derive the all-important *wave equation*. We then discuss what the possible waves can look like.

2.1 Two masses

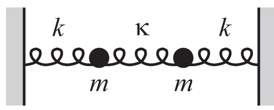


Figure 1

For a single mass on a spring, there is one natural frequency, namely $\sqrt{k/m}$. (We'll consider undamped and undriven motion for now.) Let's see what happens if we have two equal masses and three spring arranged as shown in Fig. 1. The two outside spring constants are the same, but we'll allow the middle one to be different. In general, all three spring constants could be different, but the math gets messy in that case.

Let x_1 and x_2 measure the displacements of the left and right masses from their respective equilibrium positions. We can assume that all of the springs are unstretched at equilibrium, but we don't actually have to, because the spring force is linear (see Problem [to be added]). The middle spring is stretched (or compressed) by $x_2 - x_1$, so the $F = ma$ equations on the two masses are

$$\begin{aligned} m\ddot{x}_1 &= -kx_1 - \kappa(x_1 - x_2), \\ m\ddot{x}_2 &= -kx_2 - \kappa(x_2 - x_1). \end{aligned} \tag{1}$$

Concerning the signs of the κ terms here, they are equal and opposite, as dictated by Newton's third law, so they are either both right or both wrong. They are indeed both right, as can be seen by taking the limit of, say, large x_2. The force on the left mass is then in the positive direction, which is correct.

These two $F = ma$ equations are "coupled," in the sense that both x_1 and x_2 appear in both equations. How do we go about solving for $x_1(t)$ and $x_2(t)$? There are (at least) two ways we can do this.

2.1.1 First method

This first method is quick, but it works only for simple systems with a sufficient amount of symmetry. The main goal in this method is to combine the $F = ma$ equations in well-chosen ways so that x_1 and x_2 appear only in certain unique combinations. It sometimes involves a bit of guesswork to determine what these well-chosen ways are. But in the present problem, the simplest thing to do is add the $F = ma$ equations in Eq. (1), and it turns out that this is in fact one of the two useful combinations to form. The sum yields

$$m(\ddot{x}_1 + \ddot{x}_2) = -k(x_1 + x_2) \quad \Longrightarrow \quad \frac{d^2}{dt^2}(x_1 + x_2) = -\frac{k}{m}(x_1 + x_2). \tag{2}$$

The variables x_1 and x_2 appear here only in the unique combination, $x_1 + x_2$. And furthermore, this equation is simply a harmonic-motion equation for the quantity $x_1 + x_2$. The solution is therefore

$$x_1(t) + x_2(t) = 2A_{\text{s}} \cos(\omega_{\text{s}} t + \phi_{\text{s}}), \qquad \text{where } \boxed{\omega_{\text{s}} \equiv \sqrt{\frac{k}{m}}} \tag{3}$$

The "s" here stands for "slow," to be distinguished from the "fast" frequency we'll find below. And we've defined the coefficient to be $2A_{\text{s}}$ so that we won't have a bunch of factors of $1/2$ in our final answer in Eq. (6) below.

No matter what complicated motion the masses are doing, the quantity $x_1 + x_2$ always undergoes simple harmonic motion with frequency ω_{s}. This is by no means obvious if you look at two masses bouncing back and forth in an arbitrary manner.

The other useful combination of the $F = ma$ equations is their difference, which conveniently is probably the next thing you might try. This yields

$$m(\ddot{x}_1 - \ddot{x}_2) = -(k + 2\kappa)(x_1 - x_2) \quad \Longrightarrow \quad \frac{d^2}{dt^2}(x_1 - x_2) = -\frac{k + 2\kappa}{m}(x_1 - x_2). \tag{4}$$

The variables x_1 and x_2 now appear only in the unique combination, $x_1 - x_2$. And again, we have a harmonic-motion equation for the quantity $x_1 - x_2$. So the solution is (the "f" stands for "fast")

$$x_1(t) - x_2(t) = 2A_{\text{f}} \cos(\omega_{\text{f}} t + \phi_{\text{f}}), \qquad \text{where } \boxed{\omega_{\text{f}} \equiv \sqrt{\frac{k + 2\kappa}{m}}} \tag{5}$$

As above, no matter what complicated motion the masses are doing, the quantity $x_1 - x_2$ always undergoes simple harmonic motion with frequency ω_{f}.

We can now solve for $x_1(t)$ and $x_2(t)$ by adding and subtracting Eqs. (3) and (5). The result is

$$\begin{aligned} x_1(t) &= A_s \cos(\omega_s t + \phi_s) + A_f \cos(\omega_f t + \phi_f), \\ x_2(t) &= A_s \cos(\omega_s t + \phi_s) - A_f \cos(\omega_f t + \phi_f). \end{aligned} \tag{6}$$

The four constants, A_s, A_f, ϕ_s, ϕ_f are determined by the four initial conditions, $x_1(0)$, $x_2(0)$, $\dot{x}_1(0)$, $\dot{x}_1(0)$.

The above method will clearly work only if we're able to guess the proper combinations of the $F = ma$ equations that yield equations involving unique combinations of the variables. Adding and subtracting the equations worked fine here, but for more complicated systems with unequal masses or with all the spring constants different, the appropriate combination of the equations might be far from obvious. And there is no guarantee that guessing around will get you anywhere. So before discussing the features of the solution in Eq. (6), let's take a look at the other more systematic and fail-safe method of solving for x_1 and x_2.

2.1.2 Second method

This method is longer, but it works (in theory) for any setup. Our strategy will be to look for simple kinds of motions where both masses move with the same frequency. We will then build up the most general solution from these simple motions. For all we know, such motions might not even exist, but we have nothing to lose by trying to find them. We will find that they do in fact exist. You might want to try to guess now what they are for our two-mass system, but it isn't necessary to know what they look like before undertaking this method.

Let's guess solutions of the form $x_1(t) = A_1 e^{i\omega t}$ and $x_2(t) = A_2 e^{i\omega t}$. For bookkeeping purposes, it is convenient to write these solutions in vector form:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} e^{i\omega t}. \tag{7}$$

We'll end up taking the real part in the end. We can alternatively guess the solution $e^{\alpha t}$ without the i, but then our α will come out to be imaginary. Either choice will get the job done. Plugging these guesses into the $F = ma$ equations in Eq. (1), and canceling the factor of $e^{i\omega t}$, yields

$$\begin{aligned} -m\omega^2 A_1 &= -kA_1 - \kappa(A_1 - A_2), \\ -m\omega^2 A_2 &= -kA_2 - \kappa(A_2 - A_1). \end{aligned} \tag{8}$$

In matrix form, this can be written as

$$\begin{pmatrix} -m\omega^2 + k + \kappa & -\kappa \\ -\kappa & -m\omega^2 + k + \kappa \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{9}$$

At this point, it seems like we can multiply both sides of this equation by the inverse of the matrix. This leads to $(A_1, A_2) = (0, 0)$. This is obviously a solution (the masses just sit there), but we're looking for a nontrivial solution that actually contains some motion. The only way to escape the preceding conclusion that A_1 and A_2 must both be zero is if the inverse of the matrix doesn't exist. Now, matrix inverses are somewhat messy things (involving cofactors and determinants), but for the present purposes, the only fact we need to know about them is that they involve dividing by the determinant. So if the determinant is zero, then the inverse doesn't exist. This is therefore what we want. Setting the determinant equal to zero gives the quartic equation,

$$\begin{vmatrix} -m\omega^2+k+\kappa & -\kappa \\ -\kappa & -m\omega^2+k+\kappa \end{vmatrix} = 0 \quad \Longrightarrow \quad (-m\omega^2+k+\kappa)^2-\kappa^2=0$$
$$\Longrightarrow \quad -m\omega^2+k+\kappa=\pm\kappa$$
$$\Longrightarrow \quad \omega^2=\frac{k}{m} \quad \text{or} \quad \frac{k+2\kappa}{m}. \tag{10}$$

The four solutions to the quartic equation are therefore $\omega=\pm\sqrt{k/m}$ and $\omega=\pm\sqrt{(k+2\kappa)/m}$.

For the case where $\omega^2=k/m$, we can plug this value of ω^2 back into Eq. (9) to obtain

$$\kappa\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{11}$$

Both rows of this equation yield the same result (this was the point of setting the determinant equal to zero), namely $A_1=A_2$. So (A_1,A_2) is proportional to the vector $(1,1)$.

For the case where $\omega^2=(k+2\kappa)/m$, Eq. (9) gives

$$\kappa\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{12}$$

Both rows now yield $A_1=-A_2$. So (A_1,A_2) is proportional to the vector $(1,-1)$.

With $\omega_\text{s}\equiv\sqrt{k/m}$ and $\omega_\text{f}\equiv\sqrt{(k+2\kappa)/m}$, we can write the general solution as the sum of the four solutions we have found. In vector notation, $x_1(t)$ and $x_2(t)$ are given by

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}=C_1\begin{pmatrix} 1 \\ 1 \end{pmatrix}e^{i\omega_\text{s}t}+C_2\begin{pmatrix} 1 \\ 1 \end{pmatrix}e^{-i\omega_\text{s}t}+C_3\begin{pmatrix} 1 \\ -1 \end{pmatrix}e^{i\omega_\text{f}t}+C_4\begin{pmatrix} 1 \\ -1 \end{pmatrix}e^{-i\omega_\text{f}t}. \tag{13}$$

We now perform the usual step of invoking the fact that the positions $x_1(t)$ and $x_2(t)$ must be real for all t. This yields that standard result that $C_1=C_2^*\equiv(A_\text{s}/2)e^{i\phi_\text{s}}$ and $C_3=C_4^*\equiv(A_\text{f}/2)e^{i\phi_\text{f}}$. We have included the factors of 1/2 in these definitions so that we won't have a bunch of factors of 1/2 in our final answer. The imaginary parts in Eq. (13) cancel, and we obtain

$$\boxed{\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}=A_\text{s}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\cos(\omega_\text{s}t+\phi_\text{s})+A_\text{f}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\cos(\omega_\text{f}t+\phi_\text{f})} \tag{14}$$

Therefore,

$$\begin{aligned} x_1(t) &= A_\text{s}\cos(\omega_\text{s}t+\phi_\text{s})+A_\text{f}\cos(\omega_\text{f}t+\phi_\text{f}), \\ x_2(t) &= A_\text{s}\cos(\omega_\text{s}t+\phi_\text{s})-A_\text{f}\cos(\omega_\text{f}t+\phi_\text{f}). \end{aligned} \tag{15}$$

This agrees with the result in Eq. (6).

As we discussed in Section 1.1.5, we could have just taken the real part of the $C_1(1,1)e^{i\omega_\text{s}t}$ and $C_3(1,-1)e^{i\omega_\text{f}t}$ solutions, instead of going through the "positions must be real" reasoning. However, you should continue using the latter reasoning until you're comfortable with the short cut of taking the real part.

Remark: Note that Eq. (9) can be written in the form,

$$\begin{pmatrix} k+\kappa & -\kappa \\ -\kappa & k+\kappa \end{pmatrix}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}=m\omega^2\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}. \tag{16}$$

So what we did above was solve for the *eigenvectors* and *eigenvalues* of this matrix. The eigenvectors of a matrix are the special vectors that get carried into a multiple of themselves what acted on by the matrix. And the multiple (which is $m\omega^2$ here) is called the eigenvalue. Such vectors are indeed special, because in general a vector gets both stretched (or shrunk) *and* rotated when acted on by a matrix. Eigenvectors don't get rotated at all. ♣

A third method of solving our coupled-oscillator problem is to solve for x_2 in the first equation in Eq. (1) and plug the result into the second. You will get a big mess of a fourth-order differential equation, but it's solvable by guessing $x_1 = Ae^{i\omega t}$.

2.1.3 Normal modes and normal coordinates

Normal modes

Having solved for $x_1(t)$ and $x_2(t)$ in various ways, let's now look at what we're found. If $A_\mathrm{f} = 0$ in Eq. (15), then we have

$$x_1(t) = x_2(t) = A_\mathrm{s}\cos(\omega_\mathrm{s} t + \phi_\mathrm{s}). \tag{17}$$

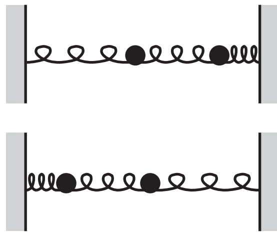


Figure 2

So both masses move in exactly the same manner. Both to the right, then both to the left, and so on. This is shown in Fig. 2. The middle spring is never stretched, so it effectively isn't there. We therefore basically have two copies of a simple spring-mass system. This is consistent with the fact that ω_s equals the standard expression $\sqrt{k/m}$, independent of κ. This nice motion, where both masses move with the same frequency, is called a *normal mode*. To specify what a normal mode looks like, you have to give the frequency and also the relative amplitudes. So this mode has frequency $\sqrt{k/m}$, and the amplitudes are equal.

If, on the other hand, $A_\mathrm{s} = 0$ in Eq. (15), then we have

$$x_1(t) = -x_2(t) = A_\mathrm{f}\cos(\omega_\mathrm{f} t + \phi_\mathrm{f}). \tag{18}$$

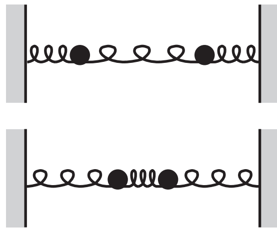


Figure 3

Now the masses move oppositely. Both outward, then both inward, and so on. This is shown in Fig. 3. The frequency is now $\omega_\mathrm{f} = \sqrt{(k + 2\kappa)/m}$. It makes sense that this is larger than ω_s, because the middle spring is now stretched or compressed, so it adds to the restoring force. This nice motion is the other normal mode. It has frequency $\sqrt{(k + 2\kappa)/m}$, and the amplitudes are equal and opposite. The task of Problem [to be added] is to deduce the frequency ω_f in a simpler way, without going through the whole process above.

Eq. (15) tells us that any arbitrary motion of the system can be thought of as a linear combination of these two normal modes. But in the general case where both coefficients A_s and A_f are nonzero, it's rather difficult to tell that the motion is actually built up from these two simple normal-mode motions.

Normal coordinates

By adding and subtracting the expressions for $x_1(t)$ and $x_1(t)$ in Eq. (15), we see that for *any* arbitrary motion of the system, the quantity $x_1 + x_2$ oscillates with frequency ω_s, and the quantity $x_1 - x_2$ oscillates with frequency ω_f. These combinations of the coordinates are known as the *normal coordinates* of the system. They are the nice combinations of the coordinates that we found advantageous to use in the first method above.

The $x_1 + x_2$ normal coordinate is associated with the normal mode $(1, 1)$, because they both have frequency ω_s. Equivalently, any contribution from the other mode (where $x_1 = -x_2$) will vanish in the sum $x_1 + x_2$. Basically, the sum $x_1 + x_2$ picks out the part of the motion with frequency ω_s and discards the part with frequency ω_f. Similarly, the $x_1 - x_2$ normal coordinate is associated with the normal mode $(1, -1)$, because they both have frequency $\omega_{\rm f}$. Equivalently, any contribution from the other mode (where $x_1 = x_2$) will vanish in the difference $x_1 - x_2$.

Note, however, that the association of the normal coordinate $x_1 + x_2$ with the normal mode $(1,1)$ does *not* follow from the fact that the coefficients in $x_1 + x_2$ are both 1. Rather, it follows from the fact that the *other* normal mode, namely $(x_1, x_2) \propto (1, -1)$, gives no contribution to the sum $x_1 + x_2$. There are a few too many 1's floating around in the present example, so it's hard to see which results are meaningful and which results are coincidence. But the following example should clear things up. Let's say we solved a problem using the determinant method, and we found the solution to be

$$\begin{pmatrix} x \\ y \end{pmatrix} = B_1 \begin{pmatrix} 3 \\ 2 \end{pmatrix} \cos(\omega_1 t + \phi_1) + B_2 \begin{pmatrix} 1 \\ -5 \end{pmatrix} \cos(\omega_2 t + \phi_2). \qquad (19)$$

Then $5x + y$ is the normal coordinate associated with the normal mode $(3, 2)$, which has frequency ω_1. This is true because there is no $\cos(\omega_2 t + \phi_2)$ dependence in the quantity $5x + y$. And similarly, $2x - 3y$ is the normal coordinate associated with the normal mode $(1, -5)$, which has frequency ω_2, because there is no $\cos(\omega_1 t + \phi_1)$ dependence in the quantity $2x - 3y$.

2.1.4 Beats

Let's now apply some initial conditions to the solution in Eq. (15). We'll take the initial conditions to be $\dot{x}_1(0) = \dot{x}_2(0) = 0$, $x_1(0) = 0$, and $x_2(0) = A$. In other words, we're pulling the right mass to the right, and then releasing both masses from rest. It's easier to apply these conditions if we write the solutions for $x_1(t)$ and $x_2(t)$ in the form,

$$\begin{aligned} x_1(t) &= a\cos\omega_{\rm s}t + b\sin\omega_{\rm s}t + c\cos\omega_{\rm f}t + d\sin\omega_{\rm f}t \\ x_2(t) &= a\cos\omega_{\rm s}t + b\sin\omega_{\rm s}t - c\cos\omega_{\rm f}t - d\sin\omega_{\rm f}t. \end{aligned} \qquad (20)$$

This form of the solution is obtained by using the trig sum formulas to expand the sines and cosines in Eq. (15). The coefficients a, b, c, d are related to the constants $A_{\rm s}$, $A_{\rm f}$, $\phi_{\rm s}$, $\phi_{\rm f}$. For example, the cosine sum formula gives $a = A_{\rm s}\cos\phi_{\rm s}$. If we now apply the initial conditions to Eq. (20), the velocities $\dot{x}_1(0) = \dot{x}_1(0) = 0$ quickly give $b = d = 0$. And the positions $x_1(0) = 0$ and $x_2(0) = A$ give $a = -c = A/2$. So we have

$$\begin{aligned} x_1(t) &= \frac{A}{2}\big(\cos\omega_{\rm s}t - \cos\omega_{\rm f}t\big), \\ x_2(t) &= \frac{A}{2}\big(\cos\omega_{\rm s}t + \cos\omega_{\rm f}t\big). \end{aligned} \qquad (21)$$

For arbitrary values of $\omega_{\rm s}$ and $\omega_{\rm f}$, this generally looks like fairly random motion, but let's look at a special case. If $\kappa \ll k$, then the $\omega_{\rm f}$ in Eq. (5) is only slightly larger than the $\omega_{\rm s}$ in Eq. (3), so something interesting happens. For frequencies that are very close to each other, it's a standard technique (for reasons that will become clear) to write $\omega_{\rm s}$ and $\omega_{\rm f}$ in terms of their average and (half) difference:

$$\begin{aligned} \omega_{\rm s} &= \frac{\omega_{\rm f} + \omega_{\rm s}}{2} - \frac{\omega_{\rm f} - \omega_{\rm s}}{2} \equiv \Omega - \epsilon, \\ \omega_{\rm f} &= \frac{\omega_{\rm f} + \omega_{\rm s}}{2} + \frac{\omega_{\rm f} - \omega_{\rm s}}{2} \equiv \Omega + \epsilon, \end{aligned} \qquad (22)$$

where

$$\Omega \equiv \frac{\omega_{\rm f} + \omega_{\rm s}}{2}, \qquad \text{and} \qquad \epsilon \equiv \frac{\omega_{\rm f} - \omega_{\rm s}}{2}. \qquad (23)$$

Using the identity $\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$, Eq. (21) becomes

$$\begin{aligned} x_1(t) &= \frac{A}{2}\Big(\cos(\Omega-\epsilon)t - \cos(\Omega+\epsilon)t\Big) = A\sin\Omega t\sin\epsilon t, \\ x_2(t) &= \frac{A}{2}\Big(\cos(\Omega-\epsilon)t + \cos(\Omega+\epsilon)t\Big) = A\cos\Omega t\cos\epsilon t. \end{aligned} \tag{24}$$

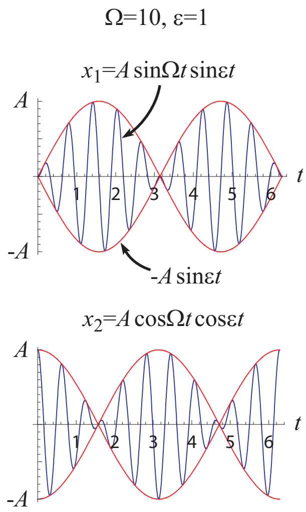


Figure 4

If ω_s is very close to ω_f, then $\epsilon \ll \Omega$, which means that the ϵt oscillation is much slower than that Ωt oscillation. The former therefore simply acts as an envelope for the latter. $x_1(t)$ and $x_2(t)$ are shown in Fig. 4 for $\Omega = 10$ and $\epsilon = 1$. The motion sloshes back and forth between the masses. At the start, only the second mass is moving. But after a time of $\epsilon t = \pi/2 \Longrightarrow t = \pi/2\epsilon$, the second mass is essentially not moving and the first mass has all the motion. Then after another time of $\pi/2\epsilon$ it switches back, and so on.

This sloshing back and forth can be understood in terms of driving forces and resonance. At the start (and until $\epsilon t = \pi/2$), x_2 looks like $\cos\Omega t$ with a slowly changing amplitude (assuming $\epsilon \ll \Omega$). And x_1 looks like $\sin\Omega t$ with a slowly changing amplitude. So x_2 is 90° *ahead* of x_1, because $\cos\Omega t = \sin(\Omega t + \pi/2)$. This 90° phase difference means that the x_2 mass basically acts like a driving force (on resonance) on the x_1 mass. Equivalently, the x_2 mass is always doing positive work on the x_1 mass, and the x_1 mass is always doing negative work on the x_2 mass. Energy is therefore transferred from x_2 to x_1.

However, right after x_2 has zero amplitude (instantaneously) at $\epsilon t = \pi/2$, the $\cos\epsilon t$ factor in x_2 *switches sign*, so x_2 now looks like $-\cos\Omega t$ (times a slowly-changing amplitude). And x_1 still looks like $\sin\Omega t$. So now x_2 is 90° *behind* x_1, because $-\cos\Omega t = \sin(\Omega t - \pi/2)$. So the x_1 mass now acts like a driving force (on resonance) on the x_2 mass. Energy is therefore transferred from x_1 back to x_2. And so on and so forth.

In the plots in Fig. 4, you can see that something goes a little haywire when the envelope curves pass through zero at $\epsilon t = \pi/2,\ \pi$, etc. The x_1 or x_2 curves skip ahead (or equivalently, fall behind) by half of a period. If you inverted the second envelope "bubble" in the first plot, the periodicity would then return. That is, the peaks of the fast-oscillation curve would occur at equal intervals, even in the transition region around $\epsilon t = \pi$.

The classic demonstration of beats consists of two identical pendulums connected by a weak spring. The gravitational restoring force mimics the "outside" springs in the above setup, so the same general results carry over (see Problem [to be added]). At the start, one pendulum moves while the other is nearly stationary. But then after a while the situation is reversed. However, if the masses of the pendulums are different, it turns out that not all of the energy is transferred. See Problem [to be added] for the details.

When people talk about the "beat frequency," they generally mean the frequency of the "bubbles" in the envelope curve. If you're listening to, say, the sound from two guitar strings that are at nearly the same frequency, then this beat frequency is the frequency of the waxing and waning that you hear. But note that this frequency is 2ϵ, and not ϵ, because two bubbles occur in each of the $\epsilon t = 2\pi$ periods of the envelope.[1]

2.1.5 Driven and damped coupled oscillators

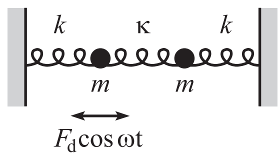


Figure 5

Consider the coupled oscillator system with two masses and three springs from Fig. 1 above, but now with a driving force acting on one of the masses, say the left one (the x_1 one); see Fig. 5. And while we're at it, let's immerse the system in a fluid, so that both masses have a drag coefficient b (we'll assume it's the same for both). Then the $F = ma$ equations are

$$m\ddot{x}_1 = -kx_1 - \kappa(x_1 - x_2) - b\dot{x}_1 + F_d\cos\omega t,$$

[1] If you want to map the spring/mass setup onto the guitar setup, then the x_1 in Eq. (21) represents the amplitude of the sound wave at your ear, and the ω_s and ω_f represent the two different nearby frequencies. The second position, x_2, doesn't come into play (or vice versa). Only one of the plots in Fig. 4 is relevant.

$$m\ddot{x}_2 = -kx_2 - \kappa(x_2 - x_1) - b\dot{x}_2. \tag{25}$$

We can solve these equations by using the same adding and subtracting technique we used in Section 2.1.1. Adding them gives

$$\begin{aligned} m(\ddot{x}_1 + \ddot{x}_2) &= -k(x_1 + x_2) - b(\dot{x}_1 + \dot{x}_2) + F_{\rm d}\cos\omega t \\ \Longrightarrow \quad \ddot{z}_{\rm s} + \gamma\dot{z}_{\rm s} + \omega_{\rm s}^2 z_{\rm s} &= F\cos\omega t, \end{aligned} \tag{26}$$

where $z_{\rm s} \equiv x_1 + x_2$, $\gamma \equiv b/m$, $\omega_{\rm s}^2 \equiv k/m$, and $F \equiv F_{\rm d}/m$. But this is our good ol' driven/dampded oscillator equation, in the variable $z_{\rm s}$. We can therefore just invoke the results from Chapter 1. The general solution is the sum of the homogeneous and particular solutions. But the let's just concentrate on the particular (steady state) solution here. We can imagine that the system has been oscillating for a long time, so that the damping has made the homogeneous solution decay to zero. For the particular solution, we can simply copy the results from Section 1.3.1. So we have

$$x_1 + x_2 \equiv z_{\rm s} = A_{\rm s}\cos(\omega t + \phi_{\rm s}), \tag{27}$$

where

$$\tan\phi_{\rm s} = \frac{-\gamma\omega}{\omega_{\rm s}^2 - \omega^2}\,, \qquad \text{and} \qquad A_{\rm s} = \frac{F}{\sqrt{(\omega_{\rm s}^2 - \omega^2)^2 + \gamma^2\omega^2}}\,. \tag{28}$$

Similarly, subtracting the $F = ma$ equations gives

$$\begin{aligned} m(\ddot{x}_1 - \ddot{x}_2) &= -(k + 2\kappa)(x_1 - x_2) - b(\dot{x}_1 - \dot{x}_2) + F_{\rm d}\cos\omega t \\ \Longrightarrow \quad \ddot{z}_{\rm f} + \gamma\dot{z}_{\rm f} + \omega_{\rm f}^2 z_{\rm f} &= F\cos\omega t, \end{aligned} \tag{29}$$

where $z_{\rm f} \equiv x_1 - x_2$ and $\omega_{\rm f}^2 \equiv (k + 2\kappa)/m$. Again, this is a nice driven/dampded oscillator equation, and the particular solution is

$$x_1 - x_2 \equiv z_{\rm f} = A_{\rm f}\cos(\omega t + \phi_{\rm f}), \tag{30}$$

where

$$\tan\phi_{\rm f} = \frac{-\gamma\omega}{\omega_{\rm f}^2 - \omega^2}\,, \qquad \text{and} \qquad A_{\rm f} = \frac{F}{\sqrt{(\omega_{\rm f}^2 - \omega^2)^2 + \gamma^2\omega^2}}\,. \tag{31}$$

Adding and subtracting Eqs. (27) and (30) to solve for $x_1(t)$ and $x_2(t)$ gives

$$\begin{aligned} x_1(t) &= C_{\rm s}\cos(\omega t + \phi_{\rm s}) + C_{\rm f}\cos(\omega t + \phi_{\rm f}), \\ x_2(t) &= C_{\rm s}\cos(\omega t + \phi_{\rm s}) - C_{\rm f}\cos(\omega t + \phi_{\rm f}), \end{aligned} \tag{32}$$

where $C_{\rm s} \equiv A_{\rm s}/2$, and $C_{\rm f} \equiv A_{\rm f}/2$.

We end up getting *two* resonant frequencies, which are simply the frequencies of the normal modes, $\omega_{\rm s}$ and $\omega_{\rm f}$. If γ is small, and if the driving frequency ω equals either $\omega_{\rm s}$ or $\omega_{\rm f}$, then the amplitudes of x_1 and x_2 are large. In the $\omega = \omega_{\rm s}$ case, x_1 and x_2 are approximately in phase with equal amplitudes (the $C_{\rm s}$ terms dominate the $C_{\rm f}$ terms). And in the $\omega = \omega_{\rm s}$ case, x_1 and x_2 are approximately out of phase with equal amplitudes (the $C_{\rm f}$ terms dominate the $C_{\rm s}$ terms, and there is a relative minus sign). But these are the normal modes we found in Section 2.1.3. The resonances therefore cause the system to be in the normal modes.

In general, if there are N masses (and hence N modes), then there are N resonant frequencies, which are the N normal-mode frequencies. So for complicated objects with more than two pieces, there are *lots* of resonances.

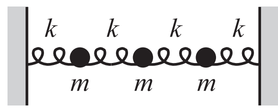


Figure 6

2.2 Three masses

As a warmup to the general case of N masses connected by springs, let's look at the case of three masses, as shown in Fig. 6. We'll just deal with undriven and undamped motion here, and we'll also assume that all the spring constants are equal, lest the math get intractable. If x_1, x_2, and x_3 are the displacements of the three masses from their equilibrium positions, then the three $F = ma$ equations are

$$\begin{aligned} m\ddot{x}_1 &= -kx_1 - k(x_1 - x_2), \\ m\ddot{x}_2 &= -k(x_2 - x_1) - k(x_2 - x_3), \\ m\ddot{x}_3 &= -k(x_3 - x_2) - kx_3. \end{aligned} \tag{33}$$

You can check that all the signs of the $k(x_i - x_j)$ terms are correct, by imagining that, say, one of the x's is very large. It isn't so obvious which combinations of these equations yield equations involving only certain unique combinations of the x's (the normal coordinates), so we won't be able to use the method of Section 2.1.1. We will therefore use the determinant method from Section 2.1.2 and guess a solution of the form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} e^{i\omega t}, \tag{34}$$

with the goal of solving for ω, and also for the amplitudes A_1, A_2, and A_3 (up to an overall factor). Plugging this guess into Eq. (33) and putting all the terms on the lefthand side, and canceling the $e^{i\omega t}$ factor, gives

$$\begin{pmatrix} -\omega^2 + 2\omega_0^2 & -\omega_0^2 & 0 \\ -\omega_0^2 & -\omega^2 + 2\omega_0^2 & -\omega_0^2 \\ 0 & -\omega_0^2 & -\omega^2 + 2\omega_0^2 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \tag{35}$$

where $\omega_0^2 \equiv k/m$. As in the earlier two-mass case, a nonzero solution for (A_1, A_2, A_3) exists only if the determinant of this matrix is zero. Setting it equal to zero gives

$$\begin{aligned} (-\omega^2 + 2\omega_0^2)\Big((-\omega^2 + 2\omega_0^2)^2 - \omega_0^4\Big) + \omega_0^2\Big(-\omega_0^2(-\omega^2 + 2\omega_0^2)\Big) &= 0 \\ \Longrightarrow \quad (-\omega^2 + 2\omega_0^2)(\omega^4 - 4\omega_0^2\omega^2 + 2\omega_0^4) &= 0. \end{aligned} \tag{36}$$

Although this is technically a 6th-order equation, it's really just a cubic equation in ω^2. But since we know that $(-\omega^2 + 2\omega_0^2)$ is a factor, in the end it boils down to a quadratic equation in ω^2.

Remark: If you had multiplied everything out and lost the information that $(-\omega^2 + 2\omega_0^2)$ is a factor, you could still easily see that $\omega^2 = 2\omega_0^2$ must be a root, because an easy-to-see normal mode is one where the middle mass stays fixed and the outer masses move in opposite directions. In this case the middle mass is essentially a brick wall, so the outer masses are connected to two springs whose other ends are fixed. The effective spring constant is then $2k$, which means that the frequency is $\sqrt{2}\omega_0$. ♣

Using the quadratic formula, the roots to Eq. (36) are

$$\omega^2 = 2\omega_0^2, \qquad \text{and} \qquad \omega^2 = (2 \pm \sqrt{2})\omega_0. \tag{37}$$

Plugging these values back into Eq. (35) to find the relations among A_1, A_2, and A_3 gives the three normal modes:[2]

$$
\begin{aligned}
\omega = \pm\sqrt{2}\,\omega_0 \quad &\Longrightarrow \quad \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \propto \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \\
\omega = \pm\sqrt{2+\sqrt{2}}\,\omega_0 \quad &\Longrightarrow \quad \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \propto \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}, \\
\omega = \pm\sqrt{2-\sqrt{2}}\,\omega_0 \quad &\Longrightarrow \quad \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \propto \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}.
\end{aligned} \tag{38}
$$

The most general solution is obtained by taking an arbitrary linear combination of the six solutions corresponding to the six possible values of ω (don't forget the three negative solutions):

$$
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = C_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^{i\sqrt{2}\omega_0 t} + C_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^{-i\sqrt{2}\omega_0 t} + \cdots. \tag{39}
$$

However, the x's must be real, so C_2 must be the complex conjugate of C_1. Likewise for the two C's corresponding to the $(1, -\sqrt{2}, 1)$ mode, and also for the two C's corresponding to the $(1, \sqrt{2}, 1)$ mode. Following the procedure that transformed Eq. (13) into Eq. (14), we see that the most general solution can be written as

$$
\begin{aligned}
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= A_{\mathrm{m}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \cos\left(\sqrt{2}\,\omega_0 t + \phi_{\mathrm{m}}\right) \\
&\quad + A_{\mathrm{f}} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix} \cos\left(\sqrt{2+\sqrt{2}}\,\omega_0 t + \phi_{\mathrm{f}}\right) \\
&\quad + A_{\mathrm{s}} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} \cos\left(\sqrt{2-\sqrt{2}}\,\omega_0 t + \phi_{\mathrm{s}}\right).
\end{aligned} \tag{40}
$$

The subscripts "m," "f," and "s" stand for middle, fast, and slow. The six unknowns, A_{m}, A_{f}, A_{s}, ϕ_{m}, ϕ_{f}, and ϕ_{s} are determined by the six initial conditions (three positions and three velocities). If A_{m} is the only nonzero coefficient, then the motion is purely in the middle mode. Likewise for the cases where only A_{f} or only A_{s} is nonzero. Snapshots of these modes are shown in Fig. 7. You should convince yourself that they qualitatively make sense. If you want to get quantitative, the task of Problem [to be added] is to give a force argument that explains the presence of the $\sqrt{2}$ in the amplitudes of the fast and slow modes.

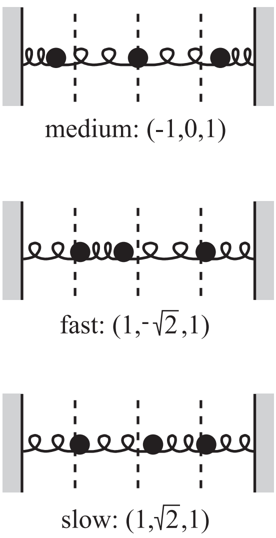


Figure 7

2.3 N masses

2.3.1 Derivation of the general result

Let's now consider the general case of N masses between two fixed walls. The masses are all equal to m, and the spring constants are all equal to k. The method we'll use below will

[2] Only two of the equations in Eq. (35) are needed. The third equation is redundant; that was the point of setting the determinant equal to zero.

actually work even if we don't have walls at the ends, that is, even if the masses extend infinitely in both directions. Let the displacements of the masses relative to their equilibrium positions be x_1, x_2,..., x_N. If the displacements of the walls are called x_0 and x_{N+1}, then the boundary conditions that we'll eventually apply are $x_0 = x_{N+1} = 0$.

The force on the nth mass is

$$F_n = -k(x_n - x_{n-1}) - k(x_n - x_{n+1}) = kx_{n-1} - 2kx_n + kx_{n+1}. \tag{41}$$

So we end up with a collection of $F = ma$ equations that look like

$$m\ddot{x}_n = kx_{n-1} - 2kx_n + kx_{n+1}. \tag{42}$$

These can all be collected into the matrix equation,

$$m\frac{d^2}{dt^2}\begin{pmatrix} \vdots \\ x_{n-1} \\ x_n \\ x_{n+1} \\ \vdots \end{pmatrix} = \begin{pmatrix} & & & \vdots & & & \\ \cdots & k & -2k & k & & & \\ & & k & -2k & k & & \\ & & & k & -2k & k & \cdots \\ & & & & \vdots & & \end{pmatrix}\begin{pmatrix} \vdots \\ x_{n-1} \\ x_n \\ x_{n+1} \\ \vdots \end{pmatrix}. \tag{43}$$

In principle, we could solve for the normal modes by guessing a solution of the form,

$$\begin{pmatrix} \vdots \\ x_{n-1} \\ x_n \\ x_{n+1} \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ A_{n-1} \\ A_n \\ A_{n+1} \\ \vdots \end{pmatrix} e^{i\omega t}, \tag{44}$$

and then setting the resulting determinant equal to zero. This is what we did in the $N = 2$ and $N = 3$ cases above. However, for large N, it would be completely intractable to solve for the ω's by using the determinant method. So we'll solve it in a different way, as follows.

We'll stick with the guess in Eq. (44), but instead of the determinant method, we'll look at each of the $F = ma$ equations individually. Consider the nth equation. Plugging $x_n(t) = A_n e^{i\omega t}$ into Eq. (42) and canceling the factor of $e^{i\omega t}$ gives

$$\begin{aligned} -\omega^2 A_n &= \omega_0^2\big(A_{n-1} - 2A_n + A_{n+1}\big) \\ \Longrightarrow \quad \frac{A_{n-1} + A_{n+1}}{A_n} &= \frac{2\omega_0^2 - \omega^2}{\omega_0^2}, \end{aligned} \tag{45}$$

where $\omega_0 = \sqrt{k/m}$, as usual. This equation must hold for all values of n from 1 to N, so we have N equations of this form. For a given mode with a given frequency ω, the quantity $(2\omega_0^2 - \omega^2)/\omega_0^2$ on the righthand side is a constant, independent of n. So the ratio $(A_{n-1}+A_{n+1})/A_n$ on the lefthand side must also be independent of n. The problem therefore reduces to finding the general form of a string of A's that has the ratio $(A_{n-1} + A_{n+1})/A_n$ being independent of n.

If someone gives you three adjacent A's, then this ratio is determined, so you can recursively find the A's for all other n (both larger and smaller than the three you were given). Or equivalently, if someone gives you two adjacent A's and also ω, so that the value of $(2\omega_0^2 - \omega^2)/\omega_0^2$ is known (we're assuming that ω_0 is given), then all the other A's can be determined. The following claim tells us what form the A's take. It is this claim that allows us to avoid using the determinant method.

Claim 2.1 *If $\omega \le 2\omega_0$, then any set of A_n's satisfying the system of N equations in Eq. (45) can be written as*

$$A_n = B\cos n\theta + C\sin n\theta, \tag{46}$$

for certain values of B, C, and θ. (The fact that there are three parameters here is consistent with the fact that three A's, or two A's and ω, determine the whole set.)

Proof: We'll start by defining

$$\cos\theta \equiv \frac{A_{n-1} + A_{n+1}}{2A_n}. \tag{47}$$

As mentioned above, the righthand side is independent of n, so θ is well defined (up to the usual ambiguities of duplicate angles; $\theta + 2\pi$, and $-\theta$, etc. also work).[3] If we're looking at a given normal mode with frequency ω, then in view of Eq. (45), an equivalent definition of θ is

$$2\cos\theta \equiv \frac{2\omega_0^2 - \omega^2}{\omega_0^2}. \tag{48}$$

These definitions are permitted only if they yield a value of $\cos\theta$ that satisfies $|\cos\theta| \le 1$. This condition is equivalent to the condition that ω must satisfy $-2\omega_0 \le \omega \le 2\omega_0$. We'll just deal with positive ω here (negative ω yields the same results, because only its square enters into the problem), but we must remember to also include the $e^{-i\omega t}$ solution in the end (as usual). So this is where the $\omega \le 2\omega_0$ condition in the claim comes from.[4]

We will find that with walls at the ends, θ (and hence ω) can take on only a certain set of discrete values. We will calculate these below. If there are no walls, that is, if the system extends infinitely in both directions, then θ (and hence ω) can take on a continuous set of values.

As we mentioned above, the N equations represented in Eq. (45) tell us that if we know two of the A's, and if we also have a value of ω, then we can use the equations to successively determine all the other A's. Let's say that we know what A_0 and A_1 are. (In the case where there are walls, we know know that $A_0 = 0$, but let's be general and not invoke this constraint yet.) The rest of the A_n's can be determined as follows. Define B by

$$A_0 \equiv B\cos(0\cdot\theta) + C\sin(0\cdot\theta) \quad\Longrightarrow\quad A_0 \equiv B. \tag{49}$$

(So $B = 0$ if there are walls.) Once B has been defined, define C by

$$A_1 \equiv B\cos(1\cdot\theta) + C\sin(1\cdot\theta) \quad\Longrightarrow\quad A_1 \equiv B\cos\theta + C\sin\theta, \tag{50}$$

For any A_0 and A_1, these two equations uniquely determine B and C (θ was already determined by ω). So to sum up the definitions: ω, A_0, and A_1 uniquely determine θ, B and C. (We'll deal with the multiplicity of the possible θ values below in the "Nyquist" subsection.) By construction of these definitions, the proposed $A_n = B\cos n\theta + C\sin n\theta$ relation holds for $n = 0$ and $n = 1$. We will now show inductively that it holds for all n.

[3]The motivation for this definition is that the fraction on the righthand side has a sort of second-derivative feel to it. The more this fraction differs from 1, the more curvature there is in the plot of the A_n's. (If the fraction equals 1, then each A_n is the average of its two neighbors, so we just have a straight line.) And since it's a good bet that we're going to get some sort of sinusoidal result out of all this, it's not an outrageous thing to define this fraction to be a sinusoidal function of a new quantity θ. But in the end, it does come a bit out of the blue. That's the way it is sometimes. However, you will find it less mysterious after reading Section 2.4, where we actually end up with a true second derivative, along with sinusoidal functions of x (the analog of n here).

[4]If $\omega > 2\omega_0$, then we have a so-called *evanescent* wave. We'll discuss these in Chapter 6. The $\omega = 0$ and $\omega = 2\omega_0$ cases are somewhat special; see Problem [to be added].

If we solve for A_{n+1} in Eq. (47) and use the inductive hypothesis that the $A_n = B\cos n\theta + C\sin n\theta$ result holds for $n-1$ and n, we have

$$\begin{aligned}
A_{n+1} &= (2\cos\theta)A_n - A_{n-1} \\
&= 2\cos\theta\Big(B\cos n\theta + C\sin n\theta\Big) - \Big(B\cos(n-1)\theta + C\sin(n-1)\theta\Big) \\
&= B\Big(2\cos n\theta\cos\theta - (\cos n\theta\cos\theta + \sin n\theta\sin\theta)\Big) \\
&\qquad + C\Big(2\sin n\theta\cos\theta - (\sin n\theta\cos\theta - \cos n\theta\sin\theta)\Big) \\
&= B\Big(\cos n\theta\cos\theta - \sin n\theta\sin\theta)\Big) + C\Big(\sin n\theta\cos\theta + \cos n\theta\sin\theta)\Big) \\
&= B\cos(n+1)\theta + C\sin(n+1)\theta, \qquad (51)
\end{aligned}$$

which is the desired expression for the case of $n+1$. (Note that this works independently for the B and C terms.) Therefore, since the $A_n = B\cos n\theta + C\sin n\theta$ result holds for $n=0$ and $n=1$, and since the inductive step is valid, the result therefore holds for all n.

If you wanted, you could have instead solved for A_{n-1} in Eq. (51) and demonstrated that the inductive step works in the negative direction too. Therefore, starting with two arbitrary masses anywhere in the line, the $A_n = B\cos n\theta + C\sin n\theta$ result holds even for an infinite number of masses extending in both directions. ■

This claim tells us that we have found a solution of the form,

$$x_n(t) = A_n e^{i\omega t} = (B\cos n\theta + C\sin n\theta)e^{i\omega t}. \qquad (52)$$

However, with the convention that ω is positive, we must remember that an $e^{-i\omega t}$ solution works just as well. So another solution is

$$x_n(t) = A_n e^{-i\omega t} = (D\cos n\theta + E\sin n\theta)e^{-i\omega t}. \qquad (53)$$

Note that the coefficients in this solution need not be the same as those in the $e^{i\omega t}$ solution. Since the $F = ma$ equations in Eq. (42) are all linear, the sum of two solutions is again a solution. So the most general solution (for a given value of ω) is the sum of the above two solutions (each of which is itself a linear combination of two solutions).

As usual, we now invoke the fact that the positions must be real. This implies that the above two solutions must be complex conjugates of each other. And since this must be true for all values of n, we see that B and D must be complex conjugates, and likewise for C and E. Let's define $B = D^* \equiv (F/2)e^{i\beta}$ and $C = E^* \equiv (G/2)e^{i\gamma}$. There is no reason why B, C, D, and E (or equivalently the A's in Eq. (44)) have to be real. The sum of the two solutions then becomes

$$x_n(t) = F\cos n\theta\cos(\omega t + \beta) + G\sin n\theta\cos(\omega t + \gamma) \qquad (54)$$

As usual, we could have just taken the real part of either of the solutions to obtain this (up to a factor of 2, which we can absorb into the definition of the constants). We can make it look a little more symmetrical be using the trig sum formula for the cosines. This gives the result (we're running out of letters, so we'll use C_i's for the coefficients here),

$$\boxed{x_n(t) = C_1\cos n\theta\cos\omega t + C_2\cos n\theta\sin\omega t + C_3\sin n\theta\cos\omega t + C_4\sin n\theta\sin\omega t} \qquad (55)$$

where θ is determined by ω via Eq. (48), which we can write in the form,

$$\boxed{\theta \equiv \cos^{-1}\left(\frac{2\omega_0^2 - \omega^2}{2\omega_0^2}\right)} \qquad (56)$$

The constants C_1, C_2, C_3, C_4 in Eq. (55) are related to the constants F, G, β, γ in Eq. (54) in the usual way ($C_1 = F\cos\beta$, etc.). There are yet other ways to write the solution, but we'll save the discussion of these for Section 2.4.

Eq. (55) is the most general form of the positions for the mode that has frequency ω. This set of the $x_n(t)$ functions (N of them) satisfies the $F = ma$ equations in Eq. (42) (N of them) for any values of C_1, C_2, C_3, C_4. These four constants are determined by four initial values, for example, $x_0(0)$, $\dot{x}_0(0)$, $x_1(0)$, and $\dot{x}_1(0)$. Of course, if $n = 0$ corresponds to a fixed wall, then the first two of these are zero.

Remarks:

1. Interestingly, we have found that $x_n(t)$ varies sinusoidally with position (that is, with n), as well as with time. However, whereas time takes on a continuous set of values, the position is relevant only at the discrete locations of the masses. For example, if the equilibrium positions are at the locations $z = na$, where a is the equilibrium spacing between the masses, then we can rewrite $x_n(t)$ in terms of z instead of n, using $n = z/a$. Assuming for simplicity that we have only, say, the $C_1 \cos n\theta \cos\omega t$ part of the solution, we have
$$x_n(t) \implies x_z(t) = C_1 \cos(z\theta/a)\cos\omega t. \tag{57}$$
For a given values of θ (which is related to ω) and a, this is a sinusoidal function of z (as well as of t). But we must remember that it is defined only at the discrete values of z of the form, $z = na$. We'll draw some nice pictures below to demonstrate the sinusoidal behavior, when we discuss a few specific values of N.

2. We should stress the distinction between z (or equivalently n) and x. z represents the equilibrium positions of the masses. A given mass is associated with a unique value of z. z doesn't change as the mass moves. $x_z(t)$, on the other hand, measures the position of a mass (the one whose equilibrium position is z) relative to its equilibrium position (namely z). So the total position of a given mass is $z + x$. The function $x_z(t)$ has dependence on both z and t, so we could very well write it as a function of two variables, $x(z,t)$. We will in fact adopt this notation in Section 2.4 when we talk about continuous systems. But in the present case where z can take on only discrete values, we'll stick with the $x_z(t)$ notation. But either notation is fine.

3. Eq. (55) gives the most general solution *for a given value of* ω, that is, for a given mode. While the most general motion of the masses is certainly not determined by $x_0(0)$, $\dot{x}_0(0)$, $x_1(0)$, and $\dot{x}_1(0)$, the motion for a single mode is. Let's see why this is true. If we apply the $x_0(0)$ and $x_1(0)$ boundary conditions to Eq. (55), we obtain $x_0(0) = C_1$ and $x_1(0) = C_1\cos\theta + C_3\sin\theta$. Since we are assuming that ω (and hence θ) is given, these two equations determine C_1 and C_3. But C_1 and C_3 in turn determine all the other $x_n(0)$ values via Eq. (55), because the $\sin\omega t$ terms are all zero at $t = 0$. So for a given mode, $x_0(0)$ and $x_1(0)$ determine all the other initial positions. In a similar manner, the $\dot{x}_0(0)$ and $\dot{x}_1(0)$ values determine C_2 and C_4, which in turn determine all the other initial velocities. Therefore, since the four values $x_0(0)$, $\dot{x}_0(0)$, $x_1(0)$, and $\dot{x}_1(0)$ give us all the initial positions and velocities, and since the accelerations depend on the positions (from the $F = ma$ equations in Eq. (42)), the future motion of all the masses is determined. ♣

2.3.2 Wall boundary conditions

Let us now see what happens when we invoke the boundary conditions due to fixed walls at the two ends. The boundary conditions at the walls are $x_0(t) = x_{N+1}(t) = 0$, for all t. These conditions are most easily applied to $x_n(t)$ written in the form in Eq. (54), although the form in Eq. (55) will work fine too. At the left wall, the $x_0(t) = 0$ condition gives
$$\begin{aligned} 0 = x_0(t) &= F\cos(0)\cos(\omega t + \beta) + G\sin(0)\cos(\omega t + \gamma) \\ &= F\cos(\omega t + \beta). \end{aligned} \tag{58}$$

If this is to be true for all t, we must have $F = 0$. So we're left with just the $G \sin n\theta \cos(\omega t + \gamma)$ term in Eq. (54). Applying the $x_{N+1}(t) = 0$ condition to this then gives

$$0 = x_{N+1}(t) = G \sin(N+1)\theta \, \cos(\omega t + \gamma). \tag{59}$$

One way for this to be true for all t is to have $G = 0$. But then all the x's are identically zero, which means that we have no motion at all. The other (nontrivial) way for this to be true is to have the $\sin(N+1)\theta$ factor be zero. This occurs when

$$(N+1)\theta = m\pi \implies \theta = \frac{m\pi}{N+1}, \tag{60}$$

where m is an integer. The solution for $x_n(t)$ is therefore

$$\boxed{x_n(t) = G \sin\left(\frac{nm\pi}{N+1}\right) \cos(\omega t + \gamma)} \tag{61}$$

The amplitudes of the masses are then

$$\boxed{A_n = G \sin\left(\frac{nm\pi}{N+1}\right)} \tag{62}$$

We've made a slight change in notation here. The A_n that we're now using for the amplitude is the *magnitude* of the A_n that we used in Eq. (44). That A_n was equal to $B \cos n\theta + C \sin n\theta$, which itself is some complex number which can be written in the form, $|A_n|e^{i\alpha}$. The solution for $x_n(t)$ is obtained by taking the real part of Eq. (52), which yields $x_n(t) = |A_n| \cos(\omega t + \alpha)$. So we're now using A_n to stand for $|A_n|$, lest we get tired if writing the absolute value bars over and over.[5] And α happens to equal the γ in Eq. (61).

If we invert the definition of θ in Eq. (48) to solve for ω in terms of θ, we find that the frequency is given by

$$\begin{aligned} 2\cos\theta \equiv \frac{2\omega_0^2 - \omega^2}{\omega_0^2} \quad \implies \quad \omega^2 &= 2\omega_0^2(1 - \cos\theta) \\ &= 4\omega_0^2 \sin^2(\theta/2) \\ \implies \quad \omega &= \boxed{2\omega_0 \sin\left(\frac{m\pi}{2(N+1)}\right)} \end{aligned} \tag{63}$$

We've taken the positive square root, because the convention is that ω is positive. We'll see below that m labels the normal mode (so the "m" stands for "mode"). If $m = 0$ or $m = N+1$, then Eq. (61) says that all the x_n's are identically zero, which means that we don't have any motion at all. So only m values in the range $1 \le m \le N$ are relevant. We'll see below in the "Nyquist" subsection that higher values of m simply give repetitions of the modes generated by the $1 \le m \le N$ values.

The most important point in the above results is that if the boundary conditions are walls at both ends, then the θ in Eq. (60), and hence the ω in Eq. (63), can take on only a certain set of *discrete* values. This is consistent with our results for the $N = 2$ and $N = 3$ cases in Sections 2.1 and 2.2, where we found that there were only two or three (respectively) allowed values of ω, that is, only two or three normal modes. Let's now show that for $N = 2$ and $N = 3$, the preceding equations quantitatively reproduce the results from Sections 2.1 and 2.2. You can examine higher values of N in Problem [to be added].

[5] If instead of taking the real part, you did the nearly equivalent thing of adding on the complex conjugate solution in Eq. (53), then $2|A_n|$ would be the amplitude. In this case, the A_n in Eq. (62) stands for $2|A_n|$.

The $N = 2$ case

If $N = 2$, there are two possible values of m:

- $m = 1$: Eqs. (62) and (63) give

$$A_n \propto \sin\left(\frac{n\pi}{3}\right), \qquad \text{and} \qquad \omega = 2\omega_0 \sin\left(\frac{\pi}{6}\right). \tag{64}$$

  So this mode is given by

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \propto \begin{pmatrix} \sin(\pi/3) \\ \sin(2\pi/3) \end{pmatrix} \propto \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \text{and} \qquad \omega = \omega_0. \tag{65}$$

  These agree with the first mode we found in Section 2.1.2. The frequency is ω_0, and the masses move in phase with each other.

- $m = 2$: Eqs. (62) and (63) give

$$A_n \propto \sin\left(\frac{2n\pi}{3}\right), \qquad \text{and} \qquad \omega = 2\omega_0 \sin\left(\frac{\pi}{3}\right). \tag{66}$$

  So this mode is given by

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \propto \begin{pmatrix} \sin(2\pi/3) \\ \sin(4\pi/3) \end{pmatrix} \propto \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad \text{and} \qquad \omega = \sqrt{3}\,\omega_0. \tag{67}$$

  These agree with the second mode we found in Section 2.1.2. The frequency is $\sqrt{3}\,\omega_0$, and the masses move exactly out of phase with each other.

To recap, the various parameters are: N (the number of masses), m (the mode number), and n (the label of each of the N masses). n runs from 1 to N, of course. And m effectively also runs from 1 to N (there are N possible modes for N masses). We say "effectively" because as we mentioned above, although m can technically take on any integer value, the values that lie outside the $1 \leq m \leq N$ range give duplications of the modes inside this range. See the "Nyquist" subsection below.

In applying Eqs. (62) and (63), things can get a little confusing because of all the parameters floating around. And this is just the simple case of $N = 2$. Fortunately, there is an extremely useful graphical way to see what's going on. This is one situation where a picture is indeed worth a thousand words (or equations).

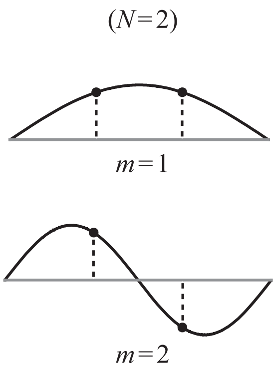
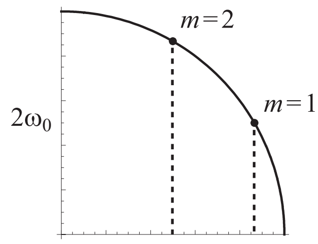


Figure 8

If we write the argument of the sin in Eq. (62) as $m\pi \cdot n/(N + 1)$, then we see that for a given N, the relative amplitudes of the masses in the mth mode are obtained by drawing a sin curve with m half oscillations, and then finding the value of this curve at equal "$1/(N + 1)$" intervals along the horizontal axis. Fig. 8 shows the results for $N = 2$. We've drawn either $m = 1$ or $m = 2$ half oscillations, and we've divided each horizontal axis into $N + 1 = 3$ equal intervals. These curves look a lot like snapshots of beads on a string oscillating *transversely* back and forth. And indeed, we will find in Chapter 4 that the $F = ma$ equations for transverse motion of beads on a string are *exactly* the same as the equations in Eq. (42) for the longitudinal motion of the spring/mass system. But for now, all of the displacements indicated in these pictures are in the longitudinal direction. And the displacements have meaning only at the discrete locations of the masses. There isn't anything actually happening at the rest of the points on the curve.

We can also easily visualize what the frequencies are. If we write the argument of the sin in Eq. (63) as $\pi/2 \cdot m/(N+1)$ then we see that for a given N, the frequency of the mth mode is obtained by breaking a quarter circle (with radius $2\omega_0$) into "$1/(N+1)$" equal intervals, and then finding the y values of the resulting points. Fig. 9 shows the results for $N = 2$. We've divided the quarter circle into $N + 1 = 3$ equal angles of $\pi/6$, which results in points at the angles of $\pi/6$ and $\pi/3$. It is *much* easier to see what's going on by looking at the pictures in Figs. 8 and 9 than by working with the algebraic expressions in Eqs. (62) and (63).



(N=2)

Figure 9

The $N = 3$ case

If $N = 3$, there are three possible values of m:

- $m = 1$: Eqs. (62) and (63) give

$$A_n \propto \sin\left(\frac{n\pi}{4}\right), \qquad \text{and} \qquad \omega = 2\omega_0 \sin\left(\frac{\pi}{8}\right). \tag{68}$$

  So this mode is given by

$$\begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \propto \begin{pmatrix} \sin(\pi/4) \\ \sin(2\pi/4) \\ \sin(3\pi/4) \end{pmatrix} \propto \begin{pmatrix} 1 \\ \sqrt{2} \\ -1 \end{pmatrix}, \qquad \text{and} \qquad \omega = \sqrt{2-\sqrt{2}}\,\omega_0, \tag{69}$$

  where we have used the half-angle formula for $\sin(\pi/8)$ to obtain ω. (Or equivalently, we just used the first line in Eq. (63).) These results agree with the "slow" mode we found in Section 2.2.

- $m = 2$: Eqs. (62) and (63) give

$$A_n \propto \sin\left(\frac{2n\pi}{4}\right), \qquad \text{and} \qquad \omega = 2\omega_0 \sin\left(\frac{2\pi}{8}\right). \tag{70}$$

  So this mode is given by

$$\begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \propto \begin{pmatrix} \sin(2\pi/4) \\ \sin(4\pi/4) \\ \sin(6\pi/4) \end{pmatrix} \propto \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \qquad \text{and} \qquad \omega = \sqrt{2}\,\omega_0. \tag{71}$$

  These agree with the "medium" mode we found in Section 2.2.

(N=3)

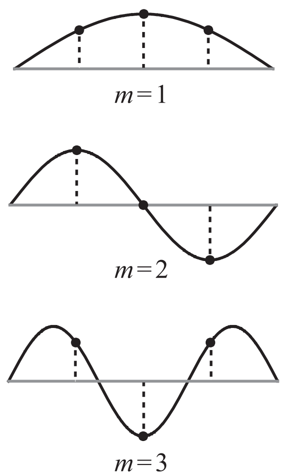


Figure 10

- $m = 3$: Eqs. (62) and (63) give

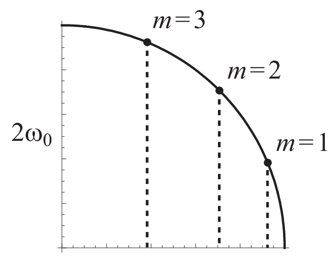
$$A_n \propto \sin\left(\frac{3n\pi}{4}\right), \qquad \text{and} \qquad \omega = 2\omega_0 \sin\left(\frac{3\pi}{8}\right). \tag{72}$$

  So this mode is given by

$$\begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \propto \begin{pmatrix} \sin(3\pi/4) \\ \sin(6\pi/4) \\ \sin(9\pi/4) \end{pmatrix} \propto \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}, \qquad \text{and} \qquad \omega = \sqrt{2+\sqrt{2}}\,\omega_0. \tag{73}$$

  These agree with the "fast" mode we found in Section 2.2..

As with the $N = 2$ case, it's much easier to see what's going on if we draw some pictures. Fig. 10 shows the relative amplitudes within the three modes, and Fig. 11 shows the associated



(N=3)

Figure 11

frequencies. Each horizontal axis in Fig. 10 is broken up into $N+1=4$ equal segments, and the quarter circle in Fig. 11 is broken up into $N+1=4$ equal arcs.

As mentioned above, although Fig. 10 looks like transverse motion on a string, remember that all the displacements indicated in this figure are in the longitudinal direction. For example, in the first $m=1$ mode, all three masses move in the same direction, but the middle one moves farther (by a factor of $\sqrt{2}$) than the outer ones. Problem [to be added] discusses higher values of N.

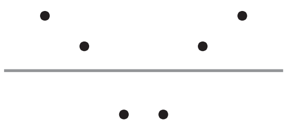


Figure 12

Aliasing, Nyquist frequency

Consider the $N=6$ case shown in Fig. 12. Assuming that this corresponds to a normal mode, which one is it? If you start with $m=1$ and keep trying different values of m until the (appropriately scaled) sin curve in Eq. (62) matches up with the six points in Fig. 12 (corresponding to n values from 1 to N), you'll find that $m=3$ works, as shown in Fig. 13.

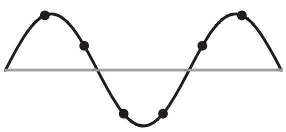


Figure 13

However, the question then arises as to whether this is the only value of m that allows the sin curve to fit the given points. If you try some higher values of m, and if you're persistent, then you'll find that $m=11$ also works, as long as you throw in a minus sign in front of the sin curve (this corresponds to the G coefficient in Eq. (61) being negative, which is fine). This is shown in Fig. 14. And if you keep going with higher m values, you'll find that $m=17$ works too, and this time there's no need for a minus sign. This is shown in Fig. 15. Note that the m value always equals the number of bumps (local maxima or minima) in the sin curve.

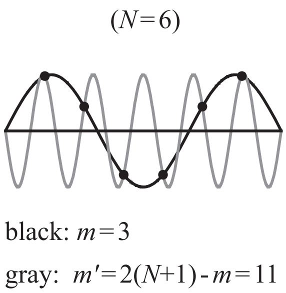


Figure 14

It turns out that there is an infinite number of m values that work, and they fall into two classes. If we start with particular values of N and m (6 and 3 here), then all m' values of the form,

$$m' = 2a(N+1) - m, \tag{74}$$

also work (with a minus sign in the sin curve), where a is any integer. And all m' values of the form,

$$m' = 2a(N+1) + m, \tag{75}$$

also work, where again a is any integer. You can verify these claims by using Eq. (62); see Problem [to be added]. For $N=6$ and $m=3$, the first of these classes contains $m'=11, 25, 39, \ldots$, and the second class contains $m'=3, 17, 31, 45, \ldots$. Negative values of a work too, but they simply reproduce the sin curves for these m' values, up to an overall minus sign.

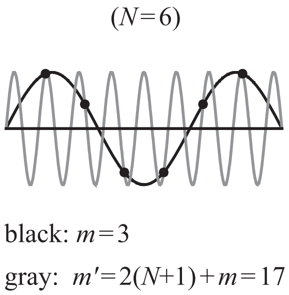


Figure 15

You can show using Eq. (63) that the frequencies corresponding to *all* of these values of m' (in both classes) are equal; see Problem [to be added] (a frequency of $-\omega$ yields the same motion as a frequency of ω). So as far as the motions of the six masses go, all of these modes yield *exactly* the same motions of the masses. (The other parts of the various sin curves don't match up, but all we care about are the locations of the masses.) It is impossible to tell which mode the masses are in. Or said more accurately, the masses aren't really in any one particular mode. There isn't one "correct" mode. Any of the above m or m' values is just as good as any other. However, by convention we label the mode with the m value in the range $1 \le m \le N$.

The above discussion pertains to a setup with N discrete masses in a line, with massless springs between them. However, if we have a continuous string/mass system, or in other words a massive spring (we'll talk about such a system in Section 2.4), then the different m' values *do* represent physically different motions. The $m=3$ and $m=17$ curves in Fig. 15 are certainly different. You can think of a continuous system as a setup with $N \longrightarrow \infty$ masses, so all the m values in the range $1 \le m \le N \Longrightarrow 1 \le m \le \infty$ yield different modes. In other words, each value of m yields a different mode.

However, if we have a continuous string, and if we only look at what is happening at equally spaced locations along it, then there is no way to tell what mode the string is really in (and in this case it really *is* in a well defined mode). If the string is in the $m = 11$ mode, and if you only look at the six equally-spaced points we considered above, then you won't be able to tell which of the $m = 3, 11, 17, 25, \ldots$ modes is the correct one.

This ambiguity is known as *aliasing*, or the *nyquist* effect. If you look at only discrete points in space, then you can't tell the true spatial frequency. Or similarly, If you look at only discrete moments in time, then you can't tell the true temporal frequency. This effect manifests itself in many ways in the real world. If you watch a car traveling by under a streetlight (which emits light in quick pulses, unlike an ordinary filament lightbulb), or if you watch a car speed by in a movie (which was filmed at a certain number of frames per second), then the "spokes" on the tires often appear to be moving at a different angular rate than the actual angular rate of the tire. They might even appear to be moving backwards. This is called the "strobe" effect. There are also countless instances of aliasing in electronics.

2.4 $N \to \infty$ and the wave equation

Let's now consider the $N \to \infty$ limit of our mass/spring setup. This means that we'll now effectively have a continuous system. This will actually make the problem easier than the finite-N case in the previous section, and we'll be able to use a quicker method to solve it. If you want, you can use the strategy of taking the $N \to \infty$ limit of the results in the previous section. This method will work (see Problem [to be added]), but we'll derive things from scratch here, because the method we will use is a very important one, and it will come up again in our study of waves.

First, a change of notation. The equilibrium position of each mass will now play a more fundamental role and appear more regularly, so we're going to label it with x instead of n (or instead of the z we used in the first two remarks at the end of Section 2.3.1). So x_n is the equilibrium position of the nth mass (we'll eventually drop the subscript n). We now need a new letter for the displacement of the masses, because we used x_n for this above. We'll use ξ now. So ξ_n is the displacement of the nth mass. The x's are constants (they just label the equilibrium positions, which don't change), and the ξ's are the things that change with time. The actual location of the nth mass is $x_n + \xi_n$, but only the ξ_n part will show up in the $F = ma$ equations, because the x_n terms don't contribute to the acceleration (because they are constant), nor do they contribute to the force (because only the displacement from equilibrium matters, since the spring force is linear).

Instead of the ξ_n notation, we'll use $\xi(x_n)$. And we'll soon drop the subscript n and just write $\xi(x)$. All three of the ξ_n, $\xi(x_n)$, $\xi(x)$ expressions stand for the same thing, namely the displacement from equilibrium of the mass whose equilibrium position is x (and whose numerical label is n). ξ is a function of t too, of course, but we won't bother writing the t dependence yet. But eventually we'll write the displacement as $\xi(x, t)$.

Let $\Delta x \equiv x_n - x_{n-1}$ be the (equal) spacing between the equilibrium positions of all the masses. The x_n values don't change with time, so neither does Δx. If the $n = 0$ mass is located at the origin, then the other masses are located at positions $x_n = n\Delta x$. In our new notation, the $F = ma$ equation in Eq. (42) becomes

$$\begin{aligned} m\ddot{\xi}_n &= k\xi_{n-1} - 2k\xi_n + k\xi_{n+1} \\ \Longrightarrow \quad m\ddot{\xi}(x_n) &= k\xi(x_n - \Delta x) + 2k\xi(x_n) + k\xi(x_n + \Delta x) \\ \Longrightarrow \quad m\ddot{\xi}(x) &= k\xi(x - \Delta x) + 2k\xi(x) + k\xi(x + \Delta x). \end{aligned} \tag{76}$$

In going from the second to the third line, we are able to drop the subscript n because the value of x uniquely determines which mass we're looking at. If we ever care to know the value of n, we can find it via $x_n = n\Delta x \Longrightarrow n = x/\Delta x$. Although the third line holds only for x values that are integral multiples of Δx, we will soon take the $\Delta x \to 0$ limit, in which case the equation holds for essentially all x.

We will now gradually transform Eq. (76) into a very nice result, which is called the *wave equation.* The first step actually involves going backward to the $F = ma$ form in Eq. (41). We have

$$\begin{aligned} m\frac{d^2\xi(x)}{dt^2} &= k\left[\Big(\xi(x+\Delta x) - \xi(x)\Big) - \Big(\xi(x) - \xi(x-\Delta x)\Big)\right] \\ \Longrightarrow \quad \frac{m}{\Delta x}\frac{d^2\xi(x)}{dt^2} &= k\Delta x\left(\frac{\frac{\xi(x+\Delta x)-\xi(x)}{\Delta x} - \frac{\xi(x)-\xi(x-\Delta x)}{\Delta x}}{\Delta x}\right). \end{aligned} \tag{77}$$

We have made these judicious divisions by Δx for the following reason. If we let $\Delta x \to 0$ (which is indeed the case if we have $N \to \infty$ masses in the system), then we can use the definitions of the first and second derivatives to obtain (with primes denoting spatial derivatives)[6]

$$\begin{aligned} \frac{m}{\Delta x}\frac{d^2\xi(x)}{dt^2} &= (k\Delta x)\frac{\xi'(x) - \xi'(x-\Delta x)}{\Delta x} \\ &= (k\Delta x)\xi''(x). \end{aligned} \tag{78}$$

But $m/\Delta x$ is the mass density ρ. And $k\Delta x$ is known as the elastic modulus, E, which happens to have the units of force. So we obtain

$$\rho\frac{d^2\xi(x)}{dt^2} = E\xi''(x). \tag{79}$$

Note that $E \equiv k\Delta x$ is a reasonable quantity to appear here, because the spring constant k for an infinitely small piece of spring is infinitely large (because if you cut a spring in half, its k doubles, etc.). The Δx in the product $k\Delta x$ has the effect of yielding a finite and informative quantity. If various people have various lengths of springs made out of a given material, then these springs have different k values, but they all have the same E value. Basically, if you buy a spring in a store, and if it's cut from a large supply on a big spool, then the spool should be labeled with the E value, because E is a property of the material and independent of the length. k depends on the length.

Since ξ is actually a function of both x and t, let's be explicit and write Eq. (79) as

$$\boxed{\rho\frac{\partial^2\xi(x,t)}{\partial t^2} = E\frac{\partial^2\xi(x,t)}{\partial x^2}} \qquad \text{(wave equation)} \tag{80}$$

This is called the *wave equation.* This equation (or analogous equations for other systems) will appear repeatedly throughout this book. Note that the derivatives are now written as partial derivatives, because ξ is a function of two arguments. Up to the factors of ρ and E, the wave equation is symmetric in x and t.

The second *time* derivative on the lefthand side of Eq. (80) comes from the "a" in $F = ma$. The second *space* derivative on the righthand side comes from the fact that it is the *differences* in the lengths of two springs that yields the net force, and each of these lengths is itself the *difference* of the positions of two masses. So it is the difference of the differences that we're concerned with. In other words, the second derivative.

[6]There is a slight ambiguity here. Is the $(\xi(x+\Delta x)-\xi(x))\Delta x$ term in Eq. (77) equal to $\xi'(x)$ or $\xi'(x+\Delta x)$? Or perhaps $\xi'(x + \Delta x/2)$? It doesn't matter which we pick, as long as we use the same convention for the $(\xi(x) - \xi(x - \Delta x))\Delta x$ term. The point is that Eq. (78) contains the first derivatives at two points (whatever they may be) that differ by Δx, and the difference of these yields the second derivative.

How do we solve the wave equation? Recall that in the finite-N case, the strategy was to guess a solution of the form (using ξ now instead of x),

$$\begin{pmatrix} \vdots \\ \xi_{n-1} \\ \xi_n \\ \xi_{n+1} \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ a_{n-1} \\ a_n \\ a_{n+1} \\ \vdots \end{pmatrix} e^{i\omega t}. \tag{81}$$

If we relabel $\xi_n \to \xi(x_n, t) \to \xi(x, t)$, and $a_n \to a(x_n) \to a(x)$, we can write the guess in the more compact form,

$$\xi(x,t) = a(x)e^{i\omega t}. \tag{82}$$

This is actually an infinite number of equations (one for each x), just as Eq. (81) is an infinite number of equations (one for each n). The $a(x)$ function gives the amplitudes of the masses, just as the original normal mode vector $(A_1, A_2, A_3, \ldots)$ did. If you want, you can think of $a(x)$ as an infinite-component vector.

Plugging this expression for $\xi(x,t)$ into the wave equation, Eq. (80), gives

$$\begin{aligned} \rho \frac{\partial^2}{\partial t^2}\left(a(x)e^{i\omega t}\right) &= E\frac{\partial^2}{\partial x^2}\left(a(x)e^{i\omega t}\right) \\ \Longrightarrow \quad -\omega^2 \rho\, a(x) &= E\frac{d^2}{dx^2}a(x) \\ \Longrightarrow \quad \frac{d^2}{dx^2}a(x) &= -\frac{\omega^2\rho}{E}a(x). \end{aligned} \tag{83}$$

But this is our good ol' simple-harmonic-oscillator equation, so the solution is

$$a(x) = Ae^{\pm ikx} \qquad \text{where} \quad \boxed{k \equiv \omega\sqrt{\frac{\rho}{E}}} \tag{84}$$

k is called the *wave number*. It is usually defined to be a positive number, so we've put in the $\pm$ by hand. Unfortunately, we've already been using k as the spring constant, but there are only so many letters! The context (and units) should make it clear which way we're using k. The wave number k has units of

$$[k] = [\omega]\sqrt{\frac{[\rho]}{[E]}} = \frac{1}{\text{s}}\sqrt{\frac{\text{kg/m}}{\text{kg m/s}^2}} = \frac{1}{\text{m}}\,. \tag{85}$$

So kx is dimensionless, as it should be, because it appears in the exponent in Eq. (84).

What is the physical meaning of k? If λ is the wavelength, then the kx exponent in Eq. (84) increases by 2π whenever x increases by λ. So we have

$$k\lambda = 2\pi \quad \Longrightarrow \quad k = \frac{2\pi}{\lambda}\,. \tag{86}$$

If k were just equal to $1/\lambda$, then it would equal the number of wavelengths (that is, the number of spatial oscillations) that fit into a unit length. With the 2π, it instead equals the number of radians of spatial oscillations that fit into a unit length.

Using Eq. (84), our solution for $\xi(x,t)$ in Eq. (82) becomes

$$\xi(x,t) = a(x)e^{i\omega t} = Ae^{i(\pm kx + \omega t)}. \tag{87}$$

As usual, we could have done all this with an $e^{-i\omega t}$ term in Eq. (81), because only the square of ω came into play (ω is generally assumed to be positive). So we really have the four different solutions,

$$\xi(x,t) = Ae^{i(\pm kx \pm \omega t)}. \tag{88}$$

The most general solution is the sum of these, which gives

$$\xi(x,t) = A_1 e^{i(kx+\omega t)} + A_1^* e^{i(-kx-\omega t)} + A_2 e^{i(kx-\omega t)} + A_2^* e^{i(-kx+\omega t)}, \tag{89}$$

where the complex conjugates appear because ξ must be real. There are many ways to rewrite this expression in terms of trig functions. Depending on the situation you're dealing with, one form is usually easier to deal with than the others, but they're all usable in theory. Let's list them out and discuss them. In the end, each form has four free parameters. We saw above in the third remark at the end of Section 2.3.1 why four was the necessary number in the discrete case, but we'll talk more about this below.

- If we let $A_1 \equiv (B_1/2)e^{i\phi_1}$ and $A_2 \equiv (B_2/2)e^{i\phi_2}$ in Eq. (89), then the imaginary parts of the exponentials cancel, and we end up with

$$\boxed{\xi(x,t) = B_1 \cos(kx + \omega t + \phi_1) + B_2 \cos(kx - \omega t + \phi_2)} \tag{90}$$

  The interpretation of these two terms is that they represent *traveling waves*. The first one moves to the left, and the second one moves to the right. We'll talk about traveling waves below.

- If we use the trig sum formulas to expand the previous expression, we obtain

$$\boxed{\xi(x,t) = C_1 \cos(kx + \omega t) + C_2 \sin(kx + \omega t) + C_3 \cos(kx - \omega t) + C_4 \sin(kx - \omega t)} \tag{91}$$

  where $C_1 = B_1 \cos\phi_1$, etc. This form has the same interpretation of traveling waves. The sines and cosines are simply 90° out of phase.

- If we use the trig sum formulas again and expand the previous expression, we obtain

$$\boxed{\xi(x,t) = D_1 \cos kx \cos \omega t + D_2 \sin kx \sin \omega t + D_3 \sin kx \cos \omega t + D_4 \cos kx \sin \omega t} \tag{92}$$

  where $D_1 = C_1 + C_3$, etc. These four terms are four different *standing waves*. For each term, the masses all oscillate in phase. All the masses reach their maximum position at the same time (the $\cos\omega t$ terms at one time, and the $\sin\omega t$ terms at another), and they all pass through zero at the same time. As a function of time, the plot of each term just grows and shrinks as a whole. The equality of Eqs. (91) and (92) implies that any traveling wave can be written as the sum of standing waves, and vice versa. This isn't terribly obvious; we'll talk about it below.

- If we collect the $\cos\omega t$ terms together in the previous expression, and likewise for the $\sin\omega t$ terms, we obtain

$$\boxed{\xi(x,t) = E_1 \cos(kx + \beta_1) \cos \omega t + E_2 \cos(kx + \beta_2) \sin \omega t} \tag{93}$$

  where $E_1 \cos\beta_1 = D_1$, etc. This form represents standing waves (the $\cos\omega t$ one is 90° ahead of the $\sin\omega t$ one in time), but they're shifted along the x axis due to the β phases. The spatial functions here could just as well be written in terms of sines, or one sine and one cosine. This would simply change the phases by $\pi/2$.

- If we collect the $\cos kx$ terms together in Eq. (92) and likewise for the $\sin kx$ terms, we obtain

$$\boxed{\xi(x,t) = F_1 \cos(\omega t + \gamma_1)\cos kx + F_2 \cos(\omega t + \gamma_2)\sin kx} \tag{94}$$

where $F_1 \cos\gamma_1 = D_1$, etc. This form represents standing waves, but they're not 90° separated in time in this case, due to the γ phases. They are, however, separated by 90° (a quarter wavelength) in space. The time functions here could just as well be written in terms of sines.

Remarks:

1. If there are no walls and the system extends infinitely in both directions (actually, infinite extent in just one direction is sufficient), then ω can take on any value. Eq. (84) then says that k is related to ω via $k = \omega\sqrt{\rho/E}$. We'll look at the various effects of boundary conditions in Chapter 4.
2. The fact that each of the above forms requires four parameters is probably most easily understood by looking at the first form given in Eq. (90). The most general wave with a given frequency ω consists of two oppositely-traveling waves, each of which is described by two parameters (magnitude and phase). So two times two is four.

   You will recall that for each of the modes in the $N = 2$ and $N = 3$ cases we discussed earlier (and any other value of N, too), only two parameters were required: an overall factor in the amplitudes, and a phase in time. Why only two there, but four now? The difference is due to the fact that we had walls in the earlier cases, but no walls now. (The difference is *not* due to the fact that we're now dealing with infinite N.) The effect of the walls (actually, only one wall is needed) is most easily seen by working with the form given in Eq. (92). Assuming that one wall is located at $x = 0$, we see that the two $\cos kx$ terms can't be present, because the displacement must always be zero at $x = 0$. So $D_1 = D_4 = 0$, and we're down to two parameters. We'll have much more to say about such matters in Chapter 4.
3. Remember that the above expressions for $\xi(x,t)$, each of which contains four parameters, represent the general solution for a *given mode* with frequency ω. If the system is undergoing arbitrary motion, then it is undoubtedly in a linear combination of many different modes, perhaps even an infinite number. So four parameters certainly don't determine the system. We need four times the number of modes, which might be infinite. ♣

Traveling waves

Consider one of the terms in Eq. (91), say, the $\cos(kx - \omega t)$ one. Let's draw the plot of $\cos(kx - \omega t)$, as a function of x, at two times separated by Δt. If we arbitrarily take the lesser time to be $t = 0$, the result is shown in Fig. 16. Basically, the left curve is a plot of $\cos kx$, and the right curve is a plot of $\cos(kx - \phi)$, where ϕ happens to be $\omega\Delta t$. It is shifted to the right because it takes a larger value of x to obtain the same phase.

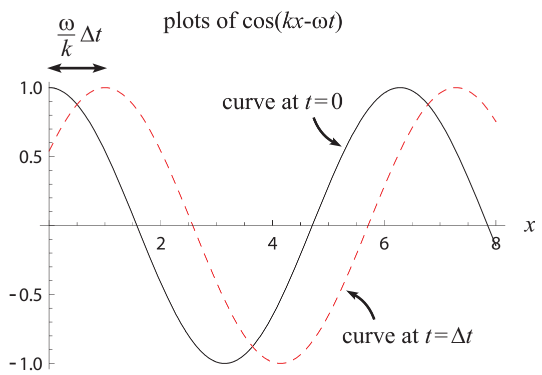


Figure 16

What is the horizontal shift between the curves? We can answer this by finding the distance between the maxima, which are achieved when the argument $kx - \omega t$ equals zero (or a multiple of 2π). If $t = 0$, then we have $kx - \omega \cdot 0 = 0 \implies x = 0$. And if $t = \Delta t$, then we have $kx - \omega \cdot \Delta t = 0 \implies x = (\omega/k)\Delta t$. So $(\omega/k)\Delta t$ is the horizontal shift. It takes a time of Δt for the wave to cover this distance, so the velocity of the wave is

$$v = \frac{(\omega/k)\Delta t}{\Delta t} \implies \boxed{v = \frac{\omega}{k}} \tag{95}$$

Likewise for the $\sin(kx - \omega t)$ function in Eq. (91). Similarly, the velocity of the $\cos(kx + \omega t)$ and $\sin(kx + \omega t)$ curves is $-\omega/k$.

We see that the wave $\cos(kx - \omega t)$ keeps its shape and travels along at speed ω/k. Hence the name "traveling wave." But note that none of the masses are actually moving with this speed. In fact, in our usual approximation of small amplitudes, the actual velocities of the masses are very small. If we double the amplitudes, then the velocities of the masses are doubled, but the speed of the waves is still ω/k.

As we discussed right after Eq. (92), the terms in that equation are standing waves. They don't travel anywhere; they just expand and contract in place. All the masses reach their maximum position at the same time, and they all pass through zero at the same time. This is certainly *not* the case with a traveling wave. Trig identities of the sort, $\cos(kx - \omega t) = \cos kx \cos \omega t + \sin kx \sin \omega t$, imply that any traveling wave can be written as the sum of two standing waves. And trig identities of the sort, $\cos kx \cos \omega t = \big(\cos(kx - \omega t) + \cos(kx + \omega t)\big)/2$, imply that any standing wave can be written as the sum of two opposite traveling waves. The latter of these facts is reasonably easy to visualize, but the former is trickier. You should convince yourself that it works.

A more general solution

We'll now present a much quicker method of finding a much more general solution (compared with our sinusoidal solutions above) to the wave equation in Eq. (80). This is a win-win combination.

From Eq. (84), we know that $k = \omega\sqrt{\rho/E}$. Combining this with Eq. (95) gives $\sqrt{E/\rho} = \omega/k = v \implies E/\rho = v^2$. In view of this relation, if we divide the wave equation in Eq. (80) by ρ, we see that it can be written as

$$\boxed{\frac{\partial^2 \xi(x,t)}{\partial t^2} = v^2 \frac{\partial^2 \xi(x,t)}{\partial x^2}} \tag{96}$$

Consider now the function $f(x - vt)$, where f is an *arbitrary* function of its argument. (The function $f(x + vt)$ will work just as well.) There is no need for f to even vaguely resemble a sinusoidal function. What happens if we plug $\xi(x,t) \equiv f(x - vt)$ into Eq. (96)? Does it satisfy the equation? Indeed it does, as we can see by using the chain rule. In what follows, we'll use the notation f'' to denote the second derivative of f. In other words, $f''(x - vt)$ equals $d^2 f(z)/dz^2$ evaluated at $z = x - vt$. (Since f is a function of only one variable, there is no need for any partial derivatives.) Eq. (96) then becomes (using the chain rule on the left, and also on the right in a trivial sense)

$$\begin{aligned} \frac{\partial^2 f(x - vt)}{\partial t^2} &\stackrel{?}{=} v^2 \frac{\partial^2 f(x - vt)}{\partial x^2} \\ \iff (-v)^2 f''(x - vt) &\stackrel{?}{=} v^2 \cdot (1)^2 f''(x - vt), \end{aligned} \tag{97}$$

which is indeed true.

There is a fairly easy way to see graphically why any function of the form $f(x - vt)$ satisfies the wave equation in Eq. (96). The function $f(x-vt)$ represents a wave moving to the right at speed v. This is true because $f(x_0 - vt_0) = f\big((x_0 + v\Delta t) - v(t_0 + \Delta t)\big)$, which says that if you increase t by Δt, then you need to increase x by $v\Delta t$ in order to obtain the same value of f. This is exactly what happens if you take a curve and move it to the right at speed v. This is basically the same reasoning that led to Eq. (95).

We now claim that any curve that moves to the right (or left) at speed v satisfies the wave equation in Eq. (96). Consider a closeup view of the curve near a given point x_0, at two nearby times separated by Δt. The curve is essentially a straight line in the vicinity of x_0, so we have the situation shown in Fig. 17.

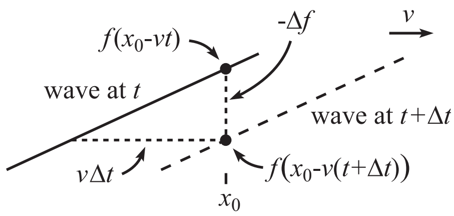


Figure 17

The solid line shows the curve at some time t, and the dotted line shows it at time $t+\Delta t$. The slope of the curve, which is by definition $\partial f/\partial x$, equals the ratio of the lengths of the legs in the right triangle shown. The length of the vertical leg equals the magnitude of the change Δf in the function. Since the change is negative here, the length is $-\Delta f$. But by the definition of $\partial f/\partial t$, the change is $\Delta f = (\partial f/\partial t)\Delta t$. So the length of the vertical leg is $-(\partial f/\partial t)\Delta t$. The length of the horizontal leg is $v\Delta t$, because the curve moves at speed v. So the statement that $\partial f/\partial x$ equals the ratio of the lengths of the legs is

$$\frac{\partial f}{\partial x} = \frac{-(\partial f/\partial t)\Delta t}{v\Delta t} \quad \Longrightarrow \quad \frac{\partial f}{\partial t} = -v\frac{\partial f}{\partial x}\,. \tag{98}$$

If we had used the function $f(x + vt)$, we would have obtained $\partial f/\partial t = v(\partial f/\partial x)$. Of course, these results follow immediately from applying the chain rule to $f(x \pm vt)$. But it's nice to also see how they come about graphically.

Eq. (98) then implies the wave equation in Eq. (96), because if we take $\partial/\partial t$ of Eq. (98), and use the fact that partial differentiation commutes (the order doesn't matter), we obtain

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\frac{\partial f}{\partial t}\right) &= -v\frac{\partial}{\partial t}\left(\frac{\partial f}{\partial x}\right) \\
\Longrightarrow \quad \frac{\partial^2 f}{\partial t^2} &= -v\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial t}\right) \\
&= -v\frac{\partial}{\partial x}\left(-v\frac{\partial f}{\partial x}\right) \\
&= v^2\frac{\partial^2 f}{\partial x^2}\,,
\end{aligned} \tag{99}$$

where we have used Eq. (98) again to obtain the third line. This result agrees with Eq. (96), as desired.

Another way of seeing why Eq. (98) implies Eq. (96) is to factor Eq. (96). Due to the fact that partial differentiation commutes, we can rewrite Eq. (96) as

$$\left(\frac{\partial}{\partial t} - v\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} + v\frac{\partial}{\partial x}\right) f = 0. \tag{100}$$

We can switch the order of these "differential operators" in parentheses, so either of them can be thought of acting on f first. Therefore, if either operator yields zero when applied to f, then the lefthand side of the equation equals zero. In other words, if Eq. (98) is true (with either a plus or a minus on the right side), then Eq. (96) is true.

We have therefore seen (in various ways) that *any* arbitrary function that takes the form of $f(x - vt)$ satisfies the wave equation. This seems too simple to be true. Why did we go through the whole procedure above that involved guessing a solution of the form $\xi(x,t) = a(x)e^{i\omega t}$? Well, that has always been our standard procedure, so the question we should be asking is: Why does an arbitrary function $f(x - vt)$ work?

Well, we gave a few reasons in Eqs. (97) and (98). But here's another reason, one that relates things back to our original sinusoidal solutions. $f(x - vt)$ works because of a combination of *Fourier analysis* and *linearity*. Fourier analysis says that any (reasonably well-behaved) function can be written as the integral (or discrete sum, if the function is periodic) of exponentials, or equivalently sines and cosines. That is,

$$f(z) = \int_{-\infty}^{\infty} C(r)e^{irz}dr. \tag{101}$$

Don't worry about the exact meaning of this; we'll discuss it at great length in the following chapter. But for now, you just need to know that any function $f(z)$ can be considered to be built up out of e^{irz} exponential functions. The coefficient $C(r)$ tells you how much of the function comes from a e^{irz} term with a particular value of r.

Let's now pretend that we haven't seen Eq. (97), but that we *do* know about Fourier analysis. Given the result in Eq. (101), if someone gives us the function $f(x-vt)$ out of the blue, we can write it as

$$f(x - vt) = \int_{-\infty}^{\infty} C(r)e^{ir(x-vt)}dr. \tag{102}$$

But $e^{ir(x-vt)}$ can be written as $e^{i(kx-\omega t)}$, where $k \equiv r$ and $\omega \equiv rv$. Since these values of k and ω satisfy $\omega/k = v$, and hence satisfy Eq. (84) (assuming that v has been chosen to equal $\sqrt{E/\rho}$), we know that all of these $e^{ir(x-vt)}$ terms satisfy the wave equation, Eq. (80). And since the wave equation is *linear* in ξ, it follows that *any* sum (or integral) of these exponentials also satisfies the wave equation. Therefore, in view of Eq. (102), we see that any arbitrary function $f(x - vt)$ satisfies the wave equation. As stated above, both Fourier analysis and linearity are essential in this result.

Fourier analysis plays an absolutely critical role in the study of waves. In fact, it is so important that we'll spend all of Chapter 3 on it. We'll then return to our study of waves in Chapter 4. We'll pick up where we left off here.

Chapter 3

Fourier analysis

 (Version 1, November 28, 2009)

This file contains the Fourier-analysis chapter of a potential book on Waves, designed for college sophomores.

Fourier analysis is the study of how general functions can be decomposed into trigonometric or exponential functions with definite frequencies. There are two types of Fourier expansions:

- *Fourier series*: If a (reasonably well-behaved) function is periodic, then it can be written as a *discrete sum* of trigonometric or exponential functions with specific frequencies.

- *Fourier transform*: A general function that isn't necessarily periodic (but that is still reasonably well-behaved) can be written as a *continuous integral* of trigonometric or exponential functions with a continuum of possible frequencies.

The reason why Fourier analysis is so important in physics is that many (although certainly not all) of the differential equations that govern physical systems are *linear*, which implies that the sum of two solutions is again a solution. Therefore, since Fourier analysis tells us that any function can be written in terms of sinusoidal functions, we can limit our attention to these functions when solving the differential equations. And then we can build up any other function from these special ones. This is a very helpful strategy, because it is invariably easier to deal with sinusoidal functions than general ones.

The outline of this chapter is as follows. We start off in Section 3.1 with Fourier trigonometric series and look at how any periodic function can be written as a discrete sum of sine and cosine functions. Then, since anything that can be written in terms of trig functions can also be written in terms of exponentials, we show in Section 3.2 how any periodic function can be written as a discrete sum of exponentials. In Section 3.3, we move on to Fourier transforms and show how an arbitrary (not necessarily periodic) function can be written as a continuous integral of trig functions or exponentials. Some specific functions come up often when Fourier analysis is applied to physics, so we discuss a few of these in Section 3.4. One very common but somewhat odd function is the *delta function*, and this is the subject of Section 3.5.

Section 3.6 deals with an interesting property of Fourier series near discontinuities called the *Gibbs phenomenon*. This isn't so critical for applications to physics, but it's a very interesting mathematical phenomenon. In Section 3.7 we discuss the conditions under which a Fourier series actually converges to the function it is supposed to describe. Again, this discussion is more just for mathematical interest, because the functions we deal with in physics are invariably well-enough behaved to prevent any issues with convergence. Finally, in Section 3.8 we look at the relation between Fourier series and Fourier transforms. Using the tools we develop in the chapter, we end up being able to derive Fourier's theorem (which says that any periodic function can be written as a discrete sum of sine and cosine functions) from scratch, whereas we simply had to accept this on faith in Section 3.1.

To sum up, Sections 3.1 through 3.5 are very important for physics, while Sections 3.6 through 3.8 are more just for your amusement.

3.1 Fourier trigonometric series

Fourier's theorem states that any (reasonably well-behaved) function can be written in terms of trigonometric or exponential functions. We'll eventually prove this theorem in Section 3.8.3, but for now we'll accept it without proof, so that we don't get caught up in all the details right at the start. But what we *will* do is derive what the coefficients of the sinusoidal functions must be, under the assumption that any function can in fact be written in terms of them.

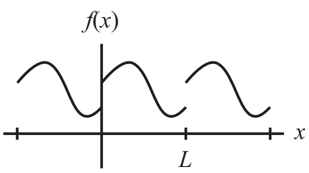


Figure 1

Consider a function $f(x)$ that is periodic on the interval $0 \leq x \leq L$. An example is shown in Fig. 1. Fourier's theorem works even if $f(x)$ isn't continuous, although an interesting thing happens at the discontinuities, which we'll talk about in Section 3.6. Other conventions for the interval are $-L \leq x \leq L$, or $0 \leq x \leq 1$, or $-\pi \leq x \leq \pi$, etc. There are many different conventions, but they all lead to the same general result in the end. If we assume $0 \leq x \leq L$ periodicity, then Fourier's theorem states that $f(x)$ can be written as

$$\boxed{f(x) = a_0 + \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{2\pi n x}{L}\right) + b_n \sin\left(\frac{2\pi n x}{L}\right) \right]} \tag{1}$$

where the a_n and b_n coefficients take on certain values that we will calculate below. This expression is the *Fourier trigonometric series* for the function $f(x)$. We could alternatively not separate out the a_0 term, and instead let the sum run from $n = 0$ to ∞, because $\cos(0) = 1$ and $\sin(0) = 0$. But the normal convention is to isolate the a_0 term.

With the 2π included in the arguments of the trig functions, the $n = 1$ terms have period L, the $n = 2$ terms have period $L/2$, and so on. So for any integer n, an integral number of oscillations fit into the period L. The expression in Eq. (1) therefore has a period of (at most) L, which is a necessary requirement, of course, for it to equal the original periodic function $f(x)$. The period can be shorter than L if, say, only the even n's have nonzero coefficients (in which case the period is $L/2$). But it can't be longer than L; the function repeats at least as often as with period L.

We're actually making two statements in Eq. (1). The first statement is that any periodic function *can* be written this way. This is by no means obvious, and it is the part of the theorem that we're accepting here. The second statement is that the a_n and b_n coefficients take on particular values, *assuming* that the function $f(x)$ can be written this way. It's reasonably straightforward to determine what these values are, in terms of $f(x)$, and we'll do this below. But we'll first need to discuss the concept of *orthogonal functions.*

Orthogonal functions

For given values of n and m, consider the integral,

$$\int_0^L \sin\left(\frac{2\pi n x}{L}\right) \cos\left(\frac{2\pi m x}{L}\right) dx. \tag{2}$$

Using the trig sum formulas, this can be written as

$$\frac{1}{2}\int_0^L \left[\sin\left((n+m)\frac{2\pi x}{L}\right) + \sin\left((n-m)\frac{2\pi x}{L}\right)\right] dx. \tag{3}$$

But this equals zero, because both of the terms in the integrand undergo an integral number of complete oscillations over the interval from 0 to L, which means that the total area under the curve is zero. (The one exception is the $(n-m)$ term if $n = m$, but in that case the sine is identically zero anyway.) Alternatively, you can simply evaluate the integrals; the results involve cosines whose values are the same at the endpoints of integration. (But again, in the case where $n = m$, the $(n-m)$ term must be treated separately.)

The same kind of reasoning shows that the integral,

$$\int_0^L \cos\left(\frac{2\pi nx}{L}\right)\cos\left(\frac{2\pi mx}{L}\right) dx = \frac{1}{2}\int_0^L \left[\cos\left((n+m)\frac{2\pi x}{L}\right) + \cos\left((n-m)\frac{2\pi x}{L}\right)\right] dx, \tag{4}$$

equals zero except in the special case where $n = m$. If $n = m$, the $(n-m)$ term is identically 1, so the integral equals $L/2$. (Technically $n = -m$ also yields a nonzero integral, but we're concerned only with positive n and m.) Likewise, the integral,

$$\int_0^L \sin\left(\frac{2\pi nx}{L}\right)\sin\left(\frac{2\pi mx}{L}\right) dx = \frac{1}{2}\int_0^L \left[-\cos\left((n+m)\frac{2\pi x}{L}\right) + \cos\left((n-m)\frac{2\pi x}{L}\right)\right] dx, \tag{5}$$

equals zero except in the special case where $n = m$, in which case the integral equals $L/2$.

To sum up, we have demonstrated the following facts:

$$\begin{aligned}
\int_0^L \sin\left(\frac{2\pi nx}{L}\right)\cos\left(\frac{2\pi mx}{L}\right) dx &= 0, \\
\int_0^L \cos\left(\frac{2\pi nx}{L}\right)\cos\left(\frac{2\pi mx}{L}\right) dx &= \frac{L}{2}\delta_{nm}, \\
\int_0^L \sin\left(\frac{2\pi nx}{L}\right)\sin\left(\frac{2\pi mx}{L}\right) dx &= \frac{L}{2}\delta_{nm},
\end{aligned} \tag{6}$$

where δ_{nm} (known as the Kronecker delta) is defined to be 1 if $n = m$, and zero otherwise. In short, the integral of the product of any two of these trig functions is zero unless the two functions are actually the same function. The fancy way of saying this is that all of these functions are *orthogonal*.

We normally use the word "orthogonal" when we talk about vectors. Two vectors are orthogonal if their inner product (also called the dot product) is zero, where the inner product is defined to be the sum of the products of corresponding components of the vectors. For example, in three dimensions we have

$$(a_x, a_y, a_z)\cdot(b_x, b_y, b_z) = a_x b_x + a_y b_y + a_z b_z. \tag{7}$$

For functions, if we define the inner product of the above sine and cosine functions to be the integral of their product, then we can say that two functions are orthogonal if their inner product is zero, just as with vectors. So our definition of the inner product of two functions, $f(x)$ and $g(x)$, is

$$\text{Inner product} \equiv \int f(x)g(x)\,dx. \tag{8}$$

This definition depends on the limits of integration, which can be chosen arbitrarily (we're taking them to be 0 and L). This definition is more than just cute terminology. The inner product between two functions defined in this way is actually *exactly the same thing* as the inner product between two vectors, for the following reason.

Let's break up the interval $0 \le x \le L$ into a thousand tiny intervals and look at the thousand values of a given function at these points. If we list out these values next to each other, then we've basically formed a thousand-component vector. If we want to calculate the inner product of two such functions, then a good approximation to the continuous integral in Eq. (8) is the discrete sum of the thousand products of the values of the two functions at corresponding points. (We then need to multiply by $dx = L/1000$, but that won't be important here.) But this discrete sum is exactly what you would get if you formed the inner product of the two thousand-component vectors representing the values of the two functions.

Breaking the interval $0 \le x \le L$ into a million, or a billion, etc., tiny intervals would give an even better approximation to the integral. So in short, a function can be thought of as an infinite-component vector, and the inner product is simply the standard inner product of these infinite-component vectors (times the tiny interval dx).

Calculating the coefficients

Having discussed the orthogonality of functions, we can now calculate the a_n and b_n coefficients in Eq. (1). The a_0 case is special, so let's look at that first.

- Calculating a_0: Consider the integral, $\int_0^L f(x)\,dx$. Assuming that $f(x)$ can be written in the form given in Eq. (1) (and we are indeed assuming that this aspect of Fourier's theorem is true), then all the sines and cosines undergo an integral number of oscillations over the interval $0 \le x \le L$, so they integrate to zero. The only surviving term is the a_0 one, and it yields an integral of $a_0 L$. Therefore,

$$\int_0^L f(x)\,dx = a_0 L \quad \Longrightarrow \quad \boxed{a_0 = \frac{1}{L}\int_0^L f(x)\,dx} \tag{9}$$

  $a_0 L$ is simply the area under the curve.

- Calculating a_n: Now consider the integral,

$$\int_0^L f(x)\cos\left(\frac{2\pi m x}{L}\right)dx, \tag{10}$$

  and again assume that $f(x)$ can be written in the form given in Eq. (1). Using the first of the orthogonality relations in Eq. (6), we see that all of the sine terms in $f(x)$ integrate to zero when multiplied by $\cos(2\pi m x/L)$. Similarly, the second of the orthogonality relations in Eq. (6) tells us that the cosine terms integrate to zero for all values of n except in the special case where $n = m$, in which case the integral equals $a_m(L/2)$. Lastly, the a_0 term integrates to zero, because $\cos(2\pi m x/L)$ undergoes an integral number of oscillations. So we end up with

$$\int_0^L f(x)\cos\left(\frac{2\pi m x}{L}\right)dx = a_m\frac{L}{2} \quad \Longrightarrow \quad \boxed{a_n = \frac{2}{L}\int_0^L f(x)\cos\left(\frac{2\pi n x}{L}\right)dx} \tag{11}$$

  where we have switched the m label to n, simply because it looks nicer.

- Calculating b_n: Similar reasoning shows that

$$\int_0^L f(x)\sin\left(\frac{2\pi m x}{L}\right)dx = b_m\frac{L}{2} \quad \Longrightarrow \quad \boxed{b_n = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{2\pi n x}{L}\right)dx} \tag{12}$$

The limits of integration in the above expressions don't actually have to be 0 and L. They can be any two values of x that differ by L. For example, $-L/4$ and $3L/4$ will work just as well. Any interval of length L will yield the same a_n and b_n coefficients, because both $f(x)$ and the trig functions are periodic. This can be seen in Fig. 2, where two intervals of length L are shown. (As mentioned above, it's fine if $f(x)$ isn't continuous.) These intervals are shaded with slightly different shades of gray, and the darkest region is common to both intervals. The two intervals yield the same integral, because for the interval on the right, we can imagine taking the darkly shaded region and moving it a distance L to the right, which causes the right interval to simply be the left interval shifted by L. Basically, no matter where we put an interval of length L, each point in the period of $f(x)$ and in the period of the trig function gets represented exactly once.

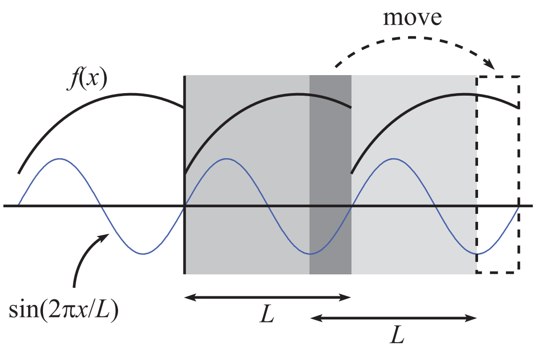


Figure 2

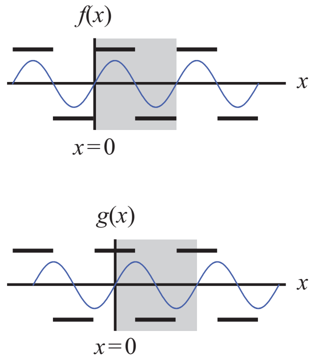


Figure 3

REMARK: However, if we actually *shift the origin* (that is, define another point to be $x = 0$), then the a_n and b_n coefficients change. Shifting the origin doesn't shift the function $f(x)$, but it does shift the sine and cosine curves horizontally, because by definition we're using the functions $\sin(2\pi mx/L)$ and $\cos(2\pi mx/L)$ with no phases. So it matters where we pick the origin. For example, let $f(x)$ be the alternating step function in the first plot in Fig. 3. In the second plot, where we have shifted the origin, the step function remains in the same position on the page, but the sine function is shifted to the right. The integral (over any interval of length L, such as the shaded one) of the product of $f(x)$ and the sine curve in the first plot is not equal to the integral of the product of $g(x)$ and the shifted (by $L/4$) sine curve in the second plot. It is nonzero in the first plot, but zero in the second plot (both of these facts follow from even/odd-ness). We're using g instead of f here to describe the given "curve" in the second plot, because g is technically a different function of x. If the horizontal shift is c (which is $L/4$ here), then g is related to f by $g(x) = f(x + c)$. So the fact that, say, the b_n coefficients don't agree is the statement that the integrals over the shaded regions in the two plots in Fig. 3 don't agree. And this in turn is the statement that

$$\int_0^L f(x)\sin\left(\frac{2\pi mx}{L}\right)dx \neq \int_0^L f(x+c)\sin\left(\frac{2\pi mx}{L}\right)dx. \tag{13}$$

If $f(x)$ has sufficient symmetry, then it is advantageous to shift the origin so that it (or technically the new function $g(x)$) is an even or odd function of x. If $g(x)$ is an even function of x, then only the a_n's survive (the cosine terms), because the b_n's equal the integral of an even function (namely $g(x)$) times an odd function (namely $\sin(2\pi mx/L)$) and therefore vanish. Similarly, if $g(x)$ is an odd function of x, then only the b_n's survive (the sine terms). In particular, the odd function $f(x)$ in the first plot in Fig. 3 has only the b_n terms, whereas the even function $g(x)$ in the second plot has only the a_n terms. But in general, a random choice of the origin will lead to a Fourier series involving both a_n and b_n terms. ♣

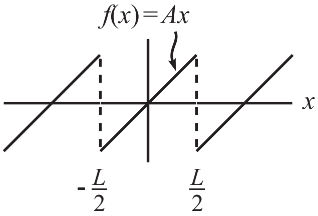


Figure 4

Example (Sawtooth function): Find the Fourier series for the periodic function shown in Fig. 4. It is defined by $f(x) = Ax$ for $-L/2 < x < L/2$, and it has period L.

Solution: Since $f(x)$ is an odd function of x, only the b_n coefficients in Eq. (1) are nonzero. Eq. (12) gives them as

$$b_n = \frac{2}{L}\int_{-L/2}^{L/2} f(x)\sin\left(\frac{2\pi nx}{L}\right)dx = \frac{2}{L}\int_{-L/2}^{L/2} Ax\sin\left(\frac{2\pi nx}{L}\right)dx. \tag{14}$$

Integrating by parts (or just looking up the integral in a table) gives the general result,

$$\begin{aligned}\int x\sin(rx)\,dx &= x\Big(-\frac{1}{r}\cos(rx)\Big)dx - \int -\frac{1}{r}\cos(rx)\,dx \\ &= -\frac{x}{r}\cos(rx) + \frac{1}{r^2}\sin(rx).\end{aligned} \tag{15}$$

Letting $r \equiv 2\pi n/L$ yields

$$\begin{aligned} b_n &= \frac{2A}{L}\left[-x\Big(\frac{L}{2\pi n}\Big)\cos\left(\frac{2\pi nx}{L}\right)\Bigg|_{-L/2}^{L/2} + \Big(\frac{L}{2\pi n}\Big)^2\sin\left(\frac{2\pi nx}{L}\right)\Bigg|_{-L/2}^{L/2}\right] \\ &= \Big(-\frac{AL}{2\pi n}\cos(\pi n) - \frac{AL}{2\pi n}\cos(-\pi n)\Big) + 0 \\ &= -\frac{AL}{\pi n}\cos(\pi n) \\ &= (-1)^{n+1}\frac{AL}{\pi n}. \end{aligned} \tag{16}$$

Eq. (1) therefore gives the Fourier trig series for $f(x)$ as

$$\begin{aligned} f(x) &= \frac{AL}{\pi}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}\sin\left(\frac{2\pi nx}{L}\right) \\ &= \frac{AL}{\pi}\left[\sin\left(\frac{2\pi x}{L}\right) - \frac{1}{2}\sin\left(\frac{4\pi x}{L}\right) + \frac{1}{3}\sin\left(\frac{6\pi x}{L}\right) - \cdots\right]. \end{aligned} \tag{17}$$

The larger the number of terms that are included in this series, the better the approximation to the periodic Ax function. The partial-series plots for 1, 3, 10, and 50 terms are shown in Fig. 5. We have arbitrarily chosen A and L to be 1. The 50-term plot gives a very good approximation to the sawtooth function, although it appears to overshoot the value at the discontinuity (and it also has some wiggles in it). This overshoot is an inevitable effect at a discontinuity, known as the *Gibbs phenomenon*. We'll talk about this in Section 3.6.

Interestingly, if we set $x = L/4$ in Eq. (17), we obtain the cool result,

$$A\cdot\frac{L}{4} = \frac{AL}{\pi}\left(1 + 0 - \frac{1}{3} + 0 + \frac{1}{5}\cdots\right) \implies \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots. \tag{18}$$

Nice expressions for π like this often pop out of Fourier series.

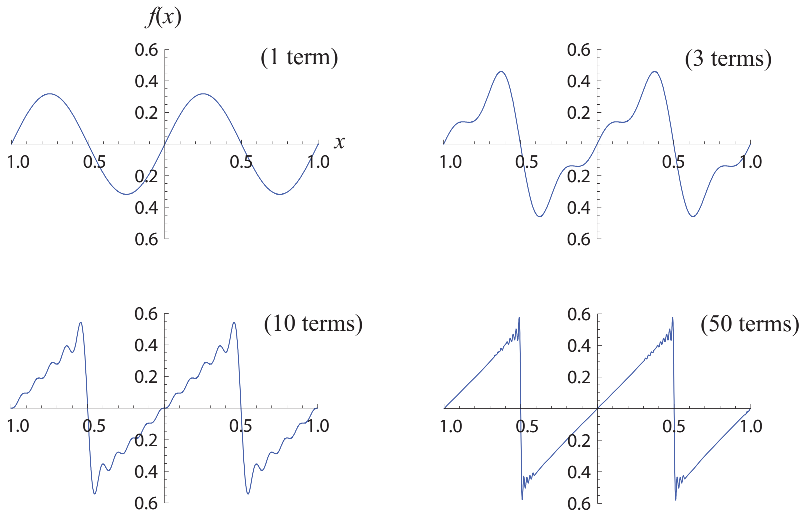


Figure 5

3.2 Fourier exponential series

Any function that can be written in terms of sines and cosines can also be written in terms of exponentials. So assuming that we can write any periodic function in the form given in Eq. (1), we can also write it as

$$\boxed{f(x) = \sum_{n=-\infty}^{\infty} C_n e^{i2\pi nx/L}} \tag{19}$$

where the C_n coefficients are given by

$$\boxed{C_m = \frac{1}{L}\int_0^L f(x)e^{-i2\pi mx/L}dx} \tag{20}$$

as we will show below. This hold for all n, including $n = 0$. Eq. (19) is the *Fourier exponential series* for the function $f(x)$. The sum runs over all the integers here, whereas it runs over only the positive integers in Eq. (1), because the negative integers there give redundant sines and cosines.

We can obtain Eq. (20) in a manner similar to the way we obtained the a_n and b_n coefficients via the orthogonality relations in Eq. (6). The analogous orthogonality relation here is

$$\begin{aligned}\int_0^L e^{i2\pi nx/L}e^{-i2\pi mx/L}dx &= \frac{L}{i2\pi(n-m)}e^{i2\pi(n-m)x/L}\Big|_0^L \\ &= 0, \quad \text{unless } m = n,\end{aligned} \tag{21}$$

because the value of the exponential is 1 at both limits. If $m = n$, then the integral is simply

$\int_0^L 1 \cdot dx = L$. So in the δ_{nm} notation, we have[1]

$$\int_0^L e^{i2\pi nx/L} e^{-i2\pi mx/L} dx = L\delta_{nm}. \tag{22}$$

Just like with the various sine and cosine functions in the previous section, the different exponential functions are orthogonal.[2] Therefore, when we plug the $f(x)$ from Eq. (19) into Eq. (20), the only term that survives from the expansion of $f(x)$ is the one where the n value equals m. And in that case the integral is simply $C_m L$. So dividing the integral by L gives C_m, as Eq. (20) states. As with a_0 in the case of the trig series, Eq. (20) tells us that $C_0 L$ is the area under the curve.

Example (Sawtooth function again): Calculate the C_n's for the periodic Ax function in the previous example.

Solution: Eq. (20) gives (switching from n to m)

$$C_n = \frac{1}{L}\int_{-L/2}^{L/2} Axe^{-i2\pi nx/L} dx. \tag{23}$$

The integral of the general form $\int xe^{-rx}dx$ can be found by integration by parts (or looking it up in a table). The result is

$$\int xe^{-rx}dx = -\frac{x}{r}e^{-rx} - \frac{1}{r^2}e^{-rx}. \tag{24}$$

This is valid unless $r = 0$, in which case the integral equals $\int x\,dx = x^2/2$. For the specific integral in Eq. (23) we have $r = i2\pi n/L$, so we obtain (for $n \neq 0$)

$$C_n = -\frac{A}{L}\left(\frac{xL}{i2\pi n}e^{-i2\pi nx/L}\bigg|_{-L/2}^{L/2} + \left(\frac{L}{i2\pi n}\right)^2 e^{-i2\pi nx/L}\bigg|_{-L/2}^{L/2}\right). \tag{25}$$

The second of these terms yields zero because the limits produce equal terms. The first term yields

$$C_n = -\frac{A}{L}\cdot\frac{(L/2)L}{i2\pi n}\left(e^{-i\pi n} + e^{i\pi n}\right). \tag{26}$$

The sum of the exponentials is simply $2(-1)^n$. So we have (getting the i out of the denominator)

$$C_n = (-1)^n \frac{iAL}{2\pi n} \qquad (\text{for } n \neq 0). \tag{27}$$

If $n = 0$, then the integral yields $C_0 = (A/L)(x^2/2)\big|_{-L/2}^{L/2} = 0$. Basically, the area under the curve is zero since Ax is an odd function. Putting everything together gives the Fourier exponential series,

$$f(x) = \sum_{n\neq 0}(-1)^n \frac{iAL}{2\pi n}e^{i2\pi nx/L}. \tag{28}$$

This sum runs over all the integers (positive and negative) with the exception of 0.

---

[1] The "L" in front of the integral in this equation is twice the "$L/2$" that appears in Eq. (6). This is no surprise, because if we use $e^{i\theta} = \cos\theta + i\sin\theta$ to write the integral in this equation in terms of sines and cosines, we end up with two $\sin\cdot\cos$ terms that integrate to zero, plus a $\cos\cdot\cos$ and a $\sin\cdot\sin$ term, each of which integrates to $(L/2)\delta_{nm}$.

[2] For complex functions, the inner product is defined to be the integral of the product of the functions, where one of them has been complex conjugated. This complication didn't arise with the trig functions in Eq. (6) because they were real. But this definition is needed here, because without the complex conjugation, the inner product of a function with itself would be zero, which wouldn't be analogous to regular vectors.

As a double check on this, we can write the exponential in terms of sines and cosines:

$$f(x) = \sum_{n\neq 0}(-1)^n \frac{iAL}{2\pi n}\left(\cos\left(\frac{2\pi nx}{L}\right) + i\sin\left(\frac{2\pi nx}{L}\right)\right). \tag{29}$$

Since $(-1)^n/n$ is an odd function of n, and since $\cos(2\pi nx/L)$ is an even function of n, the cosine terms sum to zero. Also, since $\sin(2\pi nx/L)$ is an odd function of n, we can restrict the sum to the positive integers, and then double the result. Using $i^2(-1)^n = (-1)^{n+1}$, we obtain

$$f(x) = \frac{AL}{\pi}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}\sin\left(\frac{2\pi nx}{L}\right), \tag{30}$$

in agreement with the result in Eq. (17).

Along the lines of this double-check we just performed, another way to calculate the C_n's that avoids doing the integral in Eq. (20) is to extract them from the trig coefficients, a_n and b_n, if we happen to have already calculated those (which we have, in the case of the periodic Ax function). Due to the relations,

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \qquad \text{and} \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i}, \tag{31}$$

we can replace the trig functions in Eq. (1) with exponentials. We then quickly see that the C_n coefficients can be obtained from the a_n and b_n coefficients via (getting the i's out of the denominators)

$$C_n = \frac{a_n - ib_n}{2}, \qquad \text{and} \qquad C_{-n} = \frac{a_n + ib_n}{2}, \tag{32}$$

and $C_0 = a_0$. The n's in these relations take on only positive values. For the periodic Ax function, we know that the a_n's are zero, and that the b_n's are given in Eq. (16) as $(-1)^{n+1}AL/\pi n$. So we obtain

$$C_n = -i(-1)^{n+1}\frac{AL}{2\pi n}, \qquad \text{and} \qquad C_{-n} = i(-1)^{n+1}\frac{AL}{2\pi n}. \tag{33}$$

If we define m to be $-n$ in the second expression (so that m runs over the negative integers), and then replace the arbitrary letter m with n, we see that both of these expressions can be transformed into a single formula,

$$C_n = (-1)^n\frac{iAL}{2\pi n}, \tag{34}$$

which holds for all integers (positive and negative), except 0. This result agrees with Eq. (27).

Conversely, if we've already calculated the C_n's, and if we want to determine the a_n's and b_n's in the trig series without doing any integrals, then we can use $e^{-z} = \cos z + i\sin z$ to write the exponentials in Eq. (19) in terms of sines and cosines. The result will be a trig series in the form of Eq. (1), with a_n and b_n given by

$$a_n = C_n + C_{-n}, \qquad \text{and} \qquad b_n = i(C_n - C_{-n}), \tag{35}$$

and $a_0 = C_0$. Of course, we can also obtain these relations by simply inverting the relations in Eq. (32). We essentially used these relations in the double-check we performed in the example above.

Real or imaginary, even or odd

If a function $f(x)$ is real (which is generally the case in classical physics), then the nth and $(-n)$th terms in Eq. (19) must be complex conjugates. Because the exponential parts are already complex conjugates of each other, this implies that

$$C_{-n} = C_n^* \qquad \text{(if } f(x) \text{ is real.} \tag{36}$$

For the (real) periodic Ax function we discussed above, the C_n's in Eq. (27) do indeed satisfy $C_{-n} = C_n^*$. In the opposite case where a function is purely imaginary, we must have $C_{-n} = -C_n^*$.

Concerning even/odd-ness, a function $f(x)$ is in general neither an even nor an odd function of x. But if one of these special cases holds, then we can say something about the C_n coefficients.

- If $f(x)$ is an *even* function of x, then there are only cosine terms in the trig series in Eq. (1), so the relative plus sign in the first equation in Eq. (31) implies that

$$C_n = C_{-n}. \tag{37}$$

  If additionally $f(x)$ is real, then the $C_{-n} = C_n^*$ requirement implies that the C_n's are purely real.

- If $f(x)$ is an *odd* function of x, then there are only sine terms in the trig series in Eq. (1), so the relative minus sign in the second equation in Eq. (31) implies that

$$C_n = -C_{-n}. \tag{38}$$

  This is the case for the C_n's we found in Eq. (27). If additionally $f(x)$ is real, then the $C_{-n} = C_n^*$ requirement implies that the C_n's are purely imaginary, which is also the case for the C_n's in Eq. (27).

We'll talk more about these real/imaginary and even/odd issues when we discuss Fourier transforms.

3.3 Fourier transforms

The Fourier trigonometric series given by Eq. (1) and Eqs. (9,11,12), or equivalently the Fourier exponential series given by Eqs. (19,20), work for periodic functions. But what if a function isn't periodic? Can we still write the function as a series involving trigonometric or exponential terms? It turns out that indeed we can, but the series is now a continuous integral instead of a discrete sum. Let's see how this comes about.

The key point in deriving the continuous integral is that we can consider a *nonperiodic* function to be a *periodic* function on the interval $-L/2 \le x \le L/2$, with $L \to \infty$. This might sound like a cheat, but it is a perfectly legitimate thing to do. Given this strategy, we can take advantage of all the results we derived earlier for periodic functions, with the only remaining task being to see what happens to these results in the $L \to \infty$ limit. We'll work with the exponential series here, but we could just as well work with the trigonometric series.

Our starting point will be Eqs. (19) and (20), which we'll copy here:

$$f(x) = \sum_{n=-\infty}^{\infty} C_n e^{i2\pi nx/L}, \qquad \text{where} \qquad C_n = \frac{1}{L}\int_{-L/2}^{L/2} f(x)e^{-i2\pi nx/L}dx, \tag{39}$$

where we have taken the limits to be $-L/2$ and $L/2$. Let's define $k_n \equiv 2\pi n/L$. The difference between successive k_n values is $dk_n = 2\pi(dn)/L = 2\pi/L$, because dn is simply 1. Since we are interested in the $L \to \infty$ limit, this dk_n is very small, so k_n is essentially a continuous variable. In terms of k_n, the expression for $f(x)$ in Eq. (39) becomes (where we have taken the liberty of multiplying by $dn = 1$)

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} C_n e^{ik_n x}(dn) \\ &= \sum_{n=-\infty}^{\infty} C_n e^{ik_n x}\left(\frac{L}{2\pi}\,dk_n\right). \end{aligned} \tag{40}$$

But since dk_n is very small, this sum is essentially an integral:

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty}\left(C_n\frac{L}{2\pi}\right)e^{k_n x}dk_n \\ &\equiv \int_{-\infty}^{\infty} C(k_n)e^{ik_n x}dk_n, \end{aligned} \tag{41}$$

where $C(k_n) \equiv (L/2\pi)C_n$. We can use the expression for C_n in Eq. (39) to write $C(k_n)$ as (in the $L \to \infty$ limit)

$$\begin{aligned} C(k_n) \equiv \frac{L}{2\pi}C_n &= \frac{L}{2\pi}\cdot\frac{1}{L}\int_{-\infty}^{\infty} f(x)e^{-ik_n x}dx \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ik_n x}dx. \end{aligned} \tag{42}$$

We might as well drop the index n, because if we specify k, there is no need to mention n. We can always find n from $k \equiv 2\pi n/L$ if we want. Eqs. (41) and (42) can then be written as

$$\boxed{f(x) = \int_{-\infty}^{\infty} C(k)e^{ikx}dk} \qquad \text{where} \qquad \boxed{C(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ikx}dx} \tag{43}$$

$C(k)$ is known as the *Fourier transform* of $f(x)$, and vice versa. We'll talk below about how we can write things in terms of trigonometric functions instead of exponentials. But the term "Fourier transform" is generally taken to refer to the exponential decomposition of a function.

Similar to the case with Fourier series, the first relation in Eq. (43) tells us that $C(k)$ indicates how much of the function $f(x)$ is made up of e^{ikx}. And conversely, the second relation in Eq. (43) tells us that $f(x)$ indicates how much of the function $C(k)$ is made up of e^{-ikx}. Of course, except in special cases (see Section 3.5), essentially *none* of $f(x)$ is made up of e^{ikx} for one particular value of k, because k is a continuous variable, so there is zero chance of having *exactly* a particular value of k. But the useful thing we can say is that if dk is small, then essentially $C(k)\,dk$ of the function $f(x)$ is made up of e^{ikx} terms in the range from k to $k+dk$. The smaller dk is, the smaller the value of $C(k)\,dk$ is (but see Section 3.5 for an exception to this).

Remarks:

1. These expressions in Eq. (43) are symmetric except for a minus sign in the exponent (it's a matter of convention as to which one has the minus sign), and also a 2π. If you want, you can define things so that each expression has a $1/\sqrt{2\pi}$, by simply replacing $C(k)$ with a $B(k)$ function defined by $B(k) \equiv \sqrt{2\pi}\,C(k)$. In any case, the product of the factors in front of the integrals must be $1/2\pi$.

2. With our convention that one expression has a 2π and the other doesn't, there is an ambiguity when we say, "$f(x)$ and $C(k)$ are Fourier transforms of each other." Better terminology would be to say that $C(k)$ is the Fourier transform of $f(x)$, while $f(x)$ is the "reverse" Fourier transform of $C(k)$. The "reverse" just depends on where the 2π is. If x measures a position and k measures a wavenumber, the common convention is the one in Eq. (43).

   Making both expressions have a "$1/\sqrt{2\pi}$" in them would lead to less of a need for the word "reverse," although there would still be a sign convention in the exponents which would have to be arbitrarily chosen. However, the motivation for having the first equation in Eq. (43) stay the way it is, is that the $C(k)$ there tells us how much of $f(x)$ is made up of e^{ikx}, which is a more natural thing to be concerned with than how much of $f(x)$ is made up of $e^{ikx}/\sqrt{2\pi}$. The latter is analogous to expanding the Fourier series in Eq. (1) in terms of $1/\sqrt{2\pi}$ times the trig functions, which isn't the most natural thing to do. The price to pay for the convention in Eq. (43) is the asymmetry, but that's the way it goes.

3. Note that both expressions in Eq. (43) involve integrals, whereas in the less symmetric Fourier-series results in Eq. (39), one expression involves an integral and one involves a discrete sum. The integral in the first equation in Eq. (43) is analogous to the Fourier-series sum in the first equation in Eq. (39), but unfortunately this integral doesn't have a nice concise name like "Fourier series." The name "Fourier-transform expansion" is probably the most sensible one. At any rate, the Fourier transform itself, $C(k)$, is analogous to the Fourier-series *coefficients* C_n in Eq. (39). The transform doesn't refer to the whole integral in the first equation in Eq. (43). ♣

Real or imaginary, even or odd

Anything that we can do with exponentials we can also do with trigonometric functions. So let's see what Fourier transforms look like in terms of these. Because cosines and sines are even and odd functions, respectively, the best way to go about this is to look at the even/odd-ness of the function at hand. Any function $f(x)$ can be written uniquely as the sum of an even function and an odd function. These even and odd parts (call them $f_\text{e}(x)$ and $f_\text{o}(x)$) are given by

$$f_\text{e}(x) = \frac{f(x)+f(-x)}{2} \qquad \text{and} \qquad f_\text{o}(x) = \frac{f(x)-f(-x)}{2}. \tag{44}$$

You can quickly verify that these are indeed even and odd functions. And $f(x) = f_\text{e}(x) + f_\text{o}(x)$, as desired.

Remark: These functions are unique, for the following reason. Assume that there is another pair of even and odd functions whose sum is $f(x)$. Then they must take the form of $f_\text{e}(x) + g(x)$ and $f_\text{o}(x) - g(x)$ for some function $g(x)$. The first of these relations says that $g(x)$ must be even, but the second says that $g(x)$ must be odd. The only function that is both even and odd is the zero function. So the new pair of functions must actually be the same as the pair we already had. ♣

In terms of $f_\text{e}(x)$ and $f_\text{o}(x)$, the $C(k)$ in the second equation in Eq. (43) can be written as

$$\begin{aligned} C(k) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ikx}\,dx \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \Big(f_\text{e}(x)+f_\text{o}(x)\Big)\Big(\cos(-kx)+i\sin(-kx)\Big)\,dx. \end{aligned} \tag{45}$$

If we multiply out the integrand, we obtain four terms. Two of them (the $f_\text{e}\cos$ and $f_\text{o}\sin$ ones) are even functions of x. And two of them (the $f_\text{e}\sin$ and $f_\text{o}\cos$ ones) are odd functions of x. The odd ones integrate to zero, so we are left with

$$C(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f_\text{e}(x)\cos(kx)\,dx - \frac{i}{2\pi}\int_{-\infty}^{\infty} f_\text{o}(x)\sin(kx)\,dx. \tag{46}$$

Up to this point we've been considering the even/odd-ness of $f(x)$ as a function of x. But let's now consider the even/odd-ness of $C(k)$ as a function of k. The first term in Eq. (46) is an even function of k (because k appears only in the cosine function), and the second is odd (because k appears only in the sine function). So we can read off the even and odd parts of $C(k)$:

$$\boxed{C_{\text{e}}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f_{\text{e}}(x)\cos(kx)\,dx} \quad \text{and} \quad \boxed{C_{\text{o}}(k) = -\frac{i}{2\pi}\int_{-\infty}^{\infty} f_{\text{o}}(x)\sin(kx)\,dx} \tag{47}$$

Likewise, the inverse relations are quickly obtained by letting $C(k) = C_{\text{e}}(k) + C_{\text{o}}(k)$ in the expression for $f(x)$ in Eq. (43). The result is

$$f(x) = \int_{-\infty}^{\infty} C_{\text{e}}(k)\cos(kx)\,dk + i\int_{-\infty}^{\infty} C_{\text{o}}(k)\sin(kx)\,dk, \tag{48}$$

and so

$$\boxed{f_{\text{e}}(x) = \int_{-\infty}^{\infty} C_{\text{e}}(k)\cos(kx)\,dk} \quad \text{and} \quad \boxed{f_{\text{o}}(x) = i\int_{-\infty}^{\infty} C_{\text{o}}(k)\sin(kx)\,dk} \tag{49}$$

Similar to the case with the exponential decomposition, $C_{\text{e}}(k)$ tells us how much of the function $f(x)$ is made up of $\cos(kx)$, and $iC_{\text{o}}(k)$ tells us how much of the function $f(x)$ is made up of $\sin(kx)$. In view of Eq. (48), the functions $C_{\text{e}}(k)$ and $iC_{\text{o}}(k)$ can reasonably be called the "Fourier trig transforms" of $f(x)$.

Note that if we replace the $\cos(kx)$ in the first equations Eqs. (47) and (49) by e^{ikx}, the integral is unaffected, because the sine part of the exponential is an odd function (of either x or k, as appropriate) and therefore doesn't contribute anything to the integral. These equations therefore tell us that $C_{\text{e}}(k)$ is the Fourier transform of $f_{\text{e}}(x)$. And likewise $iC_{\text{o}}(k)$ is the Fourier transform of $f_{\text{o}}(x)$. We see that the pair of functions $f_{\text{e}}(x)$ and $C_{\text{e}}(k)$ is completely unrelated to the pair $f_{\text{o}}(x)$ and $C_{\text{o}}(k)$, as far as Fourier transforms are concerned. The two pairs don't "talk" to each other.

Eq. (46) tells us that if $f(x)$ is real (which is generally the case in classical physics), then the real part of $C(k)$ is even, and the imaginary part is odd. A concise way of saying these two things is that $C(k)$ and $C(-k)$ are complex conjugates. That is, $C(-k) = C(k)^*$. Likewise, Eq. (48) tells us that if $C(k)$ is real, then $f(-x) = f(x)^*$. (Related statements hold if $f(x)$ or $C(k)$ is imaginary.)

For the special cases of purely even/odd and real/imaginary functions, the following facts follow from Eq. (47):

- If $f(x)$ is even and real, then $C(k)$ is even and real.
- If $f(x)$ is even and imaginary, then $C(k)$ is even and imaginary.
- If $f(x)$ is odd and real, then $C(k)$ is odd and imaginary.
- If $f(x)$ is odd and imaginary, then $C(k)$ is odd and real.

The converses are likewise all true, by Eq. (49).

Let's now do some examples where we find the Fourier trig series and Fourier (trig) transform of two related functions.

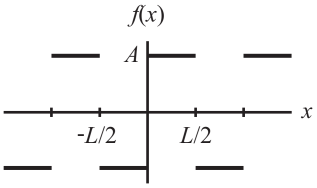


Figure 6

Example (Periodic odd step function): Calculate the Fourier trig series for the periodic odd step function shown in Fig. 6. The function takes on the values $\pm A$.

Solution: The function is odd, so only the sin terms survive in Eq. (1). Eq. (12) gives the b_n coefficients as

$$\begin{aligned} b_n &= \frac{2}{L}\int_{-L/2}^{L/2} f(x)\sin\left(\frac{2\pi n x}{L}\right)dx = 2\cdot\frac{2}{L}\int_{0}^{L/2} A\sin\left(\frac{2\pi n x}{L}\right)dx \\ &= -\frac{4A}{L}\left(\frac{L}{2\pi n}\right)\cos\left(\frac{2\pi n x}{L}\right)\Big|_0^{L/2} = \frac{2A}{\pi n}\big(1-\cos(\pi n)\big). \end{aligned} \tag{50}$$

This equals $4A/\pi n$ is n is odd, and zero if n is even. So we have

$$\begin{aligned} f(x) &= \sum_{n=1,\mathrm{odd}}^{\infty} \frac{4A}{\pi n}\sin\left(\frac{2\pi n x}{L}\right) \\ &= \frac{4A}{\pi}\left(\sin\left(\frac{2\pi x}{L}\right) + \frac{1}{3}\sin\left(\frac{6\pi x}{L}\right) + \frac{1}{5}\sin\left(\frac{10\pi x}{L}\right) + \cdots\right). \end{aligned} \tag{51}$$

Some partial plots are shown in Fig. 7. The plot with 50 terms yields a very good approximation to the step function, although there is an overshoot at the discontinuities (and also some fast wiggles). We'll talk about this in Section 3.6.

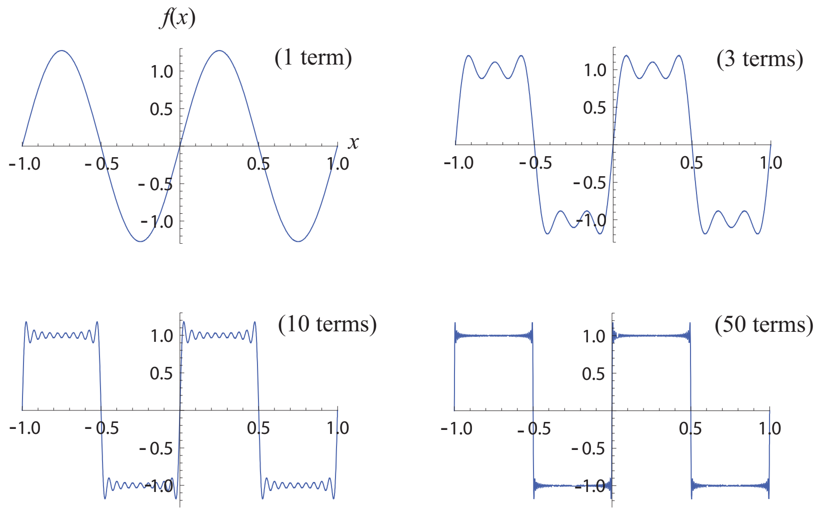


Figure 7

For the fun of it, we can plug $x = L/4$ into Eq. (51) to obtain

$$A = \frac{4A}{\pi}\left(1 - \frac{1}{3} + \frac{1}{5} - \cdots\right) \quad\Longrightarrow\quad \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots, \tag{52}$$

which is the same result we found in Eq. (18).

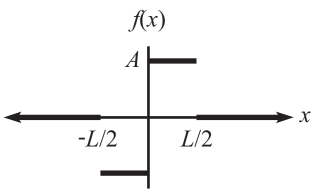


Figure 8

Example (Non-periodic odd step function): Calculate the Fourier trig transform of the non-periodic odd step function shown in Fig. 8. The value is $\pm A$ inside the region $-L/2 < x < L/2$, and zero outside. Do this in two ways:

(a) Find the Fourier series for the periodic "stretched" odd step function shown in Fig. 9, with period NL, and then take the $N \to \infty$ limit.

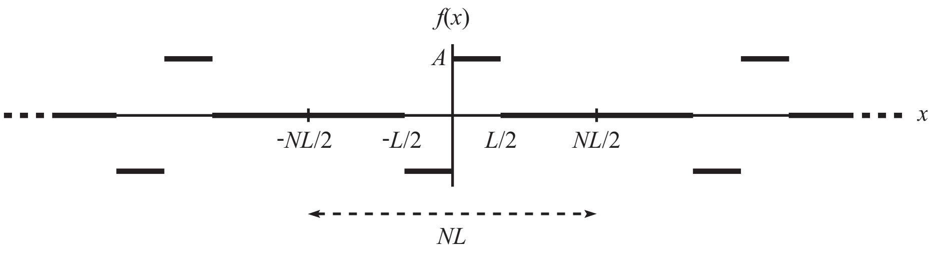


Figure 9

(b) Calculate $C(k)$ by doing the integral in Eq. (43). The result will look different from the result in part (a), but you can demonstrate that the two are equivalent.

Solution:

(a) The function in Fig. 9 is odd, so only the sin terms survive in Eq. (1). The length of the interval is NL, so Eq. (12) gives the b_n coefficients as

$$b_n = \frac{2}{NL}\int_{-NL/2}^{NL/2} f(x)\sin\left(\frac{2\pi nx}{NL}\right)dx. \tag{53}$$

But $f(x)$ is nonzero only in the interval $-L/2 < x < L/2$. We can consider just the $0 < x < L/2$ half and then multiply by 2. So we have

$$\begin{aligned} b_n = 2\cdot\frac{2}{NL}\int_0^{L/2} A\sin\left(\frac{2\pi nx}{NL}\right)dx &= -\frac{4A}{NL}\left(\frac{NL}{2\pi n}\right)\cos\left(\frac{2\pi nx}{NL}\right)\Big|_0^{L/2} \\ &= \frac{2A}{\pi n}\left(1-\cos\left(\frac{\pi n}{N}\right)\right). \end{aligned} \tag{54}$$

If $N = 1$, this agrees with the result in Eq. (50) in the previous example, as it should. The Fourier series for the function is therefore

$$f(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi nx}{NL}\right) = \sum_{n=1}^{\infty}\frac{2A}{\pi n}\left(1-\cos\left(\frac{\pi n}{N}\right)\right)\sin\left(\frac{2\pi nx}{NL}\right). \tag{55}$$

We will now find the Fourier trig transform by taking the $N \to \infty$ limit. Define z_n by $z_n \equiv n/N$. The changes in z_n and n are related by $dn = N\,dz_n$. dn is simply 1, of course, so we can multiple the result in Eq. (55) by dn without changing anything. If we then replace dn by $N\,dz_n$, and also get rid of the n's in favor of z_n's, we obtain

$$f(x) = \sum_{n=1}^{\infty}\frac{2A}{\pi(Nz_n)}\Big(1-\cos(\pi z_n)\Big)\sin\left(\frac{2\pi z_n x}{L}\right)N\,dz_n \tag{56}$$

In the $N \to \infty$ limit, dz_n is very small, so we can change this sum to an integral. And we can drop the n subscript on z, because the value of z determines the associated value of n (if we happen to care what it is). So we arrive at the desired Fourier trig transform:

$$f(x) = \frac{2A}{\pi}\int_0^{\infty}\frac{1}{z}\Big(1-\cos(\pi z)\Big)\sin\left(\frac{2\pi zx}{L}\right)dz \qquad \text{(for non-periodic } f\text{)}. \tag{57}$$

If we want to put this in the standard Fourier-transform form where the integral runs from $-\infty$ to ∞, we just need to make the coefficient be A/π. The integrand is an even function of z, so using a lower limit of $-\infty$ simply doubles the integral.

The Fourier trig *transform* in Eq. (57) for the *non-periodic* odd step function should be compared with the Fourier *series* for the *periodic* odd step function in the previous example. If we use the form of b_n in Eq. (50) (that is, without evaluating it for even and odd n), the series looks like

$$f(x) = \frac{2A}{\pi}\sum_{n=0}^{\infty}\frac{1}{n}\Big(1-\cos(\pi n)\Big)\sin\left(\frac{2\pi nx}{L}\right) \qquad \text{(for periodic } f\text{)} \tag{58}$$

The only difference between this equation and Eq. (57) is that n can take on only integer values here, whereas z is a continuous variable in Eq. (57). It is quite fascinating how the switch to a continuous variable changes the function from a periodic one to a function that is zero everywhere except in the region $-L/2 < x < L/2$.

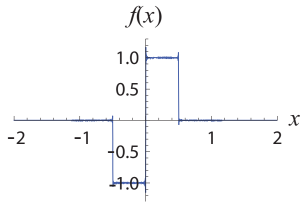


Figure 10

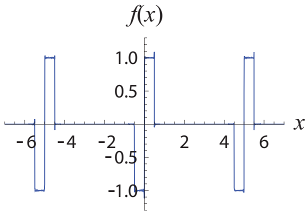


Figure 11

REMARK: What if we try to approximate the integral in Eq. (57) by discretizing z into small units and calculating a discrete sum? For example, if we take the "bin" size to be $\Delta z = 0.2$, then the resulting $f(x)$ appears to have the desired value of zero outside the $-L/2 < x < L/2$ region, as shown in Fig. 10 (we've chosen $L = 1$, and we've truncated the discrete sum at $z = 200$, which is plenty large to get a good approximation). However, the zoomed-out view in Fig. 11 shows that the function actually repeats at $x = \pm 5$. And if we zoomed out farther, we would find that it repeats at ± 10 and ± 15, etc. In retrospect this is clear, for two reasons.

First, this discrete sum is exactly the sum in Eq. (56) (after canceling the N's) with $dz_n = 0.2$, which corresponds to $N = 5$. So it's no surprise that we end up reproducing the original "stretched" periodic function with $N = 5$. Second, the lowest-frequency sine function in the discrete sum is the $z = 0.2$ one, which is $\sin\big(2\pi(0.2)x/L\big) = \sin(2\pi x/5L)$. This has a period of $5L$. And the periods of all the other sine functions are smaller than this; they take the form of $5L/m$, where m is an integer. So the sum of all the sines in the discrete sum will certainly repeat after $5L$ (and not any sooner).

This argument makes it clear that no matter how we try to approximate the Fourier-transform integral in Eq. (57) by a discrete sum, it will inevitably eventually repeat. Even if we pick the bin size to be a very small number like $\Delta z = 0.001$, the function will still repeat at some point (at $x = \pm 1000L$ in this case). We need a *truly* continuous integral in order for the function to never repeat. We'll talk more about these approximation issues when we look at the Gaussian function in Section 3.4.1. ♣

(b) Eq. (43) gives

$$\begin{aligned}
C(k) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ikx}dx \\
&= \frac{1}{2\pi}\int_{-L/2}^{0}(-A)e^{-ikx}dx + \frac{1}{2\pi}\int_{0}^{L/2}(A)e^{-ikx}dx \\
&= \frac{1}{2\pi}\cdot\frac{A}{(-ik)}\left(-e^{-ikx}\Big|_{-L/2}^{0} + e^{-ikx}\Big|_{0}^{L/2}\right) \\
&= \frac{iA}{2\pi k}\Big(-(1-e^{ikL/2}) + (e^{-ikL/2}-1)\Big) \\
&= \frac{-iA}{\pi k}\left(1-\cos\left(\frac{kL}{2}\right)\right).
\end{aligned} \tag{59}$$

So the expression for $f(x)$ in Eq. (43) becomes

$$\begin{aligned}
f(x) &= \int_{-\infty}^{\infty} C(k)e^{ikx}dk \\
&= \frac{-iA}{\pi}\int_{-\infty}^{\infty}\frac{1}{k}\left(1-\cos\left(\frac{kL}{2}\right)\right)e^{ikx}dk.
\end{aligned} \tag{60}$$

This looks different from the expression for $f(x)$ in Eq. (57), but it had better be the same thing, just in a different form. We can rewrite Eq. (60) as follows. Since the $1/k$ factor is an odd function of k, and since $\cos(kL/2)$ is even, only the odd part of e^{ikx}

(the $\sin(kx)$ part) survives in the integral. The integrand is then an even function of k, so we can have the integral run from 0 to ∞ and then multiply by 2. We obtain

$$\begin{aligned} f(x) &= \frac{-iA}{\pi} \cdot 2 \int_0^\infty \frac{1}{k}\left(1-\cos\left(\frac{kL}{2}\right)\right)\big(i\sin(kx)\big)\,dk \\ &= \frac{2A}{\pi}\int_0^\infty \frac{1}{k}\left(1-\cos\left(\frac{kL}{2}\right)\right)\sin(kx)\,dk \qquad (61) \end{aligned}$$

This still looks a little different from the expression for $f(x)$ in Eq. (57), but if we define $k \equiv 2\pi z/L \implies dk = (2\pi/L)dz$, then Eq. (61) turns into Eq. (57), as desired. So both forms generate the same function $f(x)$, as they must.

Remark: The "z" notation in Eq. (57) has the advantage of matching up with the "n" notation in the Fourier series in Eq. (58). But the "k" notation in Eqs. (60) and (61) is more widely used because it doesn't have all the π's floating around. The "n" notation in the Fourier series in Eq. (58) was reasonable to use, because it made sense to count the number of oscillations that fit into one period of the function. But when dealing with the Fourier transform of a non-periodic function, there is no natural length scale of the function, so it doesn't make sense to count the number of oscillations, so in turn there is no need for the π's. ♣

3.4 Special functions

The are a number of functions whose Fourier transforms come up often in the study of waves. Let's look at a few of the more common ones.

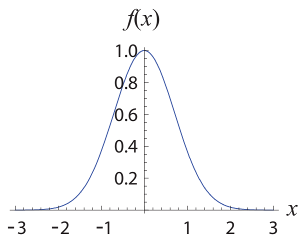


Figure 12

3.4.1 Gaussian

What is the Fourier transform of the Gaussian function, $f(x) = Ae^{-ax^2}$, shown in Fig. 12? (For the purposes of plotting a definite function, we have chosen $A = a = 1$.) If we plug this $f(x)$ into the second equation in Eq. (43), we obtain

$$C(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} Ae^{-ax^2}e^{-ikx}\,dx. \qquad (62)$$

In order to calculate this integral, we'll need to complete the square in the exponent. This gives

$$C(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} Ae^{-a(x+ik/2a)^2}e^{-k^2/4a}\,dx. \qquad (63)$$

It turns out that we can just ignore the constant (as far as x goes) term, $ik/2a$, in the exponent. This follows from a theorem in complex analysis involving the poles (divergences) of a function, but we won't get into that here. With the $ik/2a$ removed, we now have

$$C(k) = \frac{e^{-k^2/4a}}{2\pi}\int_{-\infty}^{\infty} Ae^{-ax^2}\,dx. \qquad (64)$$

Although a closed-form expression for the *indefinite* integral $\int e^{-ax^2}dx$ doesn't exist, it is fortunately possible to calculate the *definite* integral, $\int_{-\infty}^{\infty} e^{-ax^2}dx$. We can do this by using a slick trick, as follows. Let the desired integral be labeled as $I \equiv \int_{\infty}^{\infty} e^{-ax^2} dx$. Then we have (we'll explain the steps below)

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} e^{-ax^2} dx\right)\left(\int_{-\infty}^{\infty} e^{-ay^2} dy\right) \\ &= \int_{-\infty}^{\infty}\int_{\infty}^{\infty} e^{-ax^2} e^{-ay^2} dx\, dy \\ &= \int_0^{2\pi}\int_0^{\infty} e^{-ar^2} r\, dr\, d\theta \\ &= 2\pi \int_0^{\infty} re^{-ar^2} dr \\ &= -2\pi \frac{e^{-ar^2}}{2a}\bigg|_0^{\infty} \\ &= \frac{\pi}{a}. \end{aligned} \tag{65}$$

The second line follows from the fact that if you evaluate the double integral by doing the dx integral first, and then the dy one, you'll simply obtain the product of the two integrals in the first line. And the third line arises from changing the Cartesian coordinates to polar coordinates. (Note that if the limits in the original integral weren't $\pm\infty$, then the upper limit on r would depend on θ, and then the θ integral wouldn't be doable.)

Taking the square root of the result in Eq. (65) gives the nice result for the definite integral,[3]

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}. \tag{66}$$

As a check on this, if a is small, then the integral is large, which makes sense. Plugging Eq. (66) into Eq. (64) give the desired Fourier transform,

$$C(k) = \frac{e^{-k^2/4a}}{2\pi} \cdot A\sqrt{\frac{\pi}{a}} \quad\Longrightarrow\quad \boxed{C(k) = \frac{Ae^{-k^2/4a}}{2\sqrt{\pi a}}} \tag{67}$$

Interestingly, this is a Gaussian function of k. So the Fourier transform of the Gaussian function of x, $f(x) = Ae^{-ax^2}$, is a Gaussian function of k, $C(k) = (A/2\sqrt{\pi a})e^{-k^2/4a}$. So in some sense, a Gaussian function is the "nicest" function, at least as far as Fourier transforms are concerned. Gaussians remain Gaussians. Note the inverse dependences of $C(k)$ and $f(x)$ on a. For small a, $f(x)$ is very wide, whereas $C(k)$ is very tall (because of the a in the denominator) and sharply peaked (because of the a in the denominator of the exponent).

We can double check the result for $C(k)$ in Eq. (67) by going in the other direction. That is, we can plug $C(k)$ into the first equation in Eq. (43) and check that the result is the original Guassian function, $f(x)$. Indeed,

$$\begin{aligned} f(x) &= \int_{-\infty}^{\infty} C(k)e^{ikx} dk \\ &= \int_{-\infty}^{\infty} \frac{Ae^{-k^2/4a}}{2\sqrt{\pi a}}\, e^{ikx} dk \end{aligned}$$

[3] As a bonus, you can efficiently calculate integrals of the form $\int_{-\infty}^{\infty} x^{2n} e^{-ax^2} dx$ by repeatedly differentiating this expression with respect to a.

$$
\begin{aligned}
&= \frac{A}{2\sqrt{\pi a}} \int_{-\infty}^{\infty} e^{-(k-2iax)^2/4a} e^{-ax^2} dk \\
&= \frac{Ae^{-ax^2}}{2\sqrt{\pi a}} \int_{-\infty}^{\infty} e^{-k^2/4a} dk \\
&= \frac{Ae^{-ax^2}}{2\sqrt{\pi a}} \sqrt{\frac{\pi}{1/4a}} \\
&= Ae^{-ax^2}, \qquad (68)
\end{aligned}
$$

as desired. In obtaining the fourth line above, we (justifiably) ignored the constant (as far as k goes) term, $-2iax$, in the exponent. In obtaining the fifth line, we used Eq. (66) with $1/4a$ in place of a.

Making approximations to $f(x) = \int_{-\infty}^{\infty} C(k)e^{ikx} dk$

As we discussed shortly after Eq. (43), the first relation in Eq. (43) tells us that the Fourier transform $C(k)$ indicates how much of the function $f(x)$ is made up of e^{ikx}. More precisely, it tells us that $C(k)\,dk$ of the function $f(x)$ is made up of e^{ikx} terms in the range from k to $k + dk$. Since $f(x)$ is even in the present Gaussian case, we know that only the $\cos kx$ part of e^{ikx} will be relevant, but we'll keep writing the whole e^{ikx} anyway.

As a rough test to see if the Gaussian function $C(k)$ in Eq. (67) does what it is supposed to do (that is, produce the function $f(x)$ when plugged into the integral $\int_{-\infty}^{\infty} C(k)e^{ikx} dk$), let's make some approximations to this integral. There are two basic things we can do.[4]

- We can approximate the integral by performing a sum over discrete values of k. That is, we can break up the continuous integral into discrete "bins." For example, if pick the "bin size" to be $\Delta k = 1$, then this corresponds to making the approximation where all k values in the range $-0.5 < k < 0.5$ have the value of $C(k)$ at $k = 0$. And all k values in the range $0.5 < k < 1.5$ have the value of $C(k)$ at $k = 1$. And so on. This is shown in Fig. 13, for the values $A = a = 1$.

- The second way we can approximate the integral in Eq. (43) is to keep the integral a continuous one, but have the limits of integration be $k = \pm\ell$ instead of $\pm\infty$. We can then gradually make ℓ larger until we obtain the $\ell \to \infty$ limit. For example, If $\ell = 1$, then this means that we're integrating only over the shaded region shown in Fig. 14.

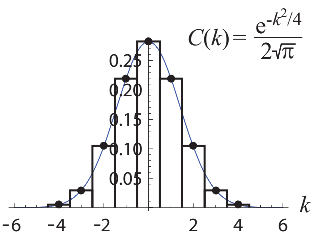


Figure 13

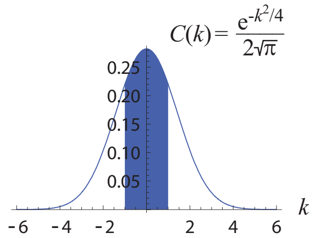
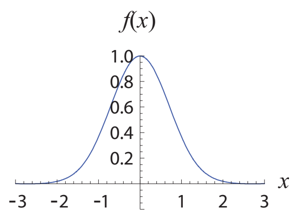


Figure 14

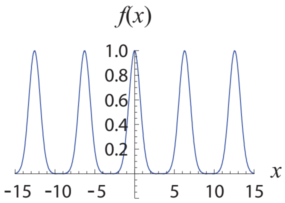
Let's see in detail what happens when we make these approximations. Consider the first kind, where we perform a sum over discrete values of k. With $A = a = 1$, we have $f(x) = e^{-x^2}$ and $C(k) = e^{-k^2/4}/2\sqrt{\pi}$. This function $C(k)$ is very small for $|k| \geq 4$, so we can truncate the sum at $k = \pm 4$ with negligible error. With $\Delta k = 1$, the discrete sum of the nine terms from $k = -4$ to $k = 4$ (corresponding to the nine bins shown in Fig. 13) yields an approximation to the $f(x)$ in Eq. (43) that is shown in Fig. 15. This provides a remarkably good approximation to the actual Gaussian function $f(x) = e^{-x^2}$, even though our discretization of k was a fairly coarse one. If we superimposed the true Gaussian on this plot, it would be hard to tell that it was actually a different curve. This is due to the high level of smoothness of the Gaussian function. For other general functions, the agreement invariably isn't this good. And even for the Gaussian, if we had picked a larger bin size of say, 2, the resulting $f(x)$ would have been noticeably non-Gaussian.



Sum from k=-4 to k=4 in steps of Δk=1

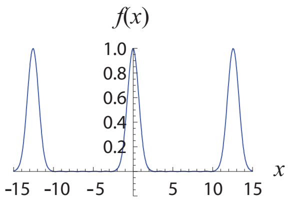
Figure 15

[4]Although the following discussion is formulated in terms of the Gaussian function, you can of course apply it to any function. You are encouraged to do so for some of the other functions discussed in this chapter.



Sum from k=-4 to k=4
in steps of Δk=1

Figure 16



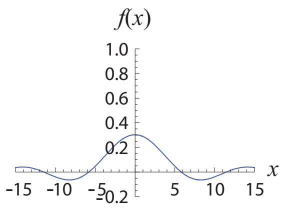
Sum from k=-4 to k=4
in steps of Δk=0.5

Figure 17

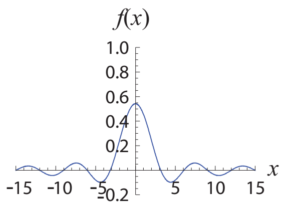
However, all is not as well as it seems. If we zoom out and let the plot of $f(x)$ run from, say, $x = -15$ to $x = 15$, then we obtain the curve shown in Fig. 16. This is hardly a Gaussian function. It is an (infinite) periodic string of nearly Gaussian functions. In retrospect, this should be no surprise. As we saw above in Fig. 11 and in the accompanying remark, the sum over discrete values of k will inevitably be periodic, because there is a lowest-frequency e^{ikx} component, and all other components have frequencies that are multiples of this. In the present scenario, the $k = 1$ component has the lowest frequency (call is k_{min}), and the period is $2\pi/k_{\text{min}} = 2\pi$. And indeed, the bumps in Fig. 16 occur at multiples of 2π.

What if we make the bin size smaller? Fig. 17 shows the case where k again runs from -4 to 4, but now with a bin size of $\Delta k = 0.5$. So there are 17 terms in the sum, with each term being multiplied by the $\Delta k = 0.5$ weighting factor. The smallest (nonzero) value of k is $k_{\text{min}} = 0.5$. We obtain basically the same shape for the bumps (they're slightly improved, although there wasn't much room for improvement over the $\Delta k = 1$ case), but the period is now $2\pi/k_{\text{min}} = 12\pi$. So the bumps are spread out more. This plot is therefore better than the one in Fig. 16, in the sense that it looks like a Gaussian for a larger range of x. If we kept decreasing the bin size Δk (while still letting the bins run from $k = -4$ to $k = 4$), the bumps would spread out farther and farther, until finally in the $\Delta k \to 0$ limit they would be infinitely far apart. In other words, there would be no repetition at all, and we would have an exact Gaussian function.

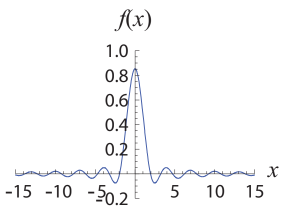
Consider now the second of the above approximation strategies, where we truncate the (continuous) integral at $k = \pm\ell$. If we let $\ell = 0.5$, then we obtain the first plot in Fig. 18. This doesn't look much like a Gaussian. In particular, it dips below zero, and the value at $x = 0$ isn't anywhere near the desired value of 1. This is because we simply haven't integrated over enough of the $C(k)$ curve. In the limit of very small ℓ, the resulting $f(x)$ is very small, because we have included only a very small part of the entire integral.



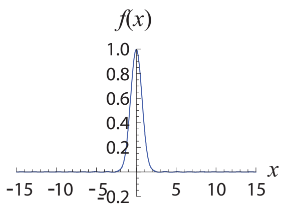
Integral from k=-0.5 to k=0.5



Integral from k=-1 to k=1



Integral from k=-2 to k=2



Integral from k=-4 to k=4

Figure 18

We can improve things by increasing ℓ to 1. The result is shown in the second plot in Fig. 18. The third plot corresponds to $\ell = 2$, and finally the fourth plot corresponds to $\ell = 4$. This looks very much like a Gaussian. Hardly any of the $C(k)$ curve lies outside $k = \pm 4$, so $\ell = 4$ provides a very good approximation to the full $\ell = \infty$ integral.

Compared with the first approximation method involving discrete sums, the plots generated with this second method have the disadvantage of not looking much like the desired

Gaussian for small values of ℓ. But they have the advantage of not being periodic (because there is no smallest value of k, since k is a now continuous variable). So, what you see is what you get. You don't have to worry about the function repeating at some later point.

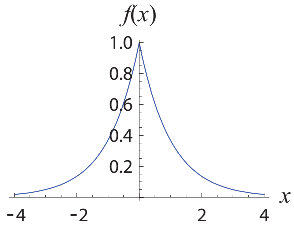


Figure 19

3.4.2 Exponential, Lorentzian

What is the Fourier transform of the exponential function, $Ae^{-b|x|}$, shown in Fig. 19? (We have chosen $A = b = 1$ in the figure.) If we plug this $f(x)$ into Eq. (43), we obtain

$$\begin{aligned} C(k) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} Ae^{-b|x|}e^{-ikx}dx \\ &= \frac{1}{2\pi}\int_{-\infty}^{0} Ae^{bx}e^{-ikx}dx + \frac{1}{2\pi}\int_{0}^{\infty} Ae^{-bx}e^{-ikx}dx. \end{aligned} \tag{69}$$

The imaginary terms in the exponents don't change the way these integrals are done (the fundamental theorem of calculus still holds), so we obtain

$$\begin{aligned} C(k) &= \frac{Ae^{(b-ik)x}}{2\pi(b-ik)}\bigg|_{-\infty}^{0} + \frac{Ae^{(-b-ik)x}}{2\pi(-b-ik)}\bigg|_{0}^{\infty} \\ &= \frac{A}{2\pi}\left(\frac{1}{b-ik} + \frac{1}{b+ik}\right) \\ &= \boxed{\frac{Ab}{\pi(b^2+k^2)}} \end{aligned} \tag{70}$$

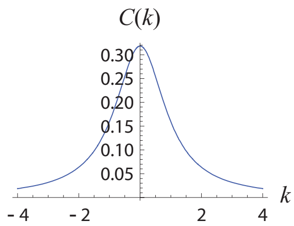


Figure 20

A function of the general form $c_1/(c_2 + c_3k^2)$ is called a *Lorentzian* function of k. From the plot of $C(k)$ shown in Fig. 20 (with $A = b = 1$), we see that a Lorentzian looks somewhat like a Gaussian. (The main difference is that it goes to zero like $1/k^2$ instead of exponentially.) This isn't too much of a surprise, because the exponential function in Fig. 19 looks vaguely similar to the Gaussian function in Fig. 12, in that they both are peaked at $x = 0$ and then taper off to zero.

As in the case of the Gaussian, the $f(x)$ and $C(k)$ functions here have inverse dependences on b. For small b, $f(x)$ is very wide, whereas $C(k)$ is sharply peaked and very tall. $C(k)$ is very tall because the value at $k = 0$ is $A/\pi b$, which is large if b is small. And it is sharply peaked in a relative sense because it decreases to half-max at $k = b$, which is small if b is small. Furthermore, it is sharply peaked in an absolute sense because for any given value of k, the value of $C(k)$ is less than $Ab/\pi k^2$, which goes to zero as $b \to 0$.[5]

If you want to go in the other direction, it is possible to plug the Lorentzian $C(k)$ into the first equation in Eq. (43) and show that the result is in fact $f(x)$. However, the integral requires doing a contour integration, so we'll just accept here that it does in fact work out.

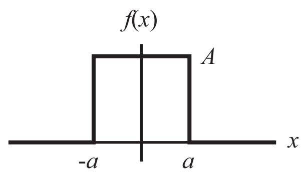


Figure 21

3.4.3 Square wave, sinc

What is the Fourier transform of the "square wave" function shown in Fig. 21? This square wave is defined by $f(x) = A$ for $-a \le x \le a$, and $f(x) = 0$ otherwise. If we plug this $f(x)$ into Eq. (43), we obtain

$$C(k) = \frac{1}{2\pi}\int_{-a}^{a} Ae^{-ikx}dx$$

[5]More precisely, for a given k, if you imagine decreasing b, starting with a value larger than k, then $C(k)$ increases to a maximum value of $A/2\pi k$ at $b = k$, and then decreases to zero as $b \to 0$.

$$= \frac{1}{2\pi} \cdot \frac{Ae^{-ikx}}{-ik}\Big|_{-a}^{a}$$
$$= \frac{A}{2\pi} \cdot \frac{e^{-ika} - e^{ika}}{-ik}$$
$$= \boxed{\frac{A\sin(ka)}{\pi k}} \qquad (71)$$

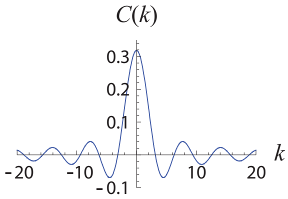


Figure 22

This is called a *sinc* function of k. $\text{sinc}(z)$ is defined generically to be $\sin(z)/z$. A plot of $C(k)$ is shown in Fig. 22 (with $A = a = 1$). It looks vaguely similar to a Gaussian and a Lorentzian near $k = 0$, but then for larger values of k it oscillates around zero, unlike the Gaussian and Lorentzian curves.

As in the Lorentzian case, it is possible to go in the other direction and plug $C(k)$ into the first equation in Eq. (43) and show that the result is $f(x)$. But again it requires a contour integration, so we'll just accept that it works out.

The dependence of the $C(k)$ sinc function on a isn't as obvious as with the Gaussian and Lorentzian functions. But using $\sin\epsilon \approx \epsilon$ for small ϵ, we see that $C(k) \approx Aa/\pi$ for k values near zero. So if a is large (that is, $f(x)$ is very wide), then $C(k)$ is very tall near $k = 0$. This is the same behavior that the Gaussian and Lorentzian functions exhibit.

However, the "sharply peaked" characteristic seems to be missing from $C(k)$. Fig. 23 shows plots of $C(k)$ for the reasonably large a values of 10 and 40. We're still taking A to be 1, and we've chosen to cut off the vertical axis at 2 and not show the peak value of Aa/π at $k = 0$. The envelope of the oscillatory sine curve is simply the function $A/\pi k$, and this is independent of a. As a increases, the oscillations become quicker, but the envelope doesn't fall off any faster. However, for the purposes of the Fourier transform, the important property of $C(k)$ is its integral, and if $C(k)$ oscillates very quickly, then it essentially averages out to zero. So for all practical purposes, $C(k)$ is basically the zero function except right near the origin. So in that sense it is sharply peaked.

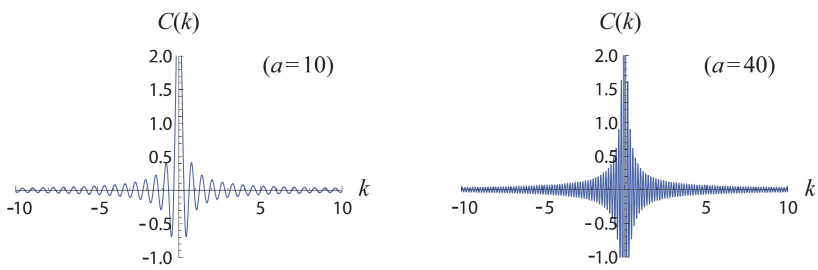


Figure 23

Remark: The wild oscillatory behavior of the $C(k)$ sinc function in Fig. 23, which is absent in the $C(k)$ Gaussian and Lorentzian functions, is due to the discontinuity in the $f(x)$ square wave. A more complicated combination of e^{ikx} functions are required to generate the discontinuity, especially if the function stays at a constant value of A for a very large interval before abruptly dropping to zero.

So, the lower degree of "niceness" of the sinc function, compared with the Gaussian and Lorentzian functions, can be traced to the discontinuity in the *value* of the $f(x)$ square wave. Similarly, a Lorentzian isn't quite as nice as a Gaussian. It may be a matter of opinion, but the Lorentzian in Fig. 20 isn't quite as smooth as the Gaussian in Fig. 12; it's a little less rounded. This lower degree of niceness of the Lorentzian can be traced to the discontinuity in the *first derivative* of the $f(x) = Ae^{-b|x|}$ exponential function. In a sense, the Gaussian function is the nicest and smoothest of all functions, and this is why its Fourier transform ends up being as nice and as smooth as possible. In other words, it is another Gaussian function. ♣

3.5 The delta function

3.5.1 Definition

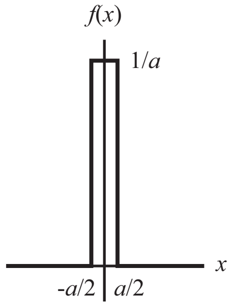


Figure 24

Consider the tall and thin square-wave function shown in Fig. 24. The height and width are related so that the area under the "curve" is 1. In the $a \to 0$ limit, this function has an interesting combination of properties: It is zero everywhere except at the origin, but it still has area 1 because the area is independent of a. A function with these properties is called a *delta function* and is denoted by $\delta(x)$. (It is sometimes called a "Dirac" delta function, to distinguish it from the "Kronecker" delta, δ_{nm}, that came up in Sections 3.1 and 3.2.) So we have the following definition of a delta function:

- *A delta function, $\delta(x)$, is a function that is zero everywhere except at the origin and has area 1.*

Strictly speaking, it's unclear what sense it makes to talk about the area under a curve, if the curve is nonzero at only one point. But there's no need to think too hard about it, because you can avoid the issue by simply thinking of a delta function as a square wave or a Gaussian (or various other functions, as we'll see below) in the sharply-peaked limit.

Consider what happens when we integrate the product of a delta function $\delta(x)$ with some other arbitrary function $f(x)$. We have (we'll explain the steps below)

$$\int_{-\infty}^{\infty} \delta(x) f(x)\, dx = \int_{-\infty}^{\infty} \delta(x) f(0)\, dx = f(0) \int_{-\infty}^{\infty} \delta(x)\, dx = f(0). \tag{72}$$

The first equality comes from the fact that the delta function is zero everywhere except at the origin, so replacing $f(x)$ with $f(0)$ everywhere else has no effect on the integral; the only relevant value of $f(x)$ is the value at the origin. The second equality comes from pulling the constant value, $f(0)$, outside the integral. And the third equality comes from the fact that the area under the delta function "curve" is 1. More generally, we have

$$\int_{-\infty}^{\infty} \delta(x-x_0) f(x)\, dx = \int_{-\infty}^{\infty} \delta(x-x_0) f(x_0)\, dx = f(x_0) \int_{-\infty}^{\infty} \delta(x-x_0)\, dx = f(x_0). \tag{73}$$

The reasoning is the same as in Eq. (72), except that now the only value of $f(x)$ that matters is $f(x_0)$, because the delta function is zero everywhere except at $x = x_0$ (because this is where the argument of the delta function is zero). This result is interesting. It says that a delta function can be used to "pick out" the value of a function at any point $x = x_0$, by simply integrating the product $\delta(x - x_0)f(x)$.

A note on terminology: A delta function is technically not a function, because there is no way for a function that is nonzero at only one point to have an area of 1. A delta function is actually a *distribution*, which is something that is defined through its integral when multiplied by another function. So a more proper definition of a delta function (we'll still call it that, since that's the accepted convention) is the relation in Eq. (72) (or equivalently Eq. (73)):

- *A delta function, $\delta(x)$, is a distribution that satisfies*

$$\int_{-\infty}^{\infty} \delta(x) f(x)\, dx = f(0). \tag{74}$$

However, having said this, the most convenient way to visualize a delta function is usually to think of it as a very tall and very thin function whose area is 1. This function could be the thin rectangle in Fig. 24, but there are infinitely many other functions that work just as well. For example, we can have a tall and thin triangle with base $2a$ and height $1/a$. Or a tall and thin Gaussian whose parameters are related in a certain way (as we'll see below). And so on.

As an exercise, you can verify the following properties of the delta function:

$$\begin{aligned}
\delta(x) &= \delta(-x) \\
\delta(ax) &= \frac{\delta(x)}{|a|} \\
\delta(g(x)) &= \frac{\delta(x-x_0)}{|g'(x_0)|} \qquad \text{if} \quad g(x_0)=0 \\
\int \delta'(x) f(x)\,dx &= -f'(0).
\end{aligned} \tag{75}$$

The second property follows from a change of variables in Eq. (74). The third follows from expanding $g(x)$ in a Taylor series around x_0 and then using the second property. The fourth follows from integration by parts. In the third property, if $g(x)$ has multiple zeros, then $\delta(g(x))$ consists of a delta function at each zero, which means that $\delta(g(x)) = \sum \delta(x-x_i)/|g'(x_i)|$.

3.5.2 The Fourier transform of $f(x) = 1$

It turns out that the delta function plays an important role with regard to Fourier transforms, in that it is the Fourier transform of one of the simplest functions of all, namely the constant function, $f(x) = 1$. Consistent with the general "inverse shape" behavior of $f(x)$ and $C(k)$ that we observed in the examples in Section 3.4, the function $f(x) = 1$ is infinitely spread out, and the delta function is infinitely sharply peaked.

To calculate the Fourier transform of $f(x) = 1$, it seems like all we have to do is plug $f(x) = 1$ into the second equation in Eq. (43). This gives

$$C(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} (1) e^{-ikx} dx. \tag{76}$$

This is an interesting integral. If $k = 0$, then we have $\int_{-\infty}^{\infty}(1)(1)\,dx$, which is infinite. And if $k \neq 0$, then the integrand oscillates indefinitely in both the positive and negative directions, so it's unclear what the integral is. If we use large but finite limits, and if e^{-ikx} undergoes an integral number of cycles within these limits, then the integral is zero. But if there is a "leftover" partial cycle, then the integral is nonzero. The integral therefore depends on the limits of integration, and hence is not well defined. Therefore, our first task in calculating the integral is to make sense of it. We can do this as follows.

Instead of having the function be identically 1 for all x, let's have it be essentially equal to 1 for a very large interval and then slowly taper off to zero for very large x. If it tapers off to zero on a length scale of ℓ, then after we calculate the integral in terms of ℓ, we can take the $\ell \to \infty$ limit. In this limit the function equals the original constant function, $f(x) = 1$. There are (infinitely) many ways to make the function taper off to zero. We'll consider a few of these below, and in the process we'll generate various representations of the delta function. We'll take advantage of the results for the special functions we derived in Section 3.4.

3.5.3 Representations of the delta function

Gaussian

A standard way to make the $f(x) = 1$ function taper off is to use the Gaussian function, e^{-ax^2}, instead of the constant function, 1. If a is very small, then this Gaussian function is essentially equal to 1 for a large range of x values, and then it eventually tapers off to zero for large x. It goes to zero on a length scale of $\ell \sim 1/\sqrt{a}$. This is true because the function is very close to 1 for, say, $x = 1/(10\sqrt{a})$, where it equals $e^{-1/100} \approx 1 - 1/100$. And it is essentially equal to zero for, say, $x = 5/\sqrt{a}$, where it equals $e^{-25} \approx 0$. In the end, we will take the $a \to 0$ limit, which will make the Gaussian function be essentially equal to 1 for all x. The point is that if we jumped right to the $a = 0$ case, there would be no way to get a handle on the integral. But if we work with a finite value of a, we can actually calculate the integral, and *then* we can take the $a \to 0$ limit.

We calculated the Fourier transform of the Gaussian function $f(x) = Ae^{-ax^2}$ in Eq. (67). With $A = 1$ here, the result is the Gaussian function of k,

$$C(k) = \frac{e^{-k^2/4a}}{2\sqrt{\pi a}}\,. \tag{77}$$

In the $a \to 0$ limit, the a in the exponent makes this Gaussian fall off to zero very quickly (it goes to zero on a length scale of $2\sqrt{a}$). And the a in the denominator makes the value at $k = 0$ very large. The Gaussian is therefore sharply peaked around $k = 0$. So the Fourier transform of a very wide Gaussian is a sharply peaked one, as shown in Fig. 25, for $a = 0.02$. This makes sense, because if $f(x)$ is very wide, then it is made up of only very slowly varying exponential/sinusoidal functions. In other words, $C(k)$ is dominated by k values near zero.

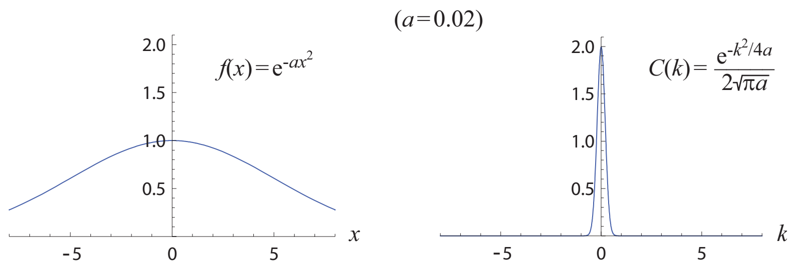


Figure 25

What is the area under the $C(k)$ curve? Using the $\int_{-\infty}^{\infty} e^{-ax^2}\,dx = \sqrt{\pi/a}$ result from Eq. (66), we have

$$\begin{aligned} \int_{-\infty}^{\infty} C(k)\,dk &= \int_{-\infty}^{\infty} \frac{e^{-k^2/4a}}{2\sqrt{\pi a}}\,dk \\ &= \frac{1}{2\sqrt{\pi a}} \cdot \sqrt{\frac{\pi}{1/4a}} \\ &= 1. \end{aligned} \tag{78}$$

So the area under the Fourier transform of the Gaussian function $f(x) = e^{-ax^2}$ equals 1. This is a nice result. And it is independent of a. Even if a is very small, so that the Fourier transform $C(k)$ is very sharply peaked, the area is still 1. We have therefore found the following representation of the delta function:

$$\boxed{\delta(k) = \lim_{a \to 0} \frac{e^{-k^2/4a}}{2\sqrt{\pi a}}} \tag{79}$$

Remarks:

1. We should say that the above process of finding the Fourier transform of the constant function $f(x) = 1$ was by no means necessary for obtaining this representation of the delta function. Even if we just pulled Eq. (79) out of a hat, it would still be a representation, because it has the two required properties of being infinitely sharply peaked and having area 1. But we wanted to motivate it by considering it to be the Fourier transform of the constant function, $f(x) = 1$.
2. It is no coincidence that the integral in Eq. (78) equals 1. It is a consequence of the fact that the value of the $f(x)$ Gaussian function equals 1 at $x = 0$. To see why, let's plug the $C(k)$ from Eq. (77) into the first equation in Eq. (43). And let's look at what this equation says when x takes on the particular value of $x = 0$. We obtain

$$f(0) = \int_{-\infty}^{\infty} \frac{e^{-k^2/4a}}{2\sqrt{\pi a}} e^{ik(0)}\, dk \implies 1 = \int_{-\infty}^{\infty} \frac{e^{-k^2/4a}}{2\sqrt{\pi a}}\, dk, \tag{80}$$

which is just what Eq. (78) says. This is a general result: if $C(k)$ is the Fourier transform of $f(x)$, then the area under the $C(k)$ curve equals $f(0)$. ♣

Exponential, Lorentzian

Another way to make the $f(x) = 1$ function taper off is to use the exponential function, $e^{-b|x|}$. If b is very small, then this exponential function is essentially equal to 1 for a large range of x values, and then it eventually tapers off to zero for large x. It goes to zero on a length scale of $1/b$.

We calculated the Fourier transform of the exponential function $Ae^{-b|x|}$ in Eq. (70). With $A = 1$ here, the result is the Lorentzian function of k,

$$C(k) = \frac{b}{\pi(b^2 + k^2)}\,. \tag{81}$$

In the $b \to 0$ limit, this function is very tall, because the value at $k = 0$ is $1/\pi b$. And it is very narrow, because as we mentioned in Section 3.4.2, for any given value of k, the value of $C(k)$ is less than $b/\pi k^2$, which goes to zero as $b \to 0$. The Fourier transform of a wide exponential is therefore a sharply peaked Lorentzian, as shown in Fig. 26.

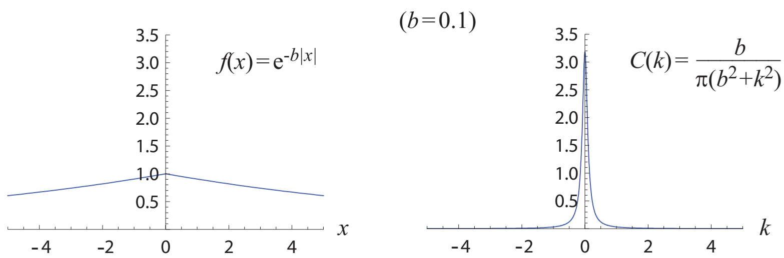


Figure 26

What is the area under the Lorentzian curve? From the above remark containing Eq. (80), we know that the area must equal 1, but let's check this by explicitly evaluating the integral. Using $\int_{-\infty}^{\infty} b\,dk/(b^2+k^2) = \tan^{-1}(k/b)$, we have

$$\int_{-\infty}^{\infty} \frac{b\,dk}{\pi(b^2+k^2)} = \frac{1}{\pi}\Big(\tan^{-1}(\infty) - \tan^{-1}(\infty)\Big) = \frac{1}{\pi}\Big(\frac{\pi}{2} - \frac{-\pi}{2}\Big) = 1. \tag{82}$$

Since the area is 1, and since the curve is sharply peaked in the $b \to 0$ limit, the Lorentzian function in Eq. (81) therefore gives another representation of the delta function:

$$\boxed{\delta(k) = \lim_{b\to 0} \frac{b}{\pi(b^2+k^2)}} \tag{83}$$

This makes sense, because the $b \to 0$ limit of the exponential function $e^{-b|x|}$ is essentially the same function as the $a \to 0$ limit of the Gaussian function a^{-ax^2} (they both have $f(x) = 1$ as the limit), and we found above in Eq. (79) that the Fourier transform of a Gaussian is a delta function in the $a \to 0$ limit.

We can look at things the other way, too. We can approximate the $f(x) = 1$ function by a very wide Lorentzian, and then take the Fourier transform of that. The Lorentzian function that approximates $f(x) = 1$ is the $b \to \infty$ limit of $f(x) = b^2/(b^2+x^2)$, because this falls off from 1 when x is of order b. Assuming (correctly) that the Fourier transform process is invertible, we know that the Fourier transform of $f(x) = b^2/(b^2+x^2)$ must be an exponential function. The only question is what the factors are. We could do a contour integration if we wanted to produce the Fourier transform from scratch, but let's do it the easy way, as follows.

The known Fourier-transform relation between $f(x) = e^{-b|x|}$ and $C(k) = b/\pi(b^2+k^2)$, which we derived in Eqs. (69) and (70), is given by Eq. (43) as

$$e^{-b|x|} = \int_{-\infty}^{\infty} \frac{b}{\pi(b^2+k^2)}\, e^{ikx} dk, \qquad \text{where} \qquad \frac{b}{\pi(b^2+k^2)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-b|x|} e^{-ikx} dx. \tag{84}$$

Interchanging the x and k letters in the first of these equations, shifting some factors around, and changing e^{ikx} to e^{-ikx} (which doesn't affect things since the rest of the integrand is an even function of x) gives

$$(b/2)e^{-b|k|} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{b^2}{(b^2+x^2)}\, e^{-ikx} dx. \tag{85}$$

But this is the statement that the Fourier transform of $f(x) = b^2/(b^2+x^2)$ (which is essentially the constant function $f(x) = 1$ if $b \to \infty$) is $C(k) = (b/2)e^{-b|k|}$.[6]

This $C(k) = (b/2)e^{-b|k|}$ exponential function is infinitely sharply-peaked in the $b \to \infty$ limit. And you can quickly show that it has an area of 1 (but again, this follows from the above remark containing Eq. (80)). So we have found another representation of the delta function:

$$\boxed{\delta(k) = \lim_{b\to\infty} (b/2)e^{-b|k|}} \tag{86}$$

Just like with the wide Gaussian and exponential functions, the wide Lorentzian function is essentially equal to the $f(x) = 1$ function, so its Fourier transform should be roughly the same as the above ones. In other words, it should be (and is) a delta function.

[6]By "Fourier transform" here, we mean the transform in the direction that has the $1/2\pi$ in front of the integral, since that's what we set out to calculate in Eq. (76).

Square wave, sinc

Another way to make the $f(x) = 1$ function "taper off" is to use a square-wave function that takes on the constant value of 1 out to $x = \pm a$ and then abruptly drops to zero. We calculated the Fourier transform of the square wave in Eq. (71). With $A = 1$ here, the result is the "sinc" function of k,

$$C(k) = \frac{\sin(ka)}{\pi k} . \tag{87}$$

Using $\sin \epsilon \approx \epsilon$ for small ϵ, we see that $C(k) \approx a/\pi$ for k values near zero. So if a is large then $C(k)$ is very tall near $k = 0$. The Fourier transform of a very wide square wave (large a) is therefore a very tall sinc function, as shown in Fig. 27. This is the same behavior we saw with the Gaussian, exponential, and Lorentzian functions.

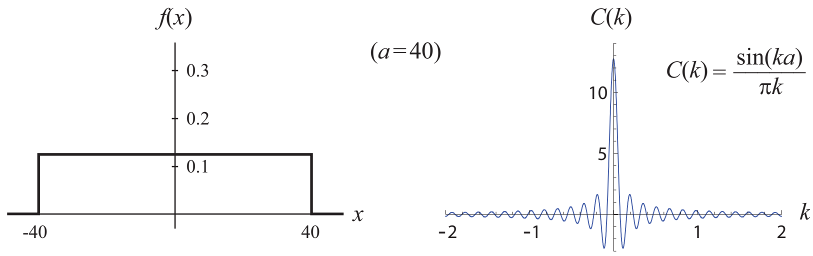


Figure 27

However, as we saw in Fig. 23 in Section 3.4.3, the "sharply peaked" characteristic seems to be missing. The envelope of the oscillatory curve is the function $1/\pi k$, and this is independent of a. As a increases, the oscillations become quicker, but the envelope doesn't fall off any faster. But when it comes to considering $C(k)$ to be a delta function, the important property is its integral. If $C(k)$ oscillates very quickly, then it essentially averages out to zero. So for all practical purposes, $C(k)$ is basically the zero function except right near the origin. So in that sense it is sharply peaked.

Furthermore, we know from the above remark containing Eq. (80) (or a contour integration would also do the trick) that the (signed) area under the $C(k)$ curve equals 1, independent of a. We therefore have yet another representation of the delta function:

$$\boxed{\delta(k) = \lim_{a\to\infty} \frac{\sin(ka)}{\pi k}} \tag{88}$$

We can look at things the other way, too. We can approximate the $f(x) = 1$ function by a very wide sinc function, and then take the Fourier transform of that. The sinc function that approximates $f(x) = 1$ is the $a \to 0$ limit of $\sin(ax)/(ax)$, because this equals 1 at $x = 0$, and it falls off from 1 when x is of order $1/a$. Assuming (correctly) that the Fourier-transform process is invertible, you can show that the above Fourier-transform relation between the square wave and $\sin(ka)/\pi k$ implies (with the 2π in the opposite place; see the step going from Eq. (84) to Eq. (85)) that the Fourier transform of $f(x) = \sin(ax)/(ax)$ is a square-wave function of k that has height $1/2a$ and goes from $-a$ to a. This is a very tall and thin rectangle in the $a \to 0$ limit. And the area is 1. So we have returned full circle to our original representation of the delta function in Fig. 24 (with the a in that figure now $2a$).

Again, just like with the other wide functions we discussed above, a wide sinc function is essentially equal to the $f(x) = 1$ function, so its Fourier transform should be roughly the same as the above ones. In other words, it should be (and is) a delta function.

Integral representation

There is one more representation of the delta function that we shouldn't forget to write down. We've actually been using it repeatedly in this section, so it's nothing new. As we've seen many times, the Fourier transform of the function $f(x) = 1$ is a delta function. So we should write this explicitly:

$$\boxed{\delta(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ikx}dx} \tag{89}$$

where we haven't bothered to write the "1" in the integrand. Of course, this integral doesn't quite make sense the way it is, so it's understood that whenever it comes up, we make the integrand taper off to zero by our method of choice (for example, any of the above methods), and then eventually take the limit where the tapering length scale becomes infinite. More generally, we have

$$\delta(k-k_0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i(k-k_0)x}dx. \tag{90}$$

Similar generalizations to a $k - k_0$ argument hold for all of the above representations too, of course.

Eq. (90) allows us to generalize the orthogonality relation in Eq. (22), which was

$$\int_0^L e^{i2\pi nx/L}e^{-i2\pi mx/L}dx = L\delta_{nm}. \tag{91}$$

In that case, the inner product was defined as the integral from 0 to L. But if we define a different inner product that involves the integral from $-\infty$ to ∞ (we are free to define it however we want), then Eq. (90) gives us a new orthogonality relation,

$$\int_{-\infty}^{\infty} e^{ik_1x}e^{-ik_2x}dx = 2\pi\delta(k_2-k_1). \tag{92}$$

(Alternatively, we could define the inner product to be $1/2\pi$ times the integral, in which case the 2π wouldn't appear on the righthand side.) So the e^{ikx} functions are still orthogonal; the inner product is zero unless the k values are equal, in which case the inner product is infinite. We see that whether we define the inner product as the integral from 0 to L, or from $-\infty$ to ∞, we end up with a delta function. But in the former case it is the standard "Kronecker" delta, whereas in latter case it is the "Dirac" delta function.

3.6 Gibbs phenomenon

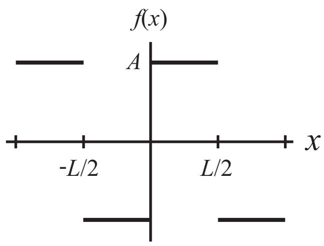


Figure 28

Let's return to the study of Fourier series, as opposed to Fourier transforms. The "Gibbs phenomenon" refers to the manner in which the Fourier series for a periodic function overshoots the values of the function on either side of a discontinuity. This effect is best illustrated (for many reasons, as we will see) by the periodic step function shown in Fig. 28. This step function is defined by

$$f(x) = \begin{cases} -A & (-L/2 < x < 0) \\ A & (0 < x < L/2), \end{cases} \tag{93}$$

and has period L. We already found the Fourier series for this function in Eq. (51). It is

$$\begin{aligned} f(x) &= \frac{4A}{\pi}\sum_{n\ \text{odd}} \frac{1}{n}\sin\left(\frac{2\pi nx}{L}\right) \\ &= \frac{4A}{\pi}\left(\sin\left(\frac{2\pi x}{L}\right) + \frac{1}{3}\sin\left(\frac{6\pi x}{L}\right) + \frac{1}{5}\sin\left(\frac{10\pi x}{L}\right) + \cdots\right). \end{aligned} \tag{94}$$

By "n odd" we mean the positive odd integers. If we plot a few partial sums (with $A = L = 1$), we obtain the plots shown in Fig. 29 (this is just a repeat of Fig. 7). The first plot includes only the first term in the series, while the last plot includes up to the 50th term (the $n = 99$ one).

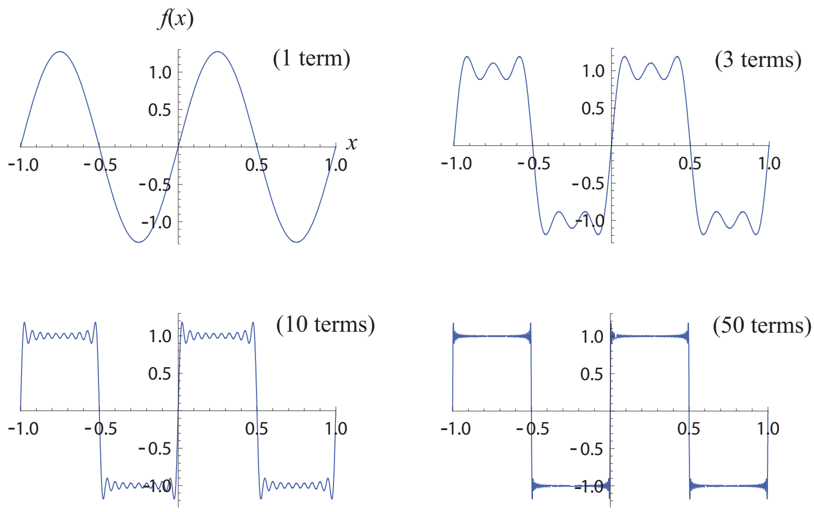


Figure 29

The more terms we include, the more the curve looks like a step function. However, there are two undesirable features. There are wiggles near the discontinuity, and there is an overshoot right next to the discontinuity. As the number of terms grows, the wiggles get pushed closer and closer to the discontinuity, in the sense that the amplitude in a given region decreases as the number of terms, N, in the partial series increases. So in some sense the wiggles go away as N approaches infinity. However, the overshoot unfortunately never goes away. Fig. 30 shows the zoomed-in pictures near the discontinuity for $N = 10$ and $N = 100$. If we included more terms, the overshoot would get pushed closer and closer to the discontinuity, but it would still always be there.

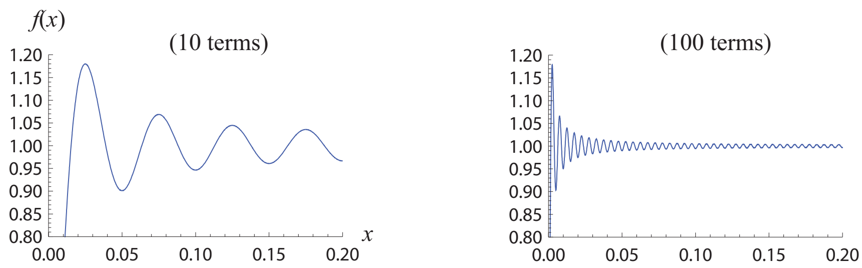


Figure 30

It turns out that as long as the number of terms in the partial series is large, the height of the overshoot is always about 9% of the jump in the function, independent of the (large) number N. This is consistent with the plots in Fig. 30, where the overshoot is about 0.18, which is 9% of the jump from -1 to 1. Furthermore, this 9% result holds for a discontinuity in *any* function, not just a step function. This general result is provable in a page or two, so let's do that here. Our strategy will be to first prove it for the above step function, and then show how the result for any other function follows from the specific result for the step function.

To demonstrate the 9% the result for the step function, our first task is to find the location of the maximum value of the overshoot. Let's assume that we're truncating the series in Eq. (94) at the Nth term (the $n = 2N - 1$ one), and then we'll take the $N \to \infty$ limit. To find the location of the maximum, we can take the derivative of the partial series (which we'll call $f_N(x)$) and set the result equal to zero. This gives

$$0 = \frac{df_N(x)}{dx} = \frac{8A}{L}\left(\cos\left(\frac{2\pi x}{L}\right) + \cos\left(\frac{6\pi x}{L}\right) + \cdots + \cos\left(\frac{2(2N-1)\pi x}{L}\right)\right). \tag{95}$$

The smallest solution to this equation is $x = L/4N$. This is a solution because it makes the arguments of the cosines in the first and last terms in Eq. (95) be supplementary, which means that the cosines cancel. And likewise for the second and second-to-last terms. And so on. So all of the terms cancel in pairs. (If N is odd, there is a middle term that is zero.) Furthermore, there is no (positive) solution that is smaller than this one, for the following reason.

For concreteness, let's take $N = 6$. If we start with a very small value of x, then the sum of the cosine terms in Eq. (95) (there are six in the $N = 6$ case) equals the sum of the horizontal components of the vectors in the first diagram in Fig. 31. (The vectors make angles $2\pi nx/L$ with the horizontal axis, where $n = 1, 3, 5, 7, 9, 11$.) This sum is clearly not zero, because the four vectors are "lopsided" in the positive direction.

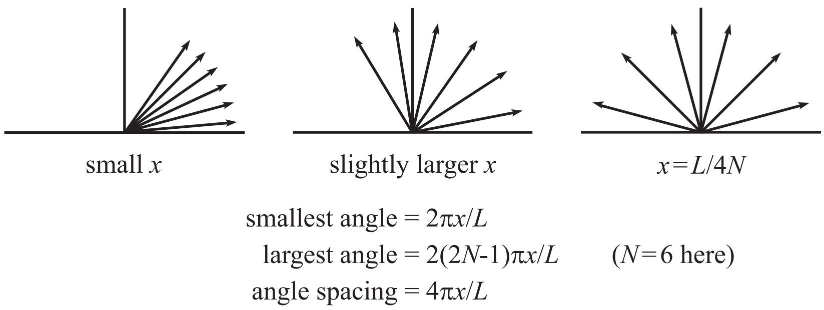


Figure 31

If we increase x a little, we get the six vectors in the second diagram in Fig. 31. The vectors are still lopsided in the positive direction, so the sum again isn't zero. But if we keep increasing x, we finally get to the situation shown in the third diagram in Fig. 31. The vectors are now symmetric with respect to the y axis, so the sum of the horizontal components is zero. This is the $x = L/4N$ case we found above, where the vectors are supplementary in pairs (setting the sum of the first and Nth angles equal to π gives $x = L/4N$). So $x = L/4N$ is indeed the smallest positive solution to Eq. (95) and is therefore the location of the overshoot peak in Fig. 30.

Having found the location of the maximum value of the overshoot, we can plug $x = L/4N$ into Eq. (94) to find this maximum value. We obtain

$$\begin{aligned} f_N\left(\frac{L}{4N}\right) &= \frac{4A}{\pi}\left(\sin\left(\frac{\pi}{2N}\right) + \frac{1}{3}\sin\left(\frac{3\pi}{2N}\right) + \cdots + \frac{1}{2N-1}\sin\left(\frac{(2N-1)\pi}{2N}\right)\right) \\ &= \frac{4A}{\pi}\sum_{m=1}^{N}\frac{1}{2m-1}\sin\left(\frac{(2m-1)\pi}{2N}\right). \end{aligned} \tag{96}$$

In the $N \to \infty$ limit, the arguments of the sines increase essentially continuously from 0 to π. Therefore, because of the $(2m-1)$ factor in the denominator, this sum is essentially the integral of $\sin(y)/y$ from 0 to π, except for some factors we need to get straight. If we multiply Eq. (96) through by 1 in the form of $(\pi/2N)(2N/\pi)$, we obtain

$$f_N\left(\frac{L}{4N}\right) = \frac{\pi}{2N}\cdot\frac{4A}{\pi}\sum_{m=1}^{N}\frac{2N}{(2m-1)\pi}\sin\left(\frac{(2m-1)\pi}{2N}\right). \tag{97}$$

Each term in this sum is weighted by a factor of 1, whereas in the integral $\int_0^\pi \big(\sin(y)/y\big)\,dy$ each term is weighted by dy. But dy is the difference between successive values of $(2m-1)\pi/2N$, which is $2\pi/2N$. So $dy = \pi/N$. The above sum is therefore N/π times the integral from 0 to π. So in the $N \to \infty$ limit we obtain

$$\begin{aligned} f_N\left(\frac{L}{4N}\right) &= \frac{\pi}{2N}\cdot\frac{4A}{\pi}\left(\frac{N}{\pi}\int_0^\pi \frac{\sin y}{y}\,dy\right) \\ &= \frac{2A}{\pi}\int_0^\pi \frac{\sin y}{y}\,dy. \end{aligned} \tag{98}$$

Alternatively, you can systematically convert the sum in Eq. (96) to the integral in Eq. (98) by defining

$$y \equiv \frac{(2m-1)\pi}{2N} \quad\Longrightarrow\quad dy = \frac{\pi\,dm}{N}. \tag{99}$$

If you multiply Eq. (96) by dm (which doesn't affect anything, since $dm = 1$) and then change variables from m to y, you will obtain Eq. (98).

The value of the actual step function just to the right of the origin is A, so to get the overshoot, we need to subtract off A from $f(L/4N)$. And then we need to divide by $2A$ (which is the total height of the jump) to obtain the fractional overshoot. The result is

$$\frac{1}{2A}\left(f\left(\frac{L}{4N}\right) - A\right) = \frac{1}{\pi}\int_0^\pi \frac{\sin y}{y}\,dy - \frac{1}{2} \approx 0.0895 \approx 9\%, \tag{100}$$

where we have calculated the integral numerically. For the simple step function, we have therefore demonstrated the desired 9% result.

As we mentioned above, this 9% result also holds for any discontinuity in any other function. Having obtained the result for the simple step function, the generalization to other arbitrary functions is surprisingly simple. It proceeds as follows.

Consider an arbitrary function $f(x)$ with period L and with a discontinuity, as shown in the first plot in Fig. 32. Without loss of generality, we can assume that the discontinuity in question (there may very well be others) is located at the origin. And we can also assume that the discontinuity is vertically symmetric around the origin, that is, it goes from $-A$ to A (or vice versa) for some value of A. This assumption is valid due to the fact that shifting the function vertically simply changes the a_0 value in Eq. (1), which doesn't affect the nature of the overshoot.

Now consider the periodic step function, $f_{\text{step}}(x)$, that jumps from $-A$ to A and also has period L, as shown in the second plot in Fig. 32. And finally consider the function $f_{\text{diff}}(x)$ defined to be the difference between $f(x)$ and $f_{\text{step}}(x)$, as shown in the third plot in Fig. 32. So we have

$$f(x) = f_{\text{step}}(x) + f_{\text{diff}}(x). \tag{101}$$

Since $f_{\text{diff}}(x) = f(x) - f_{\text{step}}(x)$, the plot of $f_{\text{diff}}(x)$ is obtained by simply subtracting or adding A from $f(x)$, depending on whether the step function is positive or negative at x, respectively. The critical point to realize is that by construction, $f_{\text{diff}}(x)$ is a *continuous*

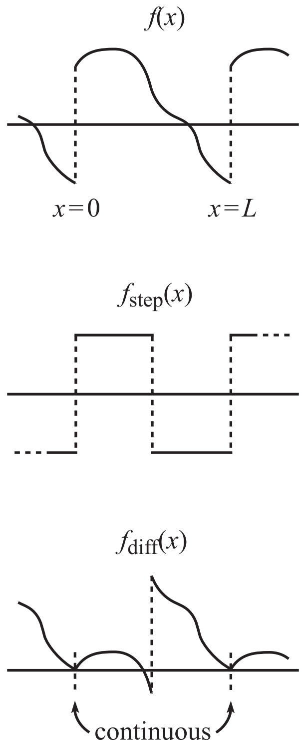


Figure 32

function at the discontinuity at $x = 0$ (and $x = L$, etc.). Its derivative might not be continuous (and in fact isn't, in the present case), but that won't matter here. Also, $f_{\text{diff}}(x)$ may very well have discontinuities elsewhere (in fact, we introduced one at $x = L/2$), but these aren't relevant as far as the discontinuity at the origin goes.

Since $f_{\text{diff}}(x)$ is continuous at $x = 0$, its Fourier series has no overshoot there (see the first remark below). The Fourier series for $f_{\text{diff}}(x)$ therefore contributes nothing to the overshoot of the Fourier series of $f(x)$ at $x = 0$. Hence, in view of Eq. (101), the overshoot in the Fourier series for $f(x)$ is exactly equal to the overshoot in the Fourier series for $f_{\text{step}}(x)$. The same 9% result therefore holds, as we wanted to show.

Remarks:

1. We claimed in the previous paragraph that since $f_{\text{diff}}(x)$ is continuous at $x = 0$, its Fourier series has no overshoot there. Said in another way, a kink in $f_{\text{diff}}(x)$ (that is, a discontinuity in its derivative) doesn't cause an overshoot. The can be shown with the same strategy that we used above, where we looked at a specific function (the step function) and then argued why any other function gave the same result.

   For the present task, we'll consider the specific triangular function shown in Fig. 35 below, which we'll discuss in Section 3.7. We'll find that the Fourier series for this triangular function is smooth at $x = 0$. There is no overshoot. Now, any function $f_{\text{diff}}(x)$ with a kink looks (at least locally near the kink) like a triangular function plus possibly a linear function of x, which has the effect of tilting the triangle so that the slopes on either side of the kink aren't equal and opposite. The triangular function has no overshoot, and the linear function certainly doesn't have one, either. So we conclude that $f_{\text{diff}}(x)$ doesn't have an overshoot, as we wanted to show.

2. The reasoning in this section can also be used to prove a quick corollary, namely, that the value of the Fourier series at a discontinuity equals the midpoint of the jump. This follows because it is true for the step function (the value at $x = 0$ is zero, because we have only sines in the series), and it is also true for $f_{\text{diff}}(x)$ (in a trivial sense, because $f_{\text{diff}}(x)$ is continuous and has no jump). Also, any a_0 constant we stripped off from the original function doesn't affect the result, because it increases the endpoints and midpoint of the jump by the same amount.

3. The overshoot (which occurs at $x = L/4N$) gets squeezed arbitrarily close to the discontinuity as N becomes large. So for any given nonzero value of x near the discontinuity, the overshoot is irrelevant as far as the value of $f(x)$ goes. Also, at any given nonzero value of x the amplitude of the wiggles in Fig. 30 goes to zero (we'll show this in the following subsection). Therefore, the Gibbs phenomenon is irrelevant for any given nonzero value of x, in the $N \to \infty$ limit. The only possibly value of x where we might have an issue is $x = 0$ (and $x = L$, etc.). From the previous remark, we know that the value of the series at $x = 0$ is the midpoint of the jump. This may or may not equal the arbitrary value we choose to assign to $f(0)$, but we can live with that. ♣

The shape of the wiggles

The strategy that we used above, where we turned the sum in Eq. (94) (or rather, the partial sum of the first N terms) into the integral in Eq. (98), actually works for any small value of x, not just the $x = L/4N$ location of the overshoot. We can therefore use this strategy to calculate the value of the Fourier series at *any* small value of x, in the large-N limit.[7] In other words, we can find the shape (and in particular the envelope) of the wiggles in Fig. 30.

[7] We need x to be small compared with L, so that the sine terms in Eq. (102) below vary essentially continuously as m varies. If this weren't the case, then we wouldn't be able to approximate the sum in Eq. (102) by an integral, which we will need to do.

The procedure follows the one above. First, write the partial sum of the first N terms in Eq. (94) as

$$f_N(x) = \frac{4A}{\pi}\sum_{m=1}^{N}\frac{1}{2m-1}\sin\Big(\frac{2(2m-1)\pi x}{L}\Big). \tag{102}$$

Then define

$$y \equiv \frac{2(2m-1)\pi x}{L} \quad\Longrightarrow\quad dy = \frac{4\pi x\, dm}{L}\,. \tag{103}$$

Then multiply Eq. (102) by dm (which doesn't affect anything, since $dm = 1$) and change variables from m to y. The result is (for large N)

$$f_N(x) \approx \frac{2A}{\pi}\int_0^{(4N-2)\pi x/L}\frac{\sin y}{y}\,dy. \tag{104}$$

As double check on this, if we let $x = L/4N$ then the upper limit of integration is equal to π (in the large-N limit), so this result reduces to the one in Eq. (98).

We were concerned with the $N \to \infty$ limit in the above derivation of the overshoot, but now we're concerned with just large (but not infinite) N. If we actually set $N = \infty$ in Eq. (104), we obtain $(2A/\pi)\int_0^\infty(\sin y)/y\cdot dy$. This had better be equal to A, because that is the value of the step function at x. And indeed it is, because this definite integral equals $\pi/2$ (see the paragraph preceding Eq. (88); it boils down to the fact that $(\sin y)/y$ is the Fourier transform of a square wave). But the point is that for the present purposes, we want to keep N finite, so that we can actually see the x dependence of $f_N(x)$. (So technically the integral approximation to the sum isn't perfect like it was in the $N = \infty$ case. But for large N it's good enough.)

To make the integral in Eq. (104) easier to deal with, let's define n via $x \equiv n(L/4N)$. The parameter n need not be an integer, but if it is, then it labels the nth local maximum or minimum of $f_N(x)$, due to the reasoning we used in Fig. 31. The number n gives the number of half revolutions the vectors in Fig. 31 make as they wrap around in the plane. If n is an integer, then the sum of the horizontal projections (the cosines in Eq. (95)) is zero, and we therefore have a local maximum or minimum of $f_N(x)$.

With this definition of n, the upper limit of the integral in Eq. (104) is essentially equal to $n\pi$ (assuming N is large), so we obtain

$$f_N(n) \approx \frac{2A}{\pi}\int_0^{n\pi}\frac{\sin y}{y}\,dy \qquad \text{(where } x \equiv n(L/4N)\text{)}. \tag{105}$$

Again, n need not be an integer. What does this function of n look like? The integral has to be computed numerically, and we obtain the first plot shown in Fig. 33 (we have chosen $A = 1$). The second plot is a zoomed-in (vertically) version of the first one. Note that there is actually no N dependence in $f_N(n)$. It is simply a function of n. Therefore, the height of the nth bump is independent of N, assuming that N is reasonably large.[8] The N (and L) dependence comes in when we convert from n back to x.

[8] If N isn't large, then we can't make the approximation that the upper limit of the integral in Eq. (105) equals $n\pi$. From Eq. (104), the limit actually equals $n\pi(1 - 1/2N)$. But if N isn't large, then the integral approximation to the sum isn't so great anyway.

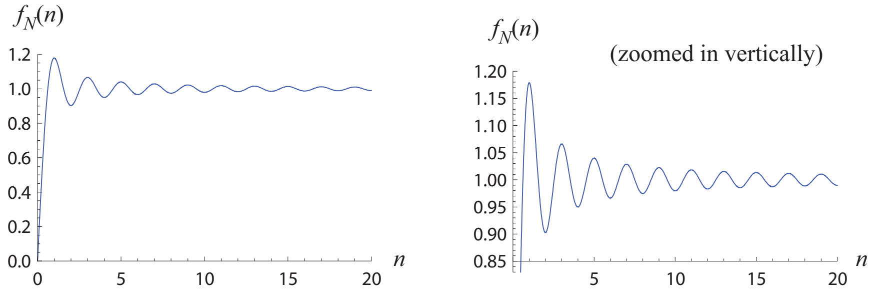


Figure 33

You can verify in Fig. 33 that the local maxima and minima occur at integer values of n. This curve has the same shape as the two curves in Fig. 30, with the only difference being the horizontal scale. Any actual step-function Fourier series such as the ones in Fig. 30 can be obtained from Fig. 33 by squashing or stretching it horizontally by the appropriate amount. This amount is determined by making the $n = 1$ maximum occur at $x = L/4N$, and then the $n = 2$ minimum occur at $x = 2L/4N$, and so on. The larger N is, the smaller these x values are, so the more squashed the wiggles are. If we increase N by a factor of, say, 10, then the wiggles in the Fourier-series plot get squashed by a factor of 10 horizontally (and are unchanged vertically). We can verify this in Fig. 30. If we look at, say, the fourth maximum, we see that in the $N = 10$ plot it occurs at about $x = 0.175$, and in the $N = 100$ plot it occurs at slightly less than $x = 0.02$ (it's hard to see exactly, but it's believable that it's about 0.0175).

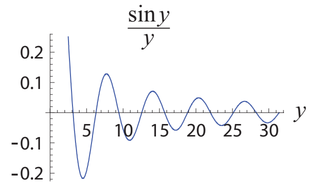


Figure 34

The envelope

What is the envelope of each of the curves in Fig. 30? (We'll be concerned with just the asymptotic behavior, that is, not right near the discontinuity.) To answer this, we must first find the envelope of the $f_N(n)$ curve in Fig. 33. Fig. 34 shows a plot of $(\sin y)/y$. Up to a factor of $2A/\pi$, the function $f_N(n)$ equals the area under the $(\sin y)/y$ curve, out to $y = n\pi$. This area oscillates around its average value (which is $\pi/2$) due to the "up" and " down" bumps of $(\sin y)/y$.

A local maximum of $f_N(n)$ is achieved when $(\sin y)/y$ has just completed an "up" bump, and a local minimum is achieved when $(\sin y)/y$ has just completed a "down" bump. The amplitude of the oscillations of $f_N(n)$ is therefore $2A/\pi$ times half of the area of one of the bumps at $y = n\pi$. The area of a simple sine bump is 2, so half of the area of a bump of $(\sin y)/y$, in a region where y is approximately equal to $n\pi$, is $(1/2) \cdot 2/n\pi = 1/n\pi$. (We're assuming that n is reasonably large, which means that the value of $1/n\pi$ is approximately constant over the span of a bump.) So the amplitude of the oscillations of $f_N(n)$ is

$$\text{Amplitude} \;=\; \frac{2A}{\pi} \cdot \frac{1}{n\pi} = \frac{2A}{n\pi^2} \approx \frac{A}{5n}\,. \tag{106}$$

You can (roughly) verify this by sight in Fig. 33. Remember that n counts the number of maxima and minima, not just the maxima.

To find the envelope of the actual Fourier-series plots in Fig. 30, we need to convert from n back to x. Using $x \equiv n(L/4N) \Longrightarrow n \equiv 4Nx/L$, the amplitude becomes

$$\text{Amplitude} \;=\; \frac{2A}{(4Nx/L)\pi^2} = \frac{AL}{2\pi^2 Nx} \propto \frac{1}{Nx}\,. \tag{107}$$

So up to some constants, this is a $1/x$ function with a $1/N$ coefficient. The larger N is, the quicker it dies off to zero.

What if $f(x)$ isn't a square wave and instead has a nonzero slope emanating from the discontinuity? The reasoning above (where we wrote $f(x)$ as $f_{\text{step}}(x) + f_{\text{diff}}(x)$) tells us that we still have the same envelope of the form, $AL/2\pi^2 Nx$, with the only difference being that the envelope is now measured with respect to a line with nonzero slope.

3.7 Convergence

Fourier's theorem states that any sufficiently well-behaved function can be written as the series in Eq. (1), where the a_n and b_n coefficients are given by Eqs. (9,11,12). But what does this statement actually mean? Does it mean only that the *value* of the Fourier series at a given x equals the value of the function $f(x)$ at this x? Or does it mean that the two functions on either side of Eq. (1) are actually the exact same function of x? In other words, is it the case that not only the values agree, but all the *derivatives* do also?

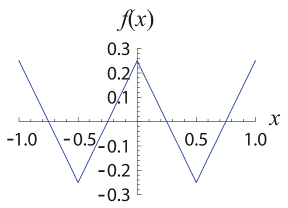


Figure 35

It turns out that Fourier's theorem makes only the first of these claims – that the values are equal. It says nothing about the agreement of the derivatives. They might agree, or they might not, depending on the function. Let's give a concrete example to illustrate this.

Consider the periodic triangular function shown in Fig. 35. For ease of notation, we have chosen the period L to be 1. The function is defined by

$$f(x) = \begin{cases} x + 1/4 & (-1/2 < x < 0) \\ -x + 1/4 & (0 < x < 1/2). \end{cases} \tag{108}$$

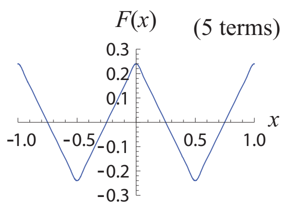


Figure 36

The task of Problem [to be added] is to find the Fourier series for this function. The result is

$$\begin{aligned} F(x) &= \frac{2}{\pi^2} \sum_{n \text{ odd}} \frac{\cos(2\pi n x)}{n^2} \\ &= \frac{2}{\pi^2} \left(\cos(2\pi x) + \frac{\cos(6\pi x)}{9} + \frac{\cos(10\pi x)}{25} + \cdots \right). \end{aligned} \tag{109}$$

We're using $F(x)$ to denote the Fourier series, to distinguish it from the actual function $f(x)$. Due to the n^2 factors in the denominators, the partial sums of $F(x)$ converge very quickly to the actual triangular function $f(x)$. This is evident from Fig. 36, which gives a very good approximation to $f(x)$ despite including only five terms in the partial sum.

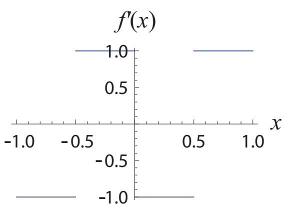


Figure 37

What if we calculate the derivative of the Fourier series, $F'(x)$, and compare it with the derivative of the actual function, $f'(x)$? The slope of $f(x)$ is either 1 or -1, so $f'(x)$ is simply the step function shown in Fig. 37. The derivative of $F(x)$ is easily calculated by differentiating each term in the infinite sum. The result is

$$\begin{aligned} F'(x) &= -\frac{4}{\pi} \sum_{n \text{ odd}} \frac{\sin(2\pi n x)}{n} \\ &= -\frac{4}{\pi} \left(\sin(2\pi x) + \frac{\sin(6\pi x)}{3} + \frac{\sin(10\pi x)}{5} + \cdots \right). \end{aligned} \tag{110}$$

This is the (negative of the) Fourier series we already calculated in Eq. (94), with $A = 1$ and $L = 1$. The plots of the partial sums of $F'(x)$ with 10 and 100 terms are shown in Fig. 38. In the infinite-sequence limit, $F'(x)$ does indeed equal the $f'(x)$ step function in Fig. 37. The Gibbs phenomenon discussed in the previous section might make you think otherwise, but the Gibbs overshoot is squeezed infinitely close to the discontinuity in the infinite-sequence limit. So for any given value of x that isn't at the discontinuity, $F'(x)$ approaches the correct value of ± 1 in the infinite-sequence limit.

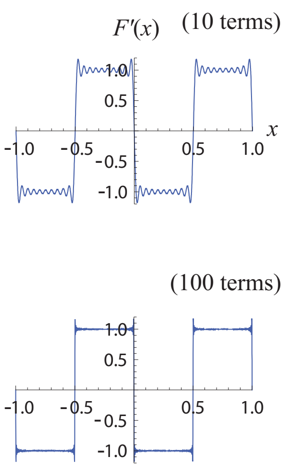


Figure 38

What if we go a step further and calculate the second derivative of the Fourier series, $F''(x)$, and compare it with the second derivative of the actual function, $f''(x)$? The slope of $f'(x)$ is zero except at the discontinuities, where it is $\pm\infty$, so $f''(x)$ is shown in Fig. 39. The arrows indicate infinite values. These infinite value are actually delta functions (see Problem [to be added]). The derivative of $F'(x)$ is again easily calculated by differentiating each term in the infinite sum. The result is

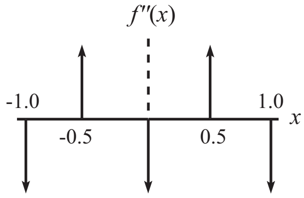


Figure 39

$$F''(x) = -8 \sum_{n \text{ odd}} \cos(2\pi n x)$$
$$= -8\Big(\cos(2\pi x) + \cos(6\pi x) + \cos(10\pi x) + \cdots\Big). \tag{111}$$

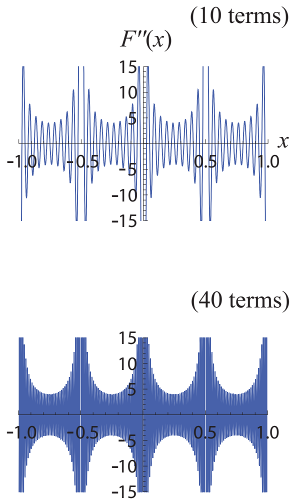


Figure 40

This can also be obtained by finding the Fourier series coefficients for $f''(x)$ via Eqs. (9,11,12) (see problem [to be added]). Plots of partial sums of $F''(x)$ with 10 and 40 terms are shown in Fig. 40. We now have a problem. These curves don't look anything like the plot of $f''(x)$ in Fig. 39. And if you plotted higher partial sums, you would find that the envelope is the same, with the only difference being the increase in the frequency of the oscillations. The curves definitely don't converge to $f''(x)$, a function that is zero everywhere except at the discontinuities.

Apparently, the Fourier series $F'(x)$ equals the function $f'(x)$ as far as the *value* is concerned, but not as far as the *derivative* is concerned. This statement might sound a little odd, in view of the fact that the second plot in Fig. 38 *seems* to indicate that except at the discontinuities, $F'(x)$ approaches a nice straight line with slope zero. But let's zoom in on a piece of this "line" and see what it looks like up close. Let's look at the middle of a step. The region between $x = 0.2$ and $x = 0.3$ is shown in Fig. 41 for partial sums with 10, 30, and 100 terms. The *value* of the function converges to -1, but slope doesn't converge to zero, because although the amplitude of the oscillations decreases, the frequency increases. So the slope ends up oscillating back and forth with the same amplitude. In short, the plot for the 100-term partial sum is simply an (essentially) scaled down version of the plot for the 10-term partial sum. This can be seen by looking at a further zoomed-in version of the 100-term partial sum, as shown in Fig. 42 (the additional zoom factor is 10 for both axes). This looks just like the plot for the 10-term partial sum. The proportions are the same, so the slope is the same.

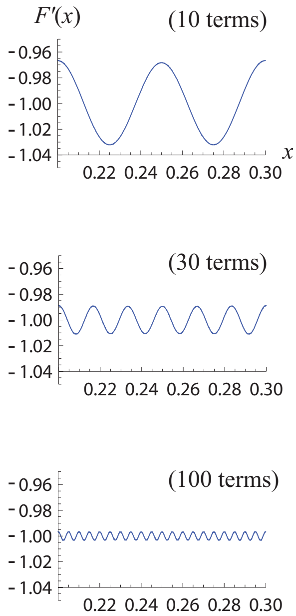


Figure 41

The general condition under which the *value* of the Fourier series $F(x)$ equals the *value* of the original function $f(x)$ (except at isolated discontinuities) is that $f(x)$ be square integrable. That is, the integral (over one period) of the square of $f(x)$ is finite. In the above example, $f(x)$ is square integrable, consistent with the fact that $F(x)$ agrees with $f(x)$. Additionally, $f'(x)$ is square integrable, consistent with the fact that $F'(x)$ agrees with $f'(x)$.

However, $f''(x)$ is *not* square integrable (because it contains delta functions), consistent with the fact that $F''(x)$ does *not* agree with $f''(x)$. A quick way to see why the square of a delta function isn't integrable (in other words, why the integral is infinite) is to consider the delta function to be a thin box with width a and height $1/a$, in the $a \to 0$ limit. The square of this function is a thin box with width a and height $1/a^2$. The area of this box is $a(1/a^2) = 1/a$, which diverges in the $a \to 0$ limit. So $(\delta(x))^2$ isn't integrable. Another even quicker way is to use the main property of a delta function, namely $\int \delta(x) f(x)\, dx = f(0)$. Letting $f(x)$ be a delta function here gives $\int \delta(x)\delta(x)\, dx = \delta(0) = \infty$. Hence, $(\delta(x))^2$ is not integrable.

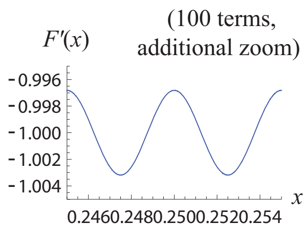


Figure 42

3.8 Relation between transforms and series

The exponential Fourier-series relations in Section 3.2 were

$$\boxed{f(x) = \sum_{n=-\infty}^{\infty} C_n e^{i2\pi nx/L}} \qquad \text{where} \qquad \boxed{C_n = \frac{1}{L}\int_0^L f(x)e^{-i2\pi nx/L}dx} \tag{112}$$

(We'll work with exponential series instead of trig series in this section.) And the Fourier-transform relations in Section 3.3 were

$$\boxed{f(x) = \int_{-\infty}^{\infty} C(k)e^{ikx}dk} \qquad \text{where} \qquad \boxed{C(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ikx}dx} \tag{113}$$

Eq. (112) says that any (reasonably well-behaved) periodic function can be written as a sum of exponentials (or sines and cosines).[9] But now that we've learned about Fourier transforms, what if we pretend that we don't know anything about Fourier *series* and instead simply try to calculate the Fourier *transform* of a periodic function using Eq. (113)? A periodic function can certainly be viewed as just a "normal" function that we might reasonably want to find the Fourier transform of, so something sensible should come out of this endeavor.

Whatever result we obtain from Eq. (113), it had better somehow reduce to Eq. (112). In particular, there had better not be any frequencies (k values) in the $C(k)$ Fourier-transform expression for $f(x)$ besides the *discrete* ones of the form $2\pi n/L$, because these are the only ones that appear in Eq. (112). So let's see what happens when we take the Fourier transform of a periodic function. Actually, let's first work backwards, starting with Eq. (112), and see what the transform has to be. Then we'll go in the "forward" direction and derive the Fourier series directly by calculating the Fourier transform.

3.8.1 From series to transform

Fourier transforms involve integrals over k, so if we work backwards from the discrete sum in Eq. (112), our goal is to somehow write this sum as an integral. This is where our knowledge of delta functions will help us out. We know that a delta function picks out a particular value of the integration variable, so if we have a sum of delta functions located at different places, then they will turn a continuous integral into a discrete sum. Therefore, assuming that we have a function $f(x)$ that can be written in the form of Eq. (112), we'll propose the following Fourier transform of $f(x)$:

$$\boxed{C_\delta(k) = \sum_{n=-\infty}^{\infty} C_n \delta(k - 2\pi n/L)} \tag{114}$$

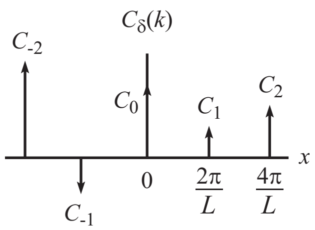


Figure 43

where the C_n's are given by Eq. (112). This $C_\delta(k)$ function is shown schematically in Fig. 43 (it's technically a distribution and not a function, but we won't worry about that). The arrows are the standard way to represent delta functions, with the height being the coefficient of the delta function, which is C_n here. The value of $C_\delta(k)$ at the delta functions is infinite, of course, so remember that the height of the arrow represents the area of the delta function (which is all that matters when doing an integral involving a delta function), and not the (infinite) value of the function.

[9]The route we followed in Sections 3.1 and 3.2 was to simply accept this as fact. However, we'll actually prove it from scratch in Section 3.8.3 below.

To determine if the $C_\delta(k)$ function in Eq. (114) is the correct Fourier transform of $f(x)$, we need to plug it into the first equation in Eq. (113) and see if it yields $f(x)$. We find

$$f(x) \stackrel{?}{=} \int_{-\infty}^{\infty} C_\delta(k)e^{ikx}\,dk = \int_{-\infty}^{\infty}\left(\sum_{n=-\infty}^{\infty} C_n\delta(k-2\pi n/L)\right)e^{ikx}\,dk. \tag{115}$$

As k runs from $-\infty$ to ∞, the integrand is zero everywhere except at the delta functions, where k takes the form, $k = 2\pi n/L$. Using the $\int f(k)\delta(k-k_0)\,dk = f(k_0)$ property of the delta function, we see that the integration across each delta function simply yields the value of the rest of the integrand evaluated at $k = 2\pi n/L$. This value is $C_n e^{i2\pi nx/L}$. Therefore, since n runs from $-\infty$ to ∞, Eq. (115) becomes

$$f(x) \stackrel{?}{=} \sum_{n=-\infty}^{\infty} C_n e^{i2\pi nx/L}. \tag{116}$$

And from the first equation in Eq. (112), the righthand side here does indeed equal $f(x)$, as desired. So the $C(k)$ in Eq. (114) does in fact make the first equation in Eq. (113) true, and is therefore the correct Fourier transform of $f(x)$.

3.8.2 From transform to series

Given a periodic function $f(x)$, let's now work in the "forward" direction and *derive* the expression for $C_\delta(k)$ in Eq. (114), which we just showed correctly leads to the series in Eq. (112). We'll derive this $C_\delta(k)$ from scratch by starting with the general Fourier-transform expression for $C(k)$ in Eq. (113). We'll pretend that we've never heard about the concept of a Fourier *series* for a periodic function.

Our goal is to show that if $f(x)$ is periodic with period L, then the $C(k)$ given in Eq. (113) reduces to the $C_\delta(k)$ given in Eq. (114), where the C_n's are given by Eq. (112). That is, we want to show that the $C(k)$ in Eq. (113) equals

$$\boxed{C(k) = \sum_{n=-\infty}^{\infty}\left(\frac{1}{L}\int_0^L f(x)e^{-i2\pi nx/L}dx\right)\delta(k-2\pi n/L)} \tag{117}$$

To demonstrate this, let's break up the integral for $C(k)$ in Eq. (113) into intervals of length L. This gives

$$C(k) = \frac{1}{2\pi}\left(\cdots\int_{-L}^{0} f(x)e^{-ikx}dx + \int_0^L f(x)e^{-ikx}dx + \int_L^{2L} f(x)e^{-ikx}dx + \cdots\right). \tag{118}$$

The $f(x)$'s in each of these integrals run through the same set of values, due to the periodicity. And the e^{-ikx} values simply shift by successive powers of e^{-ikL} in each integral. For example, if we define y by $x \equiv y + L$ and substitute this into the $\int_L^{2L}$ integral above, we obtain

$$\int_L^{2L} f(x)e^{-ikx}dx = \int_0^L f(y+L)e^{-ik(y+L)}dy = e^{-ikL}\int_0^L f(y)e^{-iky}dy, \tag{119}$$

where we have used the $f(y+L) = f(y)$ periodicity. So the $\int_L^{2L}$ integral is simply e^{-ikL} times the $\int_0^L$ integral. Likewise, the $\int_{2L}^{3L}$ integral is e^{-2ikL} times the $\int_0^L$ integral. And so on. Therefore, if we factor the $\int_0^L$ integral out of each integral in Eq. (118), we obtain

$$C(k) = \frac{1}{2\pi}\left(\int_0^L f(x)e^{-ikx}dx\right)\left(\cdots e^{2ikL} + e^{ikL} + 1 + e^{-ikL} + e^{-2ikL}\cdots\right). \tag{120}$$

We must now evaluate the sum,

$$S(k) \equiv \cdots e^{2ikL} + e^{ikL} + 1 + e^{-ikL} + e^{-2ikL} \cdots. \tag{121}$$

We'll do it quantitatively below, but first let's be qualitative to get an idea of what's going on. $S(k)$ is the sum of unit vectors in the complex plane that keep rotating around indefinitely. So they should average out to zero. The one exception occurs when $e^{ikL} = 1$, that is, when k takes the form of $k = 2\pi n/L$. In this case, all the terms in the (infinite) sum equal 1 (the unit vectors don't rotate at all), so $S(k)$ is infinite. This reasoning is basically correct, except that it doesn't get quantitative about the nature of the infinity (we'll find below that it's delta function). And it's also a little sloppy in the "averaging out to zero" part, because the sum doesn't converge, the way it stands. All the terms have magnitude 1, so it matters where you start and stop the sum.

Let's now be quantitative. Following the strategy we used many times in Section 3.5, we can get a handle on $S(k)$ by multiplying the terms by successive powers of a number r that is slightly less than 1, starting with the $e^{\pm ikL}$ terms and then working outward in both directions. We'll then take the $r \to 1$ limit to recover the original sum. So our modified sum (call it $S_r(k)$) is

$$S_r(k) = \cdots + r^3 e^{3ikL} + r^2 e^{2ikL} + re^{ikL} + 1 + re^{-ikL} + r^2 e^{-2ikL} + r^3 e^{-3ikL} + \cdots. \tag{122}$$

Summing the two geometric series on either side of the "1" turns $S_r(k)$ into

$$S_r(k) = \frac{re^{ikL}}{1 - re^{ikL}} + 1 + \frac{re^{-ikL}}{1 - re^{-ikL}}. \tag{123}$$

Getting a common denominator and combining these terms yields

$$S_r(k) = \frac{1 - r^2}{1 + r^2 - 2r\cos(kL)}. \tag{124}$$

If $\cos(kL) \neq 1$, then $S_r(k)$ equals zero in the $r \to 1$ limit, because the denominator is nonzero in this limit. So as we saw qualitatively above, if $k \neq 2\pi n/L$, then $S(k) = 0$. However, things are tricker if $\cos(kL) = 1$, that is, if $k = 2\pi n/L$ for some integer n. In this case we obtain $S_r(k) = (1 - r^2)/(1 - r)^2 = (1 + r)/(1 - r)$, which goes to infinity in the $r \to 1$ limit. We can get a handle on this infinity as follows.

Define κ by $k \equiv 2\pi n/L + \kappa$. We are concerned with very small κ values, because these values correspond to k being very close to $2\pi n/L$. In terms of κ, $S_r(k)$ becomes (using $\cos\theta \approx 1 - \theta^2/2$ to obtain the second line)

$$\begin{aligned} S_r(\kappa) &= \frac{1 - r^2}{1 + r^2 - 2r\cos(\kappa L)} \\ &\approx \frac{1 - r^2}{1 + r^2 - 2r(1 - \kappa^2 L^2/2)} \\ &= \frac{(1 + r)(1 - r)}{(1 - r)^2 + r\kappa^2 L^2}. \end{aligned} \tag{125}$$

If we now let $r \equiv 1 - \epsilon$ (so we're concerned with very small ϵ), then the $(1 + r)$ factor in the numerator is essentially equal to 2. So when we take the $\epsilon \to 0$ limit to obtain the original sum $S(k)$, we get

$$S(\kappa) = \lim_{\epsilon \to 0} S_\epsilon(\kappa) = \lim_{\epsilon \to 0} \frac{2\epsilon}{\epsilon^2 + \kappa^2 L^2}. \tag{126}$$

But from Eq. (83), this limit equals $2\pi\,\delta(\kappa L)$, which from the second equation in Eq. (75) equals $(2\pi/L)\delta(\kappa)$, which in turn equals $(2\pi/L)\delta(k - 2\pi n/L)$, from the definition of κ.[10] $S(k)$ diverges for any k of the form $2n\pi/L$, so we arrive at

$$S(k) = \sum_{n=-\infty}^{\infty} \frac{2\pi}{L}\,\delta(k - 2\pi n/L). \tag{127}$$

It's legal to just add up all the different delta functions, because each one doesn't affect any of the others (because they're zero everywhere except at the divergence). Plugging this $S(k)$ into Eq. (120) then gives

$$C(k) = \frac{1}{L}\left(\int_0^L f(x)e^{-ikx}dx\right)\sum_{n=-\infty}^{\infty} \delta(k - 2\pi n/L). \tag{128}$$

The delta functions mean that only k values of the form $2\pi n/L$ are relevant, so we can replace the k in the exponent with $2\pi n/L$ (after bringing the integral inside the sum). This gives

$$\begin{aligned} C(k) &= \sum_{n=-\infty}^{\infty}\left(\frac{1}{L}\int_0^L f(x)e^{-i2\pi nx/L}dx\right)\delta(k - 2\pi n/L), \\ &\equiv \sum_{n=-\infty}^{\infty} C_n\delta(k - 2\pi n/L), \end{aligned} \tag{129}$$

where C_n is defined as in Eq. (112). We have therefore arrived at the desired expression for $C(k)$ in Eq. (117), or equivalently Eq. (114).

And as we saw in Section 3.8.1, if we plug this Fourier *transform*, $C(k)$, into the first equation in Eq. (113) to obtain $f(x)$, we end up with the righthand side of Eq. (116), which is the desired Fourier *series* of $f(x)$. We have therefore demonstrated how the continuous Fourier-transform expansion of a periodic function leads to the discrete Fourier series.

Remarks:

1. The delta functions in Eq. (129) occur only at k values of the form $2\pi n/L$, so the above derivation explains why all the k values in the Fourier series in Eqs. (1) and (19) are multiples of $2\pi/L$. When we introduced Eq. (1) out of the blue, you may have wondered why no other frequencies were needed to obtain an arbitrary periodic function with period L.

   For example, in the Fourier-series formalism, it's not so obvious why a k value of, say, π/L isn't needed. But in the Fourier-transform formalism, it's clear that $C(\pi/L)$ equals zero, because the terms in the sum in Eq. (120) simply alternate between 1 and -1, and therefore add up to zero. (As usual, this can be made rigorous by tapering off the sum.) And similar reasoning holds for all k values not of the form $2\pi n/L$. Therefore, since non-$2\pi n/L$ values of k don't appear in the Fourier transform of $f(x)$, they don't appear in the Fourier series either.

2. For a periodic function, we found above that every value of $C(k)$ is either zero when $k \neq 2\pi n/L$, or infinite when $k = 2\pi n/L$. There is another fairly quick way to see this, as follows. First, consider a k value of the form, $k = 2\pi R/L$, where R is a rational number, a/b. Then the product $f(x)e^{-ikx}$ is periodic with period bL. If the integral of $f(x)e^{-ikx}$ over this period is *exactly* zero, then the integral from $-\infty$ to ∞ is also zero (again, with the integral tapered off). And if the integral over this period is *not* zero, then the integral from $-\infty$ to ∞ is

[10] It's no surprise that we obtained the Lozentzian representation of the delta function in Eq. (126). Our "tapering" method involving powers of r is simply a discrete version of an exponential taper. And we know from Section 3.5.3 that an exponential taper leads to a Lorentzian delta function.

infinite, because we're adding up a given nonzero number, whatever it may be, an infinite number of times. Finally, since the rational numbers are dense on the real line, this result should hold for all values of k. So for all k, $C(k)$ is either zero or infinite. As it turns out, $C(k)$ is nearly always zero. ♣

3.8.3 Derivation of Fourier's theorem

We started this chapter by invoking Fourier's theorem. That is, we accepted the fact that any (reasonably well-behaved) periodic function can be written as a Fourier-series sum of sines and cosines, or exponentials; see Eqs. (1) and (19). We then used this fact to derive (by taking the $L \to \infty$ limit) the Fourier-transform relations in Eq. (43). It turns out, however, that armed with our knowledge of delta functions, we can actually derive all of these results from scratch. And the order of the derivation is reversed; we will first derive the Fourier-transform expression for an arbitrary (reasonably well-behaved) function, and then we will use this result to derive the Fourier-series expression for a periodic function. Let's be very systematic and derive things step-by-step:

1. The first step is to use the "tapering" method of Section 3.5 to derive the integral representation of the delta function given in Eq. (90), which we will copy here:

$$\delta(k - k_0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i(k-k_0)x}dx. \tag{130}$$

   When we derived this in Section 3.5, we formulated things in terms of Fourier transforms, but this language isn't necessary. Even if we had never heard of Fourier transforms, we could have pulled the integral $\int_{-\infty}^{\infty} e^{-ikx}dx$ out of a hat, and then used the tapering method to show that it is a delta function of k. That is, it has area 1 and is zero everywhere except at $k = 0$. We need not know anything about Fourier transforms to derive this fact.

2. Given an arbitrary (not necessarily periodic) function $f(x)$, let's define (out of the blue) a function $C(k)$ by

$$C(k) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ikx}dx. \tag{131}$$

   The is a purely definitional statement and has no content. We haven't done anything with actual substance here.

3. Given this definition of $C(k)$, we will now use the expression for the delta function in Eq. (130) to show that

$$f(x) = \int_{-\infty}^{\infty} C(k)e^{ikx}dk. \tag{132}$$

   This is a statement with actual substance. To prove it, we first need to plug the above definition of $C(k)$ into the righthand side to obtain

$$f(x) \stackrel{?}{=} \int_{-\infty}^{\infty}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} f(x')e^{-ikx'}dx'\right)e^{ikx}dk, \tag{133}$$

   where we have been careful to label the dummy integration variable from Eq. (131) as x'. Switching the order of integration and rearranging things gives

$$f(x) \stackrel{?}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x')\left(\int_{-\infty}^{\infty} e^{-ik(x'-x)}dk\right)dx'. \tag{134}$$

From Eq. (130), the quantity in parentheses is simply $2\pi\delta(x'-x)$, with different letters. So we obtain

$$f(x) \stackrel{?}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x')\delta(x'-x)\,dx'. \tag{135}$$

And this is indeed true, because the righthand side equals $f(x)$, due to the properties of the delta function. We have therefore successfully derived the Fourier-transform relations in Eq. (43). More precisely, we wrote one of them, Eq. (131), down by definition. And we then derived the other one, Eq. (132).

We should mention again that whenever we deal with integrals of the type $\int_{-\infty}^{\infty} e^{-ikx}dx$, it is understood that we are technically using some sort of tapering method to make sense of them. But we know that the result will always be a delta function, so we don't actually have to go through the tapering procedure each time.

4. Having derived the Fourier-transform relations in Eq. (43) for arbitrary functions, our task is now to derive the Fourier-series relations for periodic functions, Eqs. (19) and (20), which we will copy here:

$$f(x) = \sum_{n=-\infty}^{\infty} C_n e^{i2\pi nx/L}, \tag{136}$$

where the C_n's are given by

$$C_n = \frac{1}{L}\int_0^L f(x)e^{-i2\pi nx/L}dx. \tag{137}$$

(Alternatively, we could use the trig series given in Eq. (3.1).) But we already performed this derivation in Section 3.8.2, so we won't go through it again. The main point is that the periodicity of $f(x)$ produces a collection of delta functions of k, which turns the continuous-integral Fourier-transform expression for $f(x)$ in Eq. (132) into the discrete-sum Fourier-series expression for $f(x)$ in Eq. (136).

Chapter 4

Transverse waves on a string

David Morin, morin@physics.harvard.edu

In the previous three chapters, we built up the foundation for our study of waves. In the remainder of this book, we'll investigate various types of waves, such as waves on a string, sound waves, electromagnetic waves, water waves, quantum mechanical waves, and so on. In Chapters 4 through 6, we'll discuss the properties of the two basic categories of waves, namely *dispersive* waves, and *non-dispersive* waves. The rest of the book is then largely a series of applications of these results. Chapters 4 through 6 therefore form the heart of this book.

A non-dispersive system has the property that all waves travel with the same speed, independent of the wavelength and frequency. These waves are the subject of this and the following chapter (broken up into longitudinal and transverse waves, respectively). A dispersive system has the property that the speed of a wave *does* depend on the wavelength and frequency. These waves are the subject of Chapter 6. They're a bit harder to wrap your brain around, the main reason being the appearance of the so-called *group velocity*. As we'll see in Chapter 6, the difference between non-dispersive and dispersive waves boils down to the fact that for non-dispersive waves, the frequency ω and wavelength k are related by a simple proportionality constant, whereas this is not the case for dispersive waves.

The outline of this chapter is as follows. In section 4.1 we derive the wave equation for transverse waves on a string. This equation will take exactly the same form as the wave equation we derived for the spring/mass system in Section 2.4, with the only difference being the change of a few letters. In Section 4.2 we discuss the reflection and transmission of a wave from a boundary. We will see that various things can happen, depending on exactly what the boundary looks like. In Section 4.3 we introduce the important concept of *impedance* and show how our previous results can be written in terms of it. In Section 4.4 we talk about the energy and power carried by a wave. In Section 4.5 we calculate the form of standing waves on a string that has boundary conditions that fall into the extremes (a fixed end or a "free" end). In Section 4.6 we introduce damping, and we see how the amplitude of a wave decreases with distance in a scenario where one end of the string is wiggled with a constant amplitude.

4.1 The wave equation

The most common example of a non-dispersive system is a string with transverse waves on it. We'll see below that we obtain essentially the same wave equation for *transverse* waves on a string as we obtained for the $N \to \infty$ limit of *longitudinal* waves in the mass/spring system in Section 2.4. Either of these waves could therefore be used for our discussion of the properties of non-dispersive systems. However, the reason why we've chosen to study transverse waves on a string in this chapter is that transverse waves are generally easier to visualize than longitudinal ones.

Consider a string with tension T and mass density μ (per unit length). Assume that it is infinitesimally thin and completely flexible. And assume for now that it extends infinitely in both directions. We'll eventually relax this restriction. Consider *small* transverse displacements of the string (we'll be quantitative about the word "small" below). Let x be the coordinate along the string, and let ψ be the transverse displacement. (There's no deep reason why we're using ψ for the displacement instead of the ξ we used in Chapter 2.) Our goal is to find the most general form of $\psi(x,t)$.

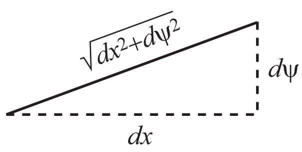


Figure 1

Consider two nearby points separated by a displacement dx in the longitudinal direction along the string, and by a displacement $d\psi$ in the transverse direction. If $d\psi$ is small (more precisely if the slope $d\psi/dx$ is small; see below), then we can make the approximation that all points in the string move only in the transverse direction. That is, there is no longitudinal motion. This is true because in Fig. 1 the length of the hypotenuse equals

$$\sqrt{dx^2 + d\psi^2} = dx\sqrt{1 + \left(\frac{d\psi}{dx}\right)^2} \approx dx\left(1 + \frac{1}{2}\left(\frac{d\psi}{dx}\right)^2\right) = dx + d\psi\frac{1}{2}\left(\frac{d\psi}{dx}\right). \tag{1}$$

This length (which is the farthest that a given point can move to the side; it's generally less that this) differs from the length of the long leg in Fig. 1 by an amount $d\psi(d\psi/dx)/2$, which is only $(d\psi/dx)/2$ times as large as the transverse displacement $d\psi$. Since we are assuming that the slope $d\psi/dx$ is small, we can neglect the longitudinal motion in comparison with the transverse motion. Hence, all points essentially move only in the transverse direction. We can therefore consider each point to be labeled with a unique value of x. That is, the ambiguity between the original and present longitudinal positions is irrelevant. The string will stretch slightly, but we can always assume that the amount of mass in any given horizontal span stays essentially constant.

We see that by the phrase "small transverse displacements" we used above, we mean that the *slope* of the string is small. The slope is a dimensionless quantity, so it makes sense to label it with the word "small." It makes no sense to say that the actual transverse displacement is small, because this quantity has dimensions.

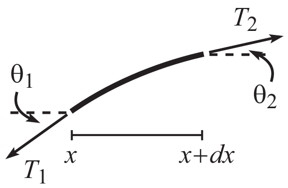


Figure 2

Our strategy for finding the wave equation for the string will be to write down the transverse $F = ma$ equation for a little piece of string in the span from x to $x + dx$. The situation is shown in Fig. 2. (We'll ignore gravity here.) Let T_1 and T_2 be the tensions in the string at the ends of the small interval. Since the slope $d\psi/dx$ is small, the slope is essentially equal to the θ angles in the figure. (So these angles are small, even though we've drawn them with reasonable sizes for the sake of clarity.) We can use the approximation $\cos\theta \approx 1 - \theta^2/2$ to say that the longitudinal components of the tensions are equal to the tensions themselves, up to small corrections of order $\theta^2 \approx (d\psi/dx)^2$. So the longitudinal components are (essentially) equal to T_1 and T_2. Additionally, from the above reasoning concerning (essentially) no longitudinal motion, we know that there is essentially no longitudinal acceleration of the little piece in Fig. 2. So the longitudinal forces must cancel. We therefore conclude that $T_1 = T_2$. Let's call this common tension T.

However, although the two tensions and their longitudinal components are all equal, the same thing *cannot* be said about the transverse components. The transverse components differ by a quantity that is *first* order in $d\psi/dx$, and this difference can't be neglected. This difference is what causes the transverse acceleration of the little piece, and it can be calculated as follows.

In Fig. 2, the "upward" transverse force on the little piece at its right end is $T\sin\theta_1$, which essentially equals T times the slope, because the angle is small. So the upward force at the right end is $T\psi'(x+dx)$. Likewise, the "downward" force at the left end is $-T\psi'(x)$. The net transverse force is therefore

$$F_{\text{net}} = T\Big(\psi'(x+dx) - \psi'(x)\Big) = T\,dx\frac{\psi'(x+dx)-\psi'(x)}{dx} \equiv T\,dx\frac{d^2\psi(x)}{dx^2}\,, \tag{2}$$

where we have assumed that dx is infinitesimal and used the definition of the derivative to obtain the third equality.[1] Basically, the difference in the first derivatives yields the second derivative. For the specific situation shown in Fig. 2, $d^2\psi/dx^2$ is negative, so the piece is accelerating in the downward direction. Since the mass of the little piece is $\mu\,dx$, the transverse $F = ma$ equation is

$$F_{\text{net}} = ma \quad\Longrightarrow\quad T\,dx\frac{d^2\psi}{dx^2} = (\mu\,dx)\frac{d^2\psi}{dt^2} \quad\Longrightarrow\quad \frac{d^2\psi}{dt^2} = \frac{T}{\mu}\frac{d^2\psi}{dx^2}\,. \tag{3}$$

Since ψ is a function of x and t, let's explicitly include this dependence and write ψ as $\psi(x,t)$. We then arrive at the desired wave equation (written correctly with partial derivatives now),

$$\boxed{\frac{\partial^2\psi(x,t)}{\partial t^2} = \frac{T}{\mu}\frac{\partial^2\psi(x,t)}{\partial x^2}} \qquad (\textit{wave equation}) \tag{4}$$

This takes exactly the same form as the wave equation we found in Section 2.4 for the $N\to\infty$ limit of the spring/mass system. The only difference is the replacement of the quantity E/ρ with the quantity T/μ. Therefore, *all* of our previous results carry over here. In particular, the solutions take the form,

$$\boxed{\psi(x,t) = Ae^{i(\pm kx\pm\omega t)}} \qquad \text{where} \qquad \boxed{\frac{\omega}{k} = \sqrt{\frac{T}{\mu}} \equiv c} \tag{5}$$

and where c is the speed of the traveling wave (from here on, we'll generally use c instead of v for the speed of the wave). This form does indeed represent a traveling wave, as we saw in Section 2.4. k and ω can take on any values, as long as they're related by $\omega/k = c$. The wavelength is $\lambda = 2\pi/k$, and the time period of the oscillation of any given point is $\tau = 2\pi/\omega$. So the expression $\omega/k = c$ can be written as

$$\frac{2\pi/\tau}{2\pi/\lambda} = c \quad\Longrightarrow\quad \frac{\lambda}{\tau} = c \quad\Longrightarrow\quad \lambda\nu = c, \tag{6}$$

where $\nu = 1/\tau$ is the frequency in cycles per second (Hertz).

For a given pair of k and ω values, the most general form of $\psi(x,t)$ can be written in many ways, as we saw in Section 2.4. From Eqs. (3.91) and (3.92) a couple of these ways are

$$\psi(x,t) = C_1\cos(kx+\omega t) + C_2\sin(kx+\omega t) + C_3\cos(kx-\omega t) + C_4\sin(kx-\omega t), \tag{7}$$

and

$$\psi(x,t) = D_1\cos kx\cos\omega t + D_2\sin kx\sin\omega t + D_3\sin kx\cos\omega t + D_4\cos kx\sin\omega t. \tag{8}$$

Of course, since the wave equation is linear, the most general solution is the sum of an arbitrary number of the expressions in, say, Eq. (8), for different pairs of k and ω values (as long as each pair is related by $\omega/k = \sqrt{T/\mu} \equiv c$). For example, a possible solution is

$$\psi(x,t) = A\cos k_1x\cos ck_1t + B\cos k_1x\sin ck_1t + C\sin k_2x\sin ck_2t + \text{etc}\ldots \tag{9}$$

[1] There is an ambiguity about whether we should write $\psi''(x)$ or $\psi''(x+dx)$ here. Or perhaps $\psi''(x+dx/2)$. But this ambiguity is irrelevant in the $dx\to 0$ limit.

Solutions of the form $f(x - ct)$

As we saw in Section 2.4, *any* function of the form $f(x - ct)$ satisfies the wave equation. There are two reasons why this functional form works. The first reason, as we showed in Eq. (2.97), is that if you simply plug $\psi(x, t) = f(x - ct)$ into the wave equation in Eq. (4), you will find that it works, provided that $c = \sqrt{T/\mu}$.

The second reason, as we also mentioned in Section 2.4, is due to Fourier analysis and linearity. Let's talk a little more about this reason. It is the combination of the following facts:

1. By Fourier analysis, we can write any function $f(z)$ as

$$f(z) = \int_{-\infty}^{\infty} C(k)e^{ikz}dk, \tag{10}$$

where $C(k)$ is given by Eq. (3.43). If we let $z \equiv x - ct$, then this becomes

$$f(x - ct) = \int_{-\infty}^{\infty} C(k)e^{ik(x-ct)}dk = \int_{-\infty}^{\infty} C(k)e^{i(kx-\omega t)}dk, \tag{11}$$

where $\omega \equiv ck$.

2. We showed in Section 2.4 (and you can quickly verify it again) that any exponential function of the form $e^{i(kx-\omega t)}$ satisfies the wave equation, provided that $\omega = ck$, which is indeed the case here.

3. Because the wave equation is linear, any linear combination of solutions is again a solution. Therefore, the integral in Eq. (11) (which is a continuous linear sum) satisfies the wave equation.

In this reasoning, it is critical that ω and k are related linearly by $\omega = ck$, where c takes on a constant value, independent of ω and k (and it must be equal to $\sqrt{T/\mu}$, or whatever constant appears in the wave equation). If this relation weren't true, then the above reasoning would be invalid, for reasons we will shortly see. When we get to dispersive waves in Chapter 6, we will find that ω does *not* equal ck. In other words, the ratio ω/k depends on k (or equivalently, on ω). Dispersive waves therefore cannot be written in the form of $f(x - ct)$. It is instructive to see exactly where the above reasoning breaks down when $\omega \neq ck$. This breakdown can be seen in mathematically, and also physically.

Mathematically, it is still true that any function $f(x - ct)$ can be written in the form of $\int_{-\infty}^{\infty} C(k)e^{ik(x-ct)}dk$. However, it is *not* true that these $e^{i(kx-(ck)t)}$ exponential functions are solutions to a dispersive wave equation (we'll see in Chapter 6 what such an equation might look like), because ω doesn't take the form of ck for dispersive waves. The actual solutions to a dispersive wave equation are exponentials of the form $e^{i(kx-\omega t)}$, where ω is some *nonlinear* function of k. That is, ω does *not* take the form of ck. If you want, you can write these exponential solutions as $e^{i(kx-c_k kt)}$, where $c_k \equiv \omega/k$ is the speed of the wave component with wavenumber k. But the point is that *a single value* of c doesn't work for all values of k.

In short, if by the ω in Eq. (11) we mean ω_k (which equals ck for nondispersive waves, but not for dispersive waves), then for dispersive waves, the above reasoning breaks down in the second equality in Eq. (11), because the coefficient of t in the first integral is ck (times $-i$), which isn't equal to the ω coefficient in the second integral. If on the other hand we want to keep the ω in Eq. (11) defined as ck, then for dispersive waves, the above reasoning breaks down in step 2. The exponential function $e^{i(kx-\omega t)}$ with $\omega = ck$ is simply not a solution to a dispersive wave equation.

The physical reason why the $f(x-ct)$ functional form doesn't work for dispersive waves is the following. Since the speed c_k of the Fourier wave components depends on k in a dispersive wave, the wave components move with different speeds. The shape of the wave at some initial time will therefore not be maintained as t increases (assuming that the wave contains more than a single Fourier component). This implies that the wave cannot be written as $f(x-ct)$, because this wave *does* keep the same shape (because the argument $x-ct$ doesn't change if t increases by Δt and x increases by $c\,\Delta t$).

The distortion of the shape of the wave can readily be seen in the case where there are just two wave components, which is much easier to visualize than the continuous infinity of components involved in a standard Fourier integral. If the two waves have (k,ω) values of $(1,1)$ and $(2,1)$, then since the speed is ω/k, the second wave moves with half the speed of the first. Fig. 3 shows the sum of these two waves at two different times. The total wave clearly doesn't keep the same shape, so it therefore can't be written in the form of $f(x-ct)$.

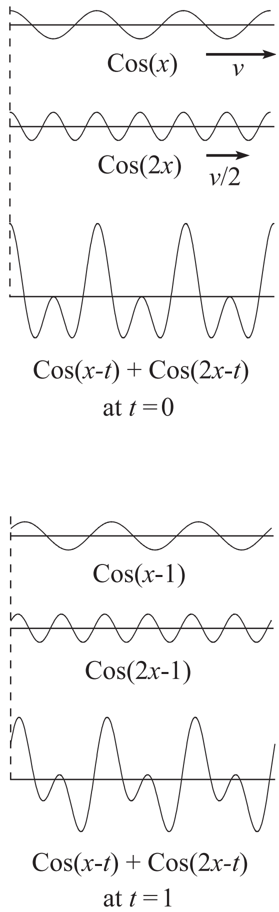


Figure 3

Fourier transform in 2-D

Having learned about Fourier transforms in Chapter 3, we can give another derivation of the fact that any solution to the wave equation in Eq. (4) can be written in terms of $e^{i(kx-\omega t)}$ exponentials, where $\omega = ck$. We originally derived these solutions in Section 2.4 by guessing exponentials, with the reasoning that since all (well enough behaved) functions can be built up from exponentials due to Fourier analysis, it suffices to consider exponentials. However, you might still feel uneasy about this "guessing" strategy, so let's be a little more systematic with the following derivation. This derivation involves looking at the Fourier transform of a function of two variables. In Chapter 3, we considered functions of only one variable, but the extension to two variables in quite straightforward.

Consider a wave $\psi(x,t)$ on a string, and take a snapshot at a given time. If $f(x)$ describes the wave at this instant, then from 1-D Fourier analysis we can write

$$f(x) = \int_{-\infty}^{\infty} C(k)e^{ikx}dk, \tag{12}$$

where $C(k)$ is given by Eq. (3.43). If we take a snapshot at a slightly later time, we can again write $\psi(x,t)$ in terms of its Fourier components, but the coefficients $C(k)$ will be slightly different. In other words, the $C(k)$'s are functions of time. So let's write them as $C(k,t)$. In general, we therefore have

$$\psi(x,t) = \int_{-\infty}^{\infty} C(k,t)e^{ikx}dk. \tag{13}$$

This equation says that at any instant we can decompose the snapshot of the string into its e^{ikx} Fourier components.

We can now do the same thing with the $C(k,t)$ function that we just did with the $\psi(x,t)$ function. But we'll now consider "slices" with constant k value instead of constant t value. If $g(t)$ describes the function $C(k,t)$ for a particular value of k, then from 1-D Fourier analysis we can write

$$g(t) = \int_{-\infty}^{\infty} \beta(\omega)e^{i\omega t}d\omega. \tag{14}$$

If we consider a slightly different value of k, we can again write $C(k,t)$ in terms of its Fourier components, but the coefficients $\beta(\omega)$ will be slightly different. That is, they are functions of k, so let's write them as $\beta(k,\omega)$. In general, we have

$$C(k,t) = \int_{-\infty}^{\infty} \beta(k,\omega)e^{i\omega t}d\omega. \tag{15}$$

This equation says that for a given value of k, we can decompose the function $C(k,t)$ into its $e^{i\omega t}$ Fourier components. Plugging this expression for $C(k,t)$ into Eq. (13) gives

$$\begin{aligned}\psi(x,t) &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\beta(k,\omega)e^{i\omega t}d\omega\right)e^{ikx}dk \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\beta(k,\omega)e^{i(kx+\omega t)}dk\,d\omega. \qquad (16)\end{aligned}$$

This is a general result for *any* function of two variables; it has nothing to do with the wave equation in Eq. (4). This result basically says that we can just take the Fourier transform in each dimension separately.

Let's now apply this general result to the problem at hand. That is, let's plug the above expression for $\psi(x,t)$ into Eq. (4) and see what it tells us. We will find that ω must be equal to ck. The function $\beta(k,\omega)$ is a constant as far as the t and x derivatives in Eq. (4) are concerned, so we obtain

$$\begin{aligned}0 &= \frac{\partial^2\psi(x,t)}{\partial t^2} - c^2\frac{\partial^2\psi(x,t)}{\partial x^2} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\beta(k,\omega)\left(\frac{\partial^2 e^{i(kx+\omega t)}}{\partial t^2} - c^2\frac{\partial^2 e^{i(kx+\omega t)}}{\partial x^2}\right)dk\,d\omega \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\beta(k,\omega)e^{i(kx+\omega t)}\Big(-\omega^2 - c^2(-k^2)\Big)\,dk\,d\omega. \qquad (17)\end{aligned}$$

Since the $e^{i(kx+\omega t)}$ exponentials here are linearly independent functions, the only way that this sum can be zero for all x and t is if the coefficient of each separate $e^{i(kx+\omega t)}$ term is zero. That is,

$$\beta(k,\omega)(\omega^2 - c^2k^2) = 0 \qquad (18)$$

for all values of k and ω.[2] There are two ways for this product to be zero. First, we can have $\beta(k,\omega) = 0$ for particular values of k and ω. This certainly works, but since $\beta(k,\omega)$ indicates how much of $\psi(x,t)$ is made up of $e^{i(kx+\omega t)}$ for these particular values of k and ω, the $\beta(k,\omega) = 0$ statement tells us that this particular $e^{i(kx+\omega t)}$ exponential doesn't appear in $\psi(x,t)$. So we don't care about how ω and k are related.

The other way for Eq. (18) to be zero is if $\omega^2 - c^2k^2 = 0$. That is, $\omega = \pm ck$, as we wanted to show. We therefore see that if $\beta(k,\omega)$ is nonzero for particular values of k and ω (that is, if $e^{i(kx+\omega t)}$ appears in $\psi(x,t)$), then ω must be equal to $\pm ck$, if we want the wave equation to be satisfied.

4.2 Reflection and transmission

4.2.1 Applying the boundary conditions

Instead of an infinite uniform string, let's now consider an infinite string with density μ_1 for $-\infty < x < 0$ and μ_2 for $0 < x < \infty$. Although the density isn't uniform, the tension is still uniform throughout the entire string, because otherwise there would be a nonzero horizontal acceleration somewhere.

Assume that a wave of the form $\psi_{\rm i}(x,t) = f_{\rm i}(x - v_1t)$ (the "i" here is for "incident") starts off far to the left and heads rightward toward $x = 0$. It turns out that it will be much

[2]Alternatively, Eq. (17) says that $\beta(k,\omega)(\omega^2 - c^2k^2)$ is the 2-D Fourier transform of zero. So it must be zero, because it can be found from the 2-D inverse-transform relations analogous to Eq. (3.43), with a zero appearing in the integrand.

more convenient to instead write the wave as

$$\psi_{\rm i}(x,t) = f_{\rm i}\left(t - \frac{x}{v_1}\right) \tag{19}$$

for a redefined function $f_{\rm i}$, so we'll use this form. Note that ψ is a function of two variables, whereas f is a function of only one. From Eq. (5), the speed v_1 equals $\sqrt{T/\mu_1}$.

What happens when the wave encounters the boundary at $x = 0$ between the different densities? The most general thing that can happen is that there is some reflected wave,

$$\psi_{\rm r}(x,t) = f_{\rm r}\left(t + \frac{x}{v_1}\right), \tag{20}$$

moving leftward from $x = 0$, and also a transmitted wave,

$$\psi_{\rm t}(x,t) = f_{\rm t}\left(t - \frac{x}{v_2}\right), \tag{21}$$

moving rightward from $x = 0$ (where $v_2 = \sqrt{T/\mu_2}$). Note the "+" sign in the argument of $f_{\rm r}$, since the reflected wave is moving leftward.

In terms of the above functions, the complete expressions for the waves on the left and right side of $x = 0$ are, respectively,

$$\begin{aligned} \psi_{\rm L}(x,t) &= \psi_{\rm i}(x,t) + \psi_{\rm r}(x,t) = f_{\rm i}(t - x/v_1) + f_{\rm r}(t + x/v_1), \\ \psi_{\rm R}(x,t) &= \psi_{\rm t}(x,t) = f_{\rm t}(t - x/v_2). \end{aligned} \tag{22}$$

If we can find the reflected and transmitted waves in terms of the incident wave, then we will know what the complete wave looks like everywhere. Our goal is therefore to find $\psi_{\rm r}(x,t)$ and $\psi_{\rm t}(x,t)$ in terms of $\psi_{\rm i}(x,t)$. To do this, we will use the two boundary conditions at $x = 0$. Using Eq. (22) to write the waves in terms of the various f functions, the two boundary conditions are:

- The string is continuous. So we must have (for all t)

$$\psi_{\rm L}(0,t) = \psi_{\rm R}(0,t) \quad \Longrightarrow \quad \boxed{f_{\rm i}(t) + f_{\rm r}(t) = f_{\rm t}(t)} \tag{23}$$

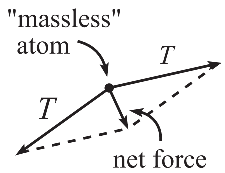


Figure 4

- The slope is continuous. This is true for the following reason. If the slope were different on either side of $x = 0$, then there would be a net (non-infinitesimal) force in some direction on the atom located at $x = 0$, as shown in Fig. 4. This (nearly massless) atom would then experience an essentially infinite acceleration, which isn't physically possible. (Equivalently, a nonzero force would have the effect of instantaneously readjusting the string to a position where the slope was continuous.) Continuity of the slope gives (for all t)

$$\left.\frac{\partial\psi_{\rm L}(x,t)}{\partial x}\right|_{x=0} = \left.\frac{\partial\psi_{\rm R}(x,t)}{\partial x}\right|_{x=0} \quad \Longrightarrow \quad -\frac{1}{v_1}f_{\rm i}'(t) + \frac{1}{v_1}f_{\rm r}'(t) = -\frac{1}{v_2}f_{\rm t}'(t). \tag{24}$$

Integrating this and getting the v's out of the denominators gives

$$\boxed{v_2 f_{\rm i}(t) - v_2 f_{\rm r}(t) = v_1 f_{\rm t}(t)} \tag{25}$$

We have set the constant of integration equal to zero because we are assuming that the string has no displacement before the wave passes by.

Solving Eqs. (23) and (25) for $f_r(t)$ and $f_t(t)$ in terms of $f_i(t)$ gives

$$f_r(s) = \frac{v_2 - v_1}{v_2 + v_1} f_i(s), \qquad \text{and} \qquad f_t(s) = \frac{2v_2}{v_2 + v_1} f_i(s), \tag{26}$$

where we have written the argument as s instead of t to emphasize that these relations hold for *any* arbitrary argument of the f functions. The argument need not have anything to do with the time t. The f's are functions of one variable, and we've chosen to call that variable s here.

4.2.2 Reflection

We derived the relations in Eq. (26) by considering how the $\psi(x,t)$'s relate at $x = 0$. But how do we relate them at other x values? We can do this in the following way. Let's look at the reflected wave first. If we replace the s in Eq. (26) by $t + x/v_1$ (which we are free to do, because the argument of the f's can be whatever we want it to be), then we can write f_r as

$$f_r\left(t + \frac{x}{v_1}\right) = \frac{v_2 - v_1}{v_2 + v_1} f_i\left(t - \frac{-x}{v_1}\right). \tag{27}$$

If we recall the definition of the f's in Eqs. (19-21), we can write this result in terms of the ψ's as

$$\boxed{\psi_r(x,t) = \frac{v_2 - v_1}{v_2 + v_1}\psi_i(-x,t)} \tag{28}$$

This is the desired relation between ψ_r and ψ_i, and its interpretation is the following. It says that at a given time t, the value of ψ_r at position x equals $(v_2 - v_1)/(v_2 + v_1)$ times the value of ψ_i at position *negative* x. This implies that the speed of the ψ_r wave equals the speed of the ψ_i wave (but with opposite velocity), and it also implies that the width of the ψ_r wave equals the width of the ψ_i wave. But the height is decreased by a factor of $(v_2 - v_1)/(v_2 + v_1)$.

Only negative values of x are relevant here, because we are dealing with the reflected wave which exists only to the left of $x = 0$. Therefore, since the expression $\psi_i(-x,t)$ appears in Eq. (28), the $-x$ here means that only *positive* position coordinates are relevant for the ψ_i wave. You might find this somewhat disconcerting, because the ψ_i function isn't applicable to the right of $x = 0$. However, we can mathematically imagine ψ_i still proceeding to the right. So we have the pictures shown in Fig. 5. For simplicity, let's say that $v_2 = 3v_1$, which means that the $(v_2 - v_1)/(v_2 + v_1)$ factor equals $1/2$. Note that in any case, this factor lies between -1 and 1. We'll talk about the various possibilities below.

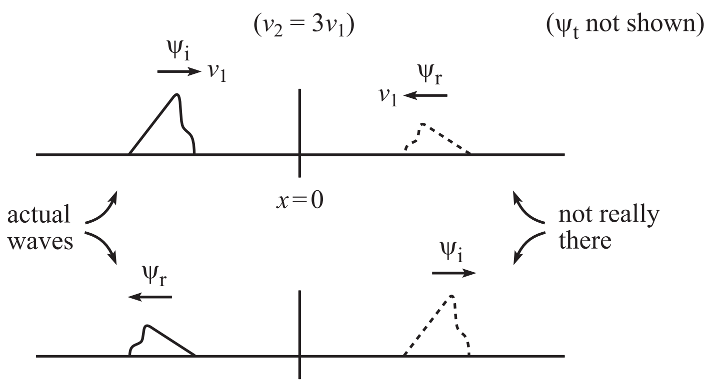


Figure 5

In the first picture in Fig. 5, the incident wave is moving in from the left, and the reflected wave is moving in from the right. The reflected wave doesn't actually exist to the right of $x = 0$, of course, but it's convenient to imagine it coming in as shown. Except for the scale factor of $(v_2 - v_1)/(v_2 + v_1)$ in the vertical direction (only), ψ_r is simply the mirror image of ψ_i.

In the second picture in Fig. 5, the incident wave has passed the origin and continues moving to the right, where it doesn't actually exist. But the reflected wave is now located on the left side of the origin and moves to the left. This is the real piece of the wave. For simplicity, we haven't shown the transmitted ψ_t wave in these pictures (we'll deal with it below), but it's technically also there.

In between the two times shown in Fig. 5, things aren't quite as clean, because there are x values near the origin (to the left of it) where both ψ_i and ψ_r are nonzero, and we need to add them to obtain the complete wave, ψ_L, in Eq. (22). But the procedure is straightforward in principle. The two ψ_i and ψ_r waves simply pass through each other, and the value of ψ_L at any point to the left of $x = 0$ is obtained by adding the values of ψ_i and ψ_r at that point. Remember that only the region to the left of $x = 0$ is real, as far as the reflected wave is concerned. The wave to the right of $x = 0$ that is generated from ψ_i and ψ_r is just a convenient mathematical construct.

Fig. 6 shows some successive snapshots that result from an easy-to-visualize incident square wave. The bold-line wave indicates the actual wave that exists to the left of $x = 0$. We haven't drawn the transmitted wave to the right of $x = 0$. You should stare at this figure until it makes sense. This wave actually isn't physical; its derivative isn't continuous, so it violates the second of the above boundary conditions (although we can imagine rounding off the corners to eliminate this issue). Also, its derivative isn't small, which violates our assumption at the beginning of this section. However, we're drawing this wave so that the important features of reflection can be seen. Throughout this book, if we actually drew realistic waves, the slopes would be so small that it would be nearly impossible to tell what was going on.

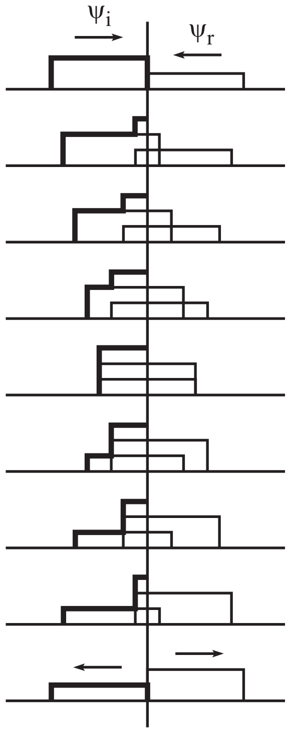


Figure 6

4.2.3 Transmission

Let's now look at the transmitted wave. If we replace the s by $t - x/v_2$ in Eq. (26), we can write f_t as

$$f_\text{t}\left(t - \frac{x}{v_2}\right) = \frac{2v_2}{v_2 + v_1} f_\text{i}\left(t - \frac{(v_1/v_2)x}{v_1}\right). \tag{29}$$

Using the definition of the f's in Eqs. (19-21), we can write this in terms of the ψ's as

$$\boxed{\psi_\text{t}(x,t) = \frac{2v_2}{v_2 + v_1}\psi_\text{i}\Big((v_1/v_2)x, t\Big)} \tag{30}$$

This is the desired relation between ψ_t and ψ_i, and its interpretation is the following. It says that at a given time t, the value of ψ_t at position x equals $2v_2/(v_2 + v_1)$ times the value of ψ_i at position $(v_1/v_2)x$. This implies that the speed of the ψ_t wave is v_2/v_1 times the speed of the ψ_i wave, and it also implies that the width of the ψ_t wave equals v_2/v_1 times the width of the ψ_i wave. These facts are perhaps a little more obvious if we write Eq. (30) as $\psi_\text{t}\big((v_2/v_1)x, t\big) = 2v_2/(v_2 + v_1) \cdot \psi_\text{i}(x, t)$.

Only positive values of x are relevant here, because we are dealing with the transmitted wave which exists only to the right of $x = 0$. Therefore, since the expression $\psi_\text{i}\big((v_1/v_2)x, t\big)$ appears in Eq. (30), only positive position coordinates are relevant for the ψ_i wave. As in the case of reflection above, although the ψ_i function isn't applicable for positive x, we can mathematically imagine ψ_i still proceeding to the right. The situation is shown in Fig. 7. As above, let's say that $v_2 = 3v_1$, which means that the $2v_2/(v_2 + v_1)$ factor equals 3/2. Note that in any case, this factor lies between 0 and 2. We'll talk about the various possibilities below.

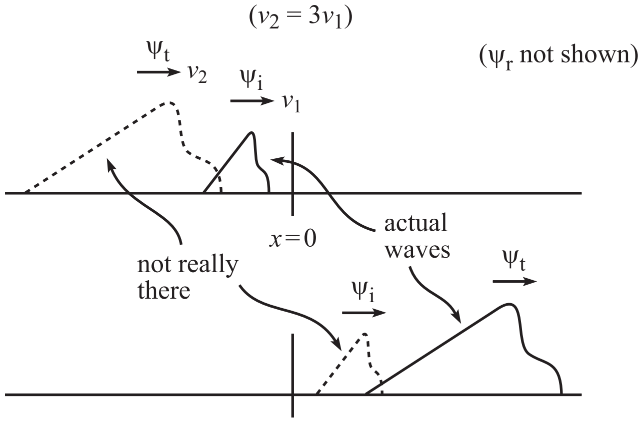


Figure 7

In the first picture in Fig. 7, the incident wave is moving in from the left, and the transmitted wave is also moving in from the left. The transmitted wave doesn't actually exist to the left of $x = 0$, of course, but it's convenient to imagine it coming in as shown. With $v_2 = 3v_1$, the transmitted wave is 3/2 as tall and 3 times as wide as the incident wave.

In the second picture in Fig. 7, the incident wave has passed the origin and continues moving to the right, where it doesn't actually exist. But the transmitted wave is now located on the right side of the origin and moves to the right. This is the real piece of the wave. For simplicity, we haven't shown the reflected ψ_r wave in these pictures, but it's technically also there.

In between the two times shown in Fig. 7, things are easier to deal with than in the reflected case, because we don't need to worry about taking the sum of two waves. The transmitted wave consists only of ψ_t. We don't have to add on ψ_i as we did in the reflected case. In short, ψ_L equals $\psi_i + \psi_r$, whereas ψ_R simply equals ψ_t. Equivalently, ψ_i and ψ_r have physical meaning only to the left of $x = 0$, whereas ψ_t has physical meaning only to the right of $x = 0$.

Fig. 8 shows some successive snapshots that result from the same square wave we considered in Fig. 6. The bold-line wave indicates the actual wave that exists to the right of $x = 0$. We haven't drawn the reflected wave to the left of $x = 0$. We've squashed the x axis relative to Fig. 6, to make a larger region viewable. These snapshots are a bit boring compared with those in Fig. 6, because there is no need to add any waves. As far as ψ_t is concerned on the right side of $x = 0$, what you see is what you get. The entire wave (on both sides of $x = 0$) is obtained by juxtaposing the bold waves in Figs. 6 and 8, after expanding Fig. 8 in the horizontal direction to make the unit sizes the same (so that the ψ_i waves have the same width).

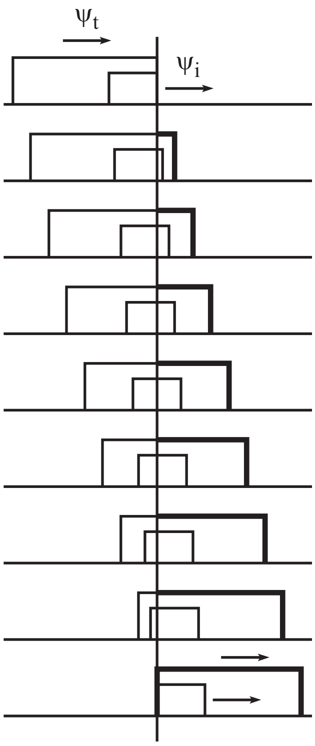


Figure 8

4.2.4 The various possible cases

For convenience, let's define the reflection and transmission coefficients as

$$R \equiv \frac{v_2 - v_1}{v_2 + v_1} \quad \text{and} \quad T \equiv \frac{2v_2}{v_2 + v_1} \tag{31}$$

With these definitions, we can write the reflected and transmitted waves in Eqs. (28) and (30) as

$$\begin{aligned} \psi_{\mathrm{r}}(x,t) &= R\psi_{\mathrm{i}}(-x,t), \\ \psi_{\mathrm{t}}(x,t) &= T\psi_{\mathrm{i}}\big((v_1/v_2)x,t\big). \end{aligned} \tag{32}$$

R and T are the amplitudes of ψ_{r} and ψ_{t} relative to ψ_{i}. Note that $1 + R = T$ always. This is just the statement of continuity of the wave at $x = 0$.

Since $v = \sqrt{T/\mu}$, and since the tension T is uniform throughout the string, we have $v_1 \propto 1/\sqrt{\mu_1}$ and $v_2 \propto 1/\sqrt{\mu_2}$. So we can alternatively write R and T in the terms of the densities on either side of $x = 0$:

$$R \equiv \frac{\sqrt{\mu_1} - \sqrt{\mu_2}}{\sqrt{\mu_1} + \sqrt{\mu_2}} \quad \text{and} \quad T \equiv \frac{2\sqrt{\mu_1}}{\sqrt{\mu_1} + \sqrt{\mu_2}} \tag{33}$$

There are various cases to consider:

- Brick wall on right: $\mu_2 = \infty \;\; (v_2 = 0) \implies R = -1,\; T = 0$. Nothing is transmitted, since $T = 0$. And the reflected wave has the same size as the incident wave, but is inverted due to the $R = -1$ value. This is shown in Fig. 9.

  The inverted nature of the wave isn't intuitively obvious, but it's believable for the following reason. When the wave encounters the wall, the wall pulls down on the string in Fig. 9. This downward force causes the string to overshoot the equilibrium position and end up in the inverted orientation. Of course, you may wonder why the downward force causes the string to overshoot the equilibrium position instead of, say, simply returning to the equilibrium position. But by the same token, you should wonder why a ball that collides elastically with a wall bounces back with the same speed, as opposed to ending up at rest.

  We can deal with both of these situations by invoking conservation of energy. Energy wouldn't be conserved if in the former case the string ended up straight, and if in the latter case the ball ended up at rest, because there would be zero energy in the final state. (The wall is "infinitely" massive, so it can't pick up any energy.)

- Light string on left, heavy string on right: $\mu_1 < \mu_2 < \infty \;\; (v_2 < v_1) \implies -1 < R < 0,\; 0 < T < 1$. This case is in between the previous and following cases. There is partial (inverted) reflection and partial transmission. See Fig. 10 for the particular case where $\mu_2 = 4\mu_1 \implies v_2 = v_1/2$. The reflection and transmission coefficients in this case are $R = -1/3$ and $T = 2/3$.

- Uniform string: $\mu_2 = \mu_1 \;\; (v_2 = v_1) \implies R = 0,\; T = 1$. Nothing is reflected. The string is uniform, so the "boundary" point is no different from any other point. The wave passes right through, as shown in Fig. 11.

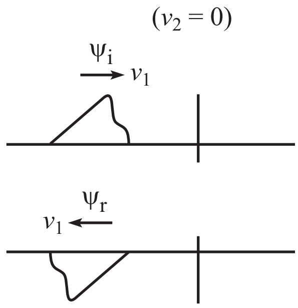


Figure 9

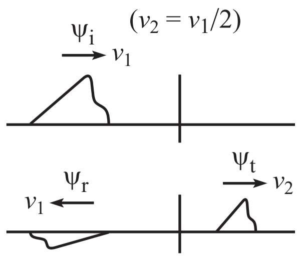


Figure 10

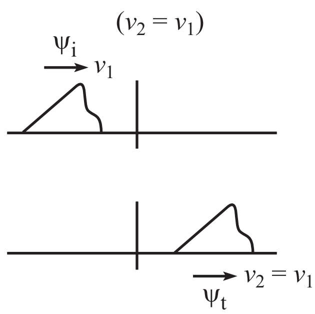


Figure 11

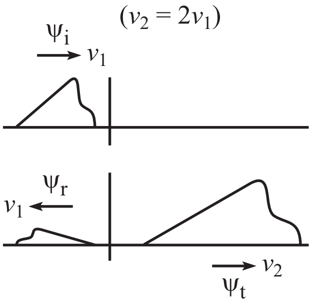


Figure 12

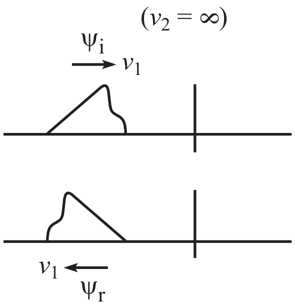


Figure 13

- Heavy string on left, light string on right: $0 < \mu_2 < \mu_1$ $(v_2 > v_1)$ $\Longrightarrow$ $0 < R < 1$, $1 < T < 2$. This case is in between the previous and following cases. There is partial reflection and partial transmission. See Fig. 12 for the particular case where $\mu_2 = \mu_1/4 \Longrightarrow v_2 = 2v_1$. The reflection and transmission coefficients in this case are $R = 1/3$ and $T = 4/3$.

- Zero-mass string on right: $\mu_2 = 0$ $(v_2 = \infty)$ $\Longrightarrow$ $R = 1$, $T = 2$. There is complete (rightside up) reflection in this case, as shown in Fig. 13. Although the string on the right side technically moves, it has zero mass so it can't carry any energy. All of the energy is therefore contained in the reflected wave (we'll talk about energy in Section 4.4). So in this sense there is total reflection. In addition to carrying no energy, the movement of the right part of the string isn't even wave motion. The whole thing always remains in a straight horizontal line and just rises and falls (technically it's a wave with infinite wavelength).[3]

  As with the brick-wall case above, the right-side-up nature of the wave isn't intuitively obvious, but it's believable for the following reason. When the wave encounters the boundary, the zero-mass string on the right side is always horizontal, so it can't apply any transverse force on the string on the left side. Since there is nothing pulling the string down, it can't end up on the other side of the equilibrium position as it did in the brick-wall case. The fact that it actually ends up with the same shape is a consequence of energy conservation, because the massless string on the right can't carry any energy.

4.3 Impedance

4.3.1 Definition of impedance

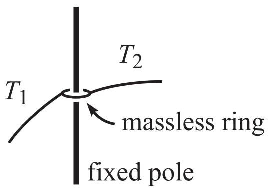


Figure 14

In the previous section, we allowed the density to change at $x = 0$, but we assumed that the tension was uniform throughout the string. Let's now relax this condition and allow the tension to also change at $x = 0$. The previous treatment is easily modified, and we will find that a new quantity, called the *impedance*, arises.

You may be wondering how we can arrange for the tension to change at $x = 0$, given that any difference in the tension should cause the atom at $x = 0$ to have "infinite" acceleration. But we can eliminate this issue by using the setup shown in Fig. 14. The boundary between the two halves of the string is a massless ring, and this ring encircles a fixed frictionless pole. The pole balances the difference in the longitudinal components of the two tensions, so the net longitudinal force on the ring is zero, consistent with the fact that it is constrained to remain on the pole and move only in the transverse direction.

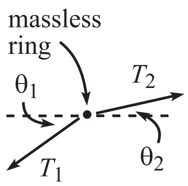


Figure 15

The net *transverse* force on the massless ring must also be zero, because otherwise it would have infinite transverse acceleration. This zero-transverse-force condition is given by $T_1 \sin\theta_1 = T_2 \sin\theta_2$, where the angles are defined in Fig. 15. In terms of the derivatives on either side of $x = 0$, this relation can be written as (assuming, as we always do, that the slope of the string is small)

$$T_1 \frac{\partial \psi_{\mathrm{L}}(x,t)}{\partial x}\bigg|_{x=0} = T_2 \frac{\partial \psi_{\mathrm{R}}(x,t)}{\partial x}\bigg|_{x=0}. \tag{34}$$

[3]The right part of the string must be straight (and hence horizontal, so that it doesn't head off to $\pm\infty$) because if it were curved, then the nonzero second derivative would imply a nonzero force on a given piece of the string, resulting in infinite acceleration, because the piece is massless. Alternatively, the transmitted wave is stretched horizontally by a factor $v_2/v_1 = \infty$ compared with the incident wave. This implies that it is essentially horizontal.

In the case of uniform tension discussed in the previous section, we had $T_1 = T_2$, so this equation reduced to the first equality in Eq. (24). With the tensions now distinct, the only modification to the second equality in Eq. (24) is the extra factors of T_1 and T_2. So in terms of the f functions, Eq. (34) becomes

$$-\frac{T_1}{v_1} f_{\mathrm{i}}'(t) + \frac{T_1}{v_1} f_{\mathrm{r}}'(t) = -\frac{T_2}{v_2} f_{\mathrm{t}}'(t). \tag{35}$$

The other boundary condition (the continuity of the string) is unchanged, so all of the results in the previous section can be carried over, with the only modification being that wherever we had a v_1, we now have v_1/T_1. And likewise for v_2. The quantity v/T can be written as

$$\frac{v}{T} = \frac{\sqrt{T/\mu}}{T} = \frac{1}{\sqrt{T\mu}} \equiv \frac{1}{Z}, \tag{36}$$

where

$$\boxed{Z \equiv \frac{T}{v} = \sqrt{T\mu}} \tag{37}$$

is called the *impedance*. We'll discuss Z in depth below, but for now we'll simply note that the results in the previous section are modified by replacing v_1 with $1/\sqrt{T_1\mu_1} \equiv 1/Z_1$, and likewise for v_2. The reflection and transmission coefficients in Eq. (31) therefore become

$$R = \frac{\frac{1}{Z_2} - \frac{1}{Z_1}}{\frac{1}{Z_2} + \frac{1}{Z_1}} = \boxed{\frac{Z_1 - Z_2}{Z_1 + Z_2}} \qquad \text{and} \qquad T = \frac{\frac{2}{Z_2}}{\frac{1}{Z_2} + \frac{1}{Z_1}} = \boxed{\frac{2Z_1}{Z_1 + Z_2}} \tag{38}$$

Note that Z grows with both T and μ.

Physical meaning of impedance

What is the meaning of the impedance? It makes our formulas look nice, but does it have any actual physical significance? Indeed it does. Consider the transverse force that the ring *applies* to the string on its *left*. Since there is zero net force on the ring, this force also equals the transverse force that the *right* string applies *to* the ring, which is

$$F_y = T_2 \frac{\partial \psi_{\mathrm{R}}(x,t)}{\partial x}\bigg|_{x=0} = T_2 \frac{\partial f_{\mathrm{t}}(t - x/v_2)}{\partial x}\bigg|_{x=0}, \tag{39}$$

where we have labeled the transverse direction as the y direction. But the chain rule tells us that the x and t partial derivatives of f_t are related by

$$\frac{\partial f_{\mathrm{t}}(t - x/v_2)}{\partial x} = -\frac{1}{v_2} \cdot \frac{\partial f_{\mathrm{t}}(t - x/v_2)}{\partial t}. \tag{40}$$

Substituting this into Eq. (39) and switching back to the $\psi_{\mathrm{R}}(x,t)$ notation gives

$$F_y = -\frac{T_2}{v_2} \cdot \frac{\partial \psi_{\mathrm{R}}(x,t)}{\partial t}\bigg|_{x=0} = -\frac{T_2}{v_2} \cdot v_y \equiv -b v_y, \tag{41}$$

where $v_y = \partial\psi_{\mathrm{R}}(x,t)/\partial t$ is the transverse velocity of the ring (at $x = 0$), and where b is defined to be T_2/v_2.

This force F_y (which again, is the force that the ring *applies* to the string on its *left*) has the interesting property that it is proportional to the (negative of the) transverse velocity. It therefore acts exactly like a *damping* force. If we removed the right string and replaced the ring with a massless plate immersed in a fluid (in other words, a piston), and if we arranged things (the thickness of the fluid and the cross-sectional area of the plate) so that the damping coefficient was b, then the left string wouldn't have any clue that the right string was removed and replaced with the damping mechanism. As far as the left string is concerned, the right string acts exactly like a resistance that is being dragged against.

Said in another way, if the left string is replaced by your hand, and if you move your hand so that the right string moves just as it was moving when the left string was there, then you can't tell whether you're driving the right string or driving a piston with an appropriately-chosen damping coefficient. This is true because by Newton's third law (and the fact that the ring is massless), the force that the ring applies to the string on the right is $+bv_y$. The plus sign signifies a driving force (that is, the force is doing positive work) instead of a damping force (where the force does negative work).

From Eq. (41), we have $F_y/v_y = -T_2/v_2$.[4] At different points in time, the ring has different velocities v_y and exerts different forces F_y on the left string. But the ratio F_y/v_y is always equal to $-T_2/v_2$, which is constant, given T_2 and μ_2 (since $v_2 = \sqrt{T_2/\mu_2}$). So, since $F_y/v_y = -T_2/v_2$ is constant in time, it makes sense to give it a name, and we call it the *impedance*, Z. This is consistent with the $Z \equiv T/v$ definition in Eq. (37). From Eq. (41), the impedance Z is simply the damping coefficient b. Large damping therefore means large impedance, so the "impedance" name makes colloquial sense.

Since the impedance $Z \equiv T/v$ equals $\sqrt{T\mu}$ from Eq. (37), it is a property of the string itself (given T and μ), and not of a particular wave motion on the string. From Eq. (38) we see that if $Z_1 = Z_2$, then $R = 0$ and $T = 1$. In other words, there is total transmission. In this case we say that the strings are "impedance matched." We'll talk much more about this below, but for now we'll just note that there are many ways to make the impedances match. One way is to simply have the left and right strings be identical, that is, $T_1 = T_2$ and $\mu_1 = \mu_2$. In this case we have a uniform string, so the wave just moves merrily along and nothing is reflected. But $T_2 = 3T_1$ and $\mu_2 = \mu_1/3$ also yields matching impedances, as do an infinite number of other scenarios. All we need is for the product $T\mu$ to be the same in the two pieces, and then the impedances match and everything is transmitted. However, in these cases, it isn't obvious that there is no reflected wave, as it was for the uniform string. The reason for zero reflection is that the left string can't tell the difference between an identical string on the right, or a piston with a damping coefficient of $\sqrt{T_1\mu_1}$, or a string with $T_2 = 3T_1$ and $\mu_2 = \mu_1/3$. They all feel exactly the same, as far as the left string is concerned; they all generate a transverse force of the form, $F_y = -\sqrt{T_1\mu_1} \cdot v_y$. So if there is no reflection in one case (and there certainly isn't in the case of an identical string), then there is no reflection in any other case. As far as reflection and transmission go, a string is completely characterized by one quantity: the impedance $Z \equiv \sqrt{T\mu}$. Nothing else matters. Other quantities such as T, μ, and $v = \sqrt{T/\mu}$ are relevant for various other considerations, but only the combination $Z \equiv \sqrt{T\mu}$ appears in the R and T coefficients.

Although the word "impedance" makes colloquial sense, there is one connotation that might be misleading. You might think that a small impedance allows a wave to transmit easily and reflect essentially nothing back. But this isn't the case. Maximal transmission occurs when the impedances *match*, not when Z_2 is small. (If Z_2 is small, say $Z_2 = 0$, then Eq. (38) tells us that we actually have *total* reflection with $R = 1$.) When we discuss energy in Section 4.4, we'll see that impedance matching yields maximal energy transfer, consistent with the fact that no energy remains on the left side, because there is no reflected wave.

[4]Remember that F_y and v_y are the *transverse* force and velocity, which are generally very small, given our usual assumption of a small slope of the string. But T_2 and v_2 are the tension and wave speed on the right side, which are "everyday-sized" quantities. What we showed in Eq. (41) was that the two ratios, F_y/v_y and $-T_2/v_2$, are always equal.

Why F_y is proportional to $\sqrt{T\mu}$

We saw above that the transverse force that the left string (or technically the ring at the boundary) applies to the right string is $F_y = +bv_y \equiv Zv_y$. So if you replace the left string with your hand, then $F_y = Zv_y$ is the transverse force than you must apply to the right string to give it the same motion that it had when the left string was there. The impedance Z gives a measure of how hard it is to wiggle the end of the string back and forth. It is therefore reasonable that $Z = \sqrt{T_2\mu_2}$ grows with both T_2 and μ_2. In particular, if μ_2 is large, then more force should be required in order to wiggle the string in a given manner.

However, although this general dependence on μ_2 seems quite intuitive, you have to be careful, because there is a common incorrect way of thinking about things. The reason why the force grows with μ_2 is *not* the same as the reason why the force grows with m in the simple case of a single point mass (with no string or anything else attached to it). In that case, if you wiggle a point mass back and forth, then the larger the mass, the larger the necessary force, due to $F = ma$.

But in the case of a string, if you grab onto the leftmost atom of the right part of the string, then this atom is essentially massless, so your force isn't doing any of the "$F = ma$" sort of acceleration. All your force is doing is simply balancing the transverse component of the T_2 tension that the right string applies to its leftmost atom. This transverse component is nonzero due to the (in general) nonzero slope. So as far as your force is concerned, all that matters are the values of T_2 and the slope. And the *slope* is where the dependence on μ_2 comes in. If μ_2 is large, then $v_2 = \sqrt{T_2/\mu_2}$ is small, which means that the wave in the right string is squashed by a factor of v_2/v_1 compared with the wave on the left string. This then means that the slope of the right part is larger by a factor that is proportional to $1/v_2 = \sqrt{\mu_2/T_2}$, which in turn means that the transverse force is larger. Since the transverse force is proportional to the product of the tension and the slope, we see that it is proportional to $T_2\sqrt{\mu_2/T_2} = \sqrt{T_2\mu_2}$. To sum up: μ_2 affects the impedance not because of an $F = ma$ effect, but rather because μ affects the wave's speed, and hence slope, which then affects the transverse component of the force.

A byproduct of this reasoning is that the dependence of the transverse force on T_2 takes the form of $\sqrt{T_2}$. This comes from the expected factor of T_2 which arises from the fact that the transverse force is proportional to the tension. But there is an additional factor of $1/\sqrt{T_2}$, because the transverse force is also proportional to the slope, which behaves like $1/\sqrt{T_2}$ from the argument in the previous paragraph.

Why F_y is proportional to v_y

We saw above that if the right string is removed and if the ring is attached to a piston with a damping coefficient of $b = \sqrt{T_2\mu_2}$, then the left string can't tell the difference. Either way, the force on the left string takes the form of $-bv_y \equiv -b\dot{y}$. If instead of a piston we attach the ring to a transverse *spring*, then the force that the ring applies to the left string (which equals the force that the spring applies to the ring, since the ring is massless) is $-ky$. And if the ring is instead simply a *mass* that isn't attached to anything, then the force it applies to the left string is $-m\ddot{y}$ (equal and opposite to the $F_y = ma_y$ force that the string applies to the mass). Neither of these scenarios mimics the correct $-b\dot{y}$ force that the right string actually applies.

This $-b\dot{y}$ force from the right string is a consequence of the "wavy-ness" of waves, for the following reason. The transverse force F_y that the right string applies to the ring is proportional to the slope of the wave; it equals the tension times the slope, assuming the slope is small:

$$F_y = T_2\frac{\partial\psi_{\rm R}}{\partial x}\,. \tag{42}$$

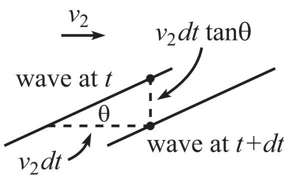


Figure 16

And the transverse velocity is also proportional to the (negative of the) slope, due to the properties of Fig. 16. The left tilted segment is a small piece of the wave at a given time t, and the right tilted segment is the piece at a later time $t + dt$ (the wave is moving to the right). The top dot is the location of a given atom at time t, and the bottom dot is the location of the same atom at time $t + dt$. The wave moves a distance $v_2\,dt$ to the right, so from the triangle shown, the dot moves a distance $(v_2\,dt)\tan\theta$ *downward*. The dot's velocity is therefore $-v_2 \tan\theta$, which is $-v_2$ times the slope. That is,

$$v_y = -v_2 \frac{\partial \psi_{\rm R}}{\partial x}\,. \tag{43}$$

We see that both the transverse force in Eq. (42) and the transverse velocity in Eq. (43) are proportional to the slope, with the constants of proportionality being T_2 and $-v_2$, respectively. The ratio F_y/v_y is therefore independent of the slope. It equals $-T_2/v_2$, which is constant in time.

Fig. 16 is the geometric explanation of the mathematical relation in Eq. (40). Written in terms of $\psi_{\rm R}$ instead of $f_{\rm t}$, Eq. (40) says that

$$\frac{\partial \psi_{\rm R}}{\partial t} = -v_2 \cdot \frac{\partial \psi_{\rm R}}{\partial x}\,. \tag{44}$$

That is, the transverse velocity is $-v_2$ times the slope, which is simply the geometrically-derived result in Eq. (43). Note that this relation holds only for a *single* traveling wave. If we have a wave that consists of, say, two different traveling waves, $\psi(x,t) = f_{\rm a}(t - x/v_{\rm a}) + f_{\rm b}(t - x/v_{\rm b})$, then

$$\begin{aligned} \frac{\partial \psi}{\partial t} &= \frac{\partial f_{\rm a}}{\partial t} + \frac{\partial f_{\rm b}}{\partial t}\,, \\ \frac{\partial \psi}{\partial x} &= -\frac{1}{v_{\rm a}} \cdot \frac{\partial f_{\rm a}}{\partial t} - \frac{1}{v_{\rm b}} \cdot \frac{\partial f_{\rm a}}{\partial t}\,. \end{aligned} \tag{45}$$

Looking at the righthand sides of these equations, we see that it is *not* the case that $\partial\psi/\partial t = -v \cdot \partial\psi/\partial x$ for a particular value of v. It doesn't work for v_1 or v_2 or anything else. This observation is relevant to the following question.

In the discussion leading up to Eq. (41), we considered the transverse force that the ring applies to the string on its *left*, and the result was $-(T_2/v_2)v_y$. From Newton's third law, the transverse force that the ring applies to the string on its *right* is therefore $+(T_2/v_2)v_y$. However, shouldn't we be able to proceed through the above derivation with "left" and "right" reversed, and thereby conclude that the force that the ring applies to the string on its *right* is equal to $+(T_1/v_1)v_y$? The answer had better be a "no," because this result isn't consistent with the $+(T_2/v_2)v_y$ result unless $T_1/v_1 = T_2/v_2$, which certainly doesn't hold for arbitrary choices of strings. Where exactly does the derivation break down? The task of Problem [to be added] is to find out.

4.3.2 Examples of impedance matching

In this section we'll present (without proof) a number of examples of impedance matching. The general point of matching impedances is to yield the maximal energy transfer from one thing to another. As we saw above in the case of the string with different densities, if the impedances are equal, then nothing is reflected. So 100% of the energy is transferred, and you can't do any better than that, due to conservation of energy.

If someone gives you two impedance-matched strings (so the product $T\mu$ is the same in both), then all of the wave is transmitted. If the above conservation-of-energy argument slips your mind, you might (erroneously) think that you could increase the amount of transmitted energy by, say, decreasing μ_2. This has the effect of decreasing Z_2 and thus increasing the transmission coefficient T in Eq. (38). And you might think that the larger amplitude of the transmitted wave implies a larger energy. However, this isn't the case, because less mass is moving now in the right string, due to the smaller value of μ_2. There are competing effects, and it isn't obvious from this reasoning which effect wins. But it *is* obvious from the conservation-of-energy argument. If the Z's aren't equal, then there is nonzero reflection, by Eq. (38), and therefore less than 100% of the energy is transmitted. This can also be demonstrated by explicitly calculating the energy. We'll talk about energy in Section 4.4 below.

The two basic ways to match two impedances are to (1) simply make one of them equal to the other, or (2) keep them as they are, but insert a large number of things (whatever type of things the two original ones are) between them with impedances that gradually change from one to the other. It isn't obvious that this causes essentially all of the energy to be transferred, but this can be shown without too much difficulty. The task of Problem [to be added] is to demonstrate this for the case of the "Gradually changing string density" mentioned below. Without going into too much detail, here are a number of examples of impedance matching:

Electrical circuits: The general expression for impedance is $Z = F/v$. In a purely resistive circuit, the analogs of Z, F, and v are, respectively, the resistance R, the voltage V, and the current I. So $Z = F/v$ becomes $R = V/I$, which is simply Ohm's law. If a source has a given impedance (resistance), and if a load has a variable impedance, then we need to make the load's impedance equal to the source's impedance, if the goal is to have the maximum power delivered to the load. In general, impedances can be complex if the circuit isn't purely resistive, but that's just a fancy way of incorporating a phase difference when V and I (or F and v in general) aren't in phase.

Gradually changing string density: If we have two strings with densities μ_1 and μ_2, and if we insert between them a long string (much longer than the wavelength of the wave being used) whose density gradually changes from μ_1 to μ_2, then essentially all of the wave is transmitted. See Problem [to be added].

Megaphone: A tapered megaphone works by the same principle. If you yell into a simple cylinder, then it turns out that the abrupt change in cross section (from a radius of r to a radius of essentially infinity) causes reflection. The impedance of a cavity depends on the cross section. However, in a megaphone the cross section varies gradually, so not much sound is reflected back toward your mouth. The same effect is relevant for a horn.

Ultrasound: The gel that is put on your skin has the effect of impedance matching the waves in the device to the waves in your body.

Ball collisions: Consider a marble that collides elastically with a bowling ball. The marble will simply bounce off and give essentially none of its energy to the bowling ball. All of the energy will remain in the marble. (And similarly, if a bowling ball collides with a marble, then the bowling ball will simply plow through and keep essentially all of the energy.) But if a series of many balls, each with slightly increasing size, is placed between them (see Fig. 17), then it turns out that essentially all of the marble's energy will end up in the bowling ball. Not obvious, but true. And conversely, if the bowling ball is the one that is initially moving (to the left), then essentially all of its energy will end up in the marble, which will therefore be moving very fast.

Figure 17

It is nebulous what impedance means for one-time evens like collisions between balls, because we defined impedance for waves. But the above string of balls is certainly similar to a longitudinal series of masses (with increasing size) and springs. The longitudinal waves that travel along this spring/mass system consist of many "collisions" between the masses. In the original setup with just the string of balls and no springs, when two balls collide they smush a little and basically act like springs. Well, sort of; they can only repel and not attract. At any rate, if you abruptly increased the size of the masses in the spring/mass system by a large factor, then not much of the wave would make it through. But gradually increasing the masses would be just like gradually increasing the density μ in the "Gradually changing string density" example above.

Lever: If you try to lift a refrigerator that is placed too far out on a lever, you're not going to be able to do it. If you jumped on your end, you would just bounce off like on a springboard. You'd keep all of the energy, and none of it would be transmitted. But if you move the refrigerator inward enough, you'll be able to lift it. However, if you move it in too far (let's assume it's a point mass), then you're back to essentially not being able to lift it, because you'd have to move your end of the lever, say, a mile to lift the refrigerator up by a foot. So there is an optimal placement.

Bicycle: The gears on a bike act basically the same way as a lever. If you're in too high a gear, you'll find that it's too hard on your muscles; you can't get going fast. And likewise, if you're in too low a gear, your legs will just spin wildly, and you'll be able to go only so fast. There is an optimal gear ratio that allows you to transfer the maximum amount of energy from chemical potential energy (from your previous meal) to kinetic energy.

Rolling a ball up a ramp: This is basically just like a lever or a bike. If the ramp is too shallow, then the ball doesn't gain much potential energy. And if it's too steep, then you might not be able to move the ball at all.

4.4 Energy

Energy

What is the energy of a wave? Or more precisely, what is the energy density per unit length? Consider a little piece of the string between x and $x + dx$. In general, this piece has both kinetic and potential energy. The kinetic energy comes from the transverse motion (we showed in the paragraph following Eq. (1) that the longitudinal motion is negligible), so it equals

$$K_{dx} = \frac{1}{2}(dm)v_y^2 = \frac{1}{2}(\mu\, dx)\left(\frac{\partial\psi}{\partial t}\right)^2. \tag{46}$$

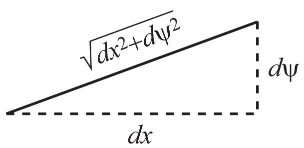


Figure 18

We have used the fact that since there is essentially no longitudinal motion, the mass within the span from x to $x + dx$ is always essentially equal to $\mu\, dx$.

The potential energy depends on the stretch of the string. In general, a given piece of the string is tilted and looks like the piece shown in Fig. 18. As we saw in Eq. (1), the Taylor series $\sqrt{1+\epsilon} \approx 1 + \epsilon/2$ gives the length of the piece as

$$dx\sqrt{1+\left(\frac{\partial\psi}{\partial x}\right)^2} \approx dx + \frac{dx}{2}\left(\frac{\partial\psi}{\partial x}\right)^2. \tag{47}$$

The piece is therefore stretched by an amount, $d\ell \approx (dx/2)(\partial\psi/\partial x)^2$.[5] This stretch is caused by the external tension forces at the two ends. These forces do an amount $T d\ell$ of

[5]Actually, I think this result is rather suspect, although it doesn't matter in the end. See the remark below.

work, and this work shows up as potential energy in the piece, exactly in the same way that a normal spring acquires potential energy if you grab an end and stretch it. So the potential energy of the piece is

$$U_{dx} = \frac{1}{2}T dx \left(\frac{\partial\psi}{\partial x}\right)^2. \tag{48}$$

The total energy per unit length (call it $\mathcal{E}$) is therefore

$$\begin{aligned}\mathcal{E}(x,t) = \frac{K_{dx}+U_{dx}}{dx} &= \frac{\mu}{2}\left(\frac{\partial\psi}{\partial t}\right)^2 + \frac{T}{2}\left(\frac{\partial\psi}{\partial x}\right)^2 \\ &= \frac{\mu}{2}\left[\left(\frac{\partial\psi}{\partial t}\right)^2 + \frac{T}{\mu}\left(\frac{\partial\psi}{\partial x}\right)^2\right] \\ &= \frac{\mu}{2}\left[\left(\frac{\partial\psi}{\partial t}\right)^2 + v^2\left(\frac{\partial\psi}{\partial x}\right)^2\right],\end{aligned} \tag{49}$$

where we have used $v = \sqrt{T/\mu}$. This expression for $\mathcal{E}(x,t)$ is valid for an arbitrary wave. But let's now look at the special case of a single traveling wave, which takes the form of $\psi(x,t) = f(x \pm vt)$. For such a wave, the energy density can be further simplified. As we have seen many times, the partial derivatives of a single traveling wave are related by $\partial\psi/\partial t = \pm v\,\partial\psi/\partial x$. So the two terms in the expression for $\mathcal{E}(x,t)$ are equal at a given point and at a given time. We can therefore write

$$\boxed{\mathcal{E}(x,t) = \mu\left(\frac{\partial\psi}{\partial t}\right)^2} \quad \text{or} \quad \boxed{\mathcal{E}(x,t) = \mu v^2\left(\frac{\partial\psi}{\partial x}\right)^2} \quad \text{(for traveling waves)} \tag{50}$$

Or equivalently, we can use $Z \equiv \sqrt{T\mu}$ and $v = \sqrt{T/\mu}$ to write

$$\mathcal{E}(x,t) = \frac{Z}{v}\left(\frac{\partial\psi}{\partial t}\right)^2, \quad \text{or} \quad \mathcal{E}(x,t) = Zv\left(\frac{\partial\psi}{\partial x}\right)^2. \tag{51}$$

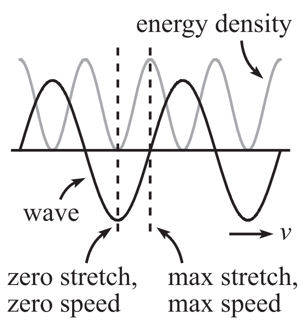


Figure 19

For sinusoidal traveling waves, the energy density is shown in Fig. 19 (with arbitrary units on the axes). The energy-density curve moves right along with the wave.

Remark : As mentioned in Footnote 5, the length of the string given in Eq. (47) and the resulting expression for the potential energy given in Eq. (48) are highly suspect. The reason is the following. In writing down Eq. (47), we made the assumption that all points on the string move in the transverse direction; we assumed that the longitudinal motion is negligible. This is certainly the case (in an exact sense) if the string consists of little masses that are hypothetically constrained to ride along rails pointing in the transverse direction. If these masses are connected by little stretchable pieces of massless string, then Eq. (48) correctly gives the potential energy.

However, all that we were able to show in the reasoning following Eq. (1) was that the points on the string move in the transverse direction, *up to errors of order* $dx(\partial\psi/\partial x)^2$. We therefore have no right to trust the result in Eq. (48), because it is of the *same* order. But even if this result is wrong and if the stretching of the string is distributed differently from Eq. (47), the *total* amount of stretching is the same. Therefore, because the work done on the string, $T d\ell$, is linear in $d\ell$, the total potential energy is independent of the particular stretching details. The $(T/2)(\partial\psi/\partial x)^2$ result therefore correctly yields the *average* potential energy density, even though it may be incorrect at individual points. And since we will rarely be concerned with more than the average, we can use Eq. (48), and everything is fine. ♣

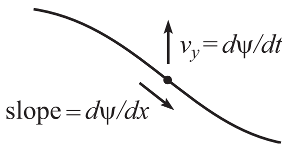


Figure 20

Power

What is the power transmitted across a given point on the string? Equivalently, what is the rate of energy flow past a given point? Equivalently again, at what rate does the string to the left of a point do work on the string to the right of the point? In Fig. 20, the left part of the string pulls on the dot with a transverse force of $F_y = -T\,\partial\psi/\partial x$. The power flow across the dot (with rightward taken to be positive) is therefore

$$P(x,t) = \frac{dW}{dt} = \frac{F_y\,d\psi}{dt} = F_y\frac{\partial\psi}{\partial t} = Fv_y = \left(-T\frac{\partial\psi}{\partial x}\right)\left(\frac{\partial\psi}{\partial t}\right). \tag{52}$$

This expression for $P(x,t)$ is valid for an arbitrary wave. But as with the energy density above, we can simplify the expression if we consider the special case of single traveling wave. If $\psi(x,t) = f(x \pm vt)$, then $\partial\psi/\partial x = \pm(1/v)\partial\psi/\partial t$. So we can write the power as (using $T/v = T/\sqrt{T/\mu} = \sqrt{T\mu} \equiv Z$)

$$P(x,t) = \mp\frac{T}{v}\left(\frac{\partial\psi}{\partial t}\right)^2 \quad\Longrightarrow\quad \boxed{P(x,t) = \mp Z\left(\frac{\partial\psi}{\partial t}\right)^2 = \mp v\mathcal{E}(x,t)} \tag{53}$$

where we have used Eq. (51). We see that the magnitude of the power is simply the wave speed times the energy density. It is positive for a rightward traveling wave $f(x - vt)$, and negative for a leftward traveling wave $f(x+vt)$ (we're assuming that v is a positive quantity here). This makes sense, because the energy plot in Fig. 19 just travels along with the wave at speed v.

Momentum

A wave on a string carries energy due to the transverse motion. Does such a wave carry momentum? Well, there is certainly nonzero momentum in the transverse direction, but it averages out to zero because half of the string is moving one way, and half is moving the other way.

What about the longitudinal direction? We saw above that the points on the string move only negligibly in the longitudinal direction, so there is no momentum in that direction. Even though a traveling wave makes it look like things are moving in the longitudinal direction, there is in fact no such motion. Every point in the string is moving only in the transverse direction. Even if the points did move non-negligible distances, the momentum would still average out to zero, consistent with the fact that there is no overall longitudinal motion of the string. The general kinematic relation $p = mv_{\text{CM}}$ holds, so if the CM of the string doesn't move, then the string has no momentum.

There are a few real-world examples that might make you think that standard traveling waves can carry momentum. One example is where you try (successfully) to move the other end of a rope (or hose, etc.) you're holding, which lies straight on the ground, by flicking the rope. This causes a wave to travel down the rope, which in turn causes the other end to move farther away from you. Everyone has probably done this at one time or another, and it works. However, you can bet that you moved your hand forward during the flick, and this is what gave the rope some longitudinal momentum. You certainly must have moved your hand forward, of course, because otherwise the far end couldn't have gotten farther away from you (assuming that the rope can't stretch significantly). If you produce a wave on a rope by moving your hand only up and down (that is, transversely), then the rope will not have any longitudinal momentum.

Another example that might make you think that waves carry momentum is the case of sound waves. Sound waves are longitudinal waves, and we'll talk about these in the following chapter. But for now we'll just note that if you're standing in front of a very large speaker (large enough so that you feel the sound vibrations), then it *seems* like the sound is applying a net force on you. But it isn't. As we'll see in the next chapter, the pressure on you alternates sign, so half the time the sound wave is pushing you away from the speaker, and half the time it's drawing you closer. So it averages out to zero.[6]

An exception to this is the case of a pulse or an explosion. In this case, something at the source of the pulse must have actually moved forward, so there is some net momentum. If we have only half (say, the positive half) of a cycle of a sinusoidal pressure wave, then this part can push you away without the (missing) other half drawing you closer. But this isn't how normal waves (with both positive and negative parts) work.

4.5 Standing waves

4.5.1 Semi-infinite string

Fixed end

Consider a leftward-moving sinusoidal wave that is incident on a brick wall at its left end, located at $x = 0$. (We've having the wave move leftward instead of the usual rightward for later convenience.) The most general form of a leftward-moving sinusoidal wave is

$$\psi_{\rm i}(x,t) = A\cos(\omega t + kx + \phi), \tag{54}$$

where ω and k satisfy $\omega/k = \sqrt{T/\mu} = v$. A is the amplitude, and ϕ depends on the arbitrary choice of the $t = 0$ time. Since the brick wall has "infinite" impedance, Eq. (38) gives $R = -1$. Eq. (32) then gives the reflected rightward-moving wave as

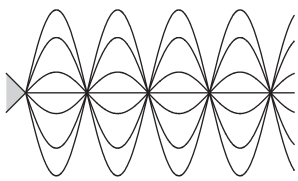
$$\psi_{\rm r}(x,t) = R\psi_{\rm i}(-x,t) = -A\cos(\omega t - kx + \phi). \tag{55}$$

The total wave is therefore

$$\begin{aligned}\psi(x,t) = \psi_{\rm i}(x,t) + \psi_{\rm r}(x,t) &= A\cos(\omega t + \phi + kx) - A\cos(\omega t + \phi - kx) \\ &= \boxed{-2A\sin(\omega t + \phi)\sin kx}\end{aligned} \tag{56}$$

As a double check, this satisfies the boundary condition $\psi(0,t) = 0$ for all t. The sine function of x is critical here. A cosine function wouldn't satisfy the boundary condition at $x = 0$. In contrast with this, it doesn't matter whether we have a sine or cosine function of t, because a phase shift in ϕ can turn one into the other.

For a given value of t, a snapshot of this wave is a sinusoidal function of x. The wavelength of this function is $\lambda = 2\pi/k$, and the amplitude is $|2A\sin(\omega t + \phi)|$. For a given value of x, each point oscillates as a sinusoidal function of t. The period of this function is $\tau = 2\pi/\omega$, and the amplitude is $|2A\sin kx|$. Points for which kx equals $n\pi$ always have $\psi(x,t) = 0$, so they never move. These points are called nodes.



(fixed left end)

Figure 21

Fig. 21 shows the wave at a number of different times. A wave such as the one in Eq. (56) is called a "standing wave." All points on the string have the same phase (or differ by π), as far as the oscillations in time go. That is, all of the points come to rest at the same time (at the maximal displacement from equilibrium), and they all pass through the origin at the same time, etc. This is *not* true for traveling waves. In particular, in a traveling

[6]The opening scene from the first *Back to the Future* movie involves Marty McFly standing in front of a humongous speaker with the power set a bit too high. He then gets blown backwards when he plays a chord. This isn't realistic, but ingenious movies are allowed a few poetic licenses.

wave, the points with $\psi = 0$ are moving with maximal speed, and the points with maximum ψ are instantaneously at rest.

If you don't want to have to invoke the $R = -1$ coefficient, as we did above, another way of deriving Eq. (56) is to apply the boundary condition at $x = 0$ to the most general form of the wave given in Eq. (8). Since $\psi(0, t) = 0$ for all t, we can have only the $\sin kx$ terms in Eq. (8). Therefore,

$$\begin{aligned} \psi(x,t) &= D_2 \sin kx \sin \omega t + D_3 \sin kx \cos \omega t \\ &= (D_2 \sin \omega t + D_3 \cos \omega t) \sin kx \\ &\equiv B \sin(\omega t + \phi) \sin kx, \end{aligned} \tag{57}$$

where B and ϕ are determined by $B\cos\phi = D_2$ and $B \sin\phi = D_3$.

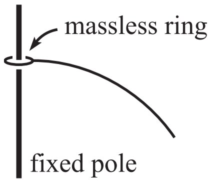


Figure 22

Free end

Consider now a leftward-moving sinusoidal wave that has its left end free (located at $x = 0$). By "free" here, we mean that a massless ring at the end of the string is looped around a fixed frictionless pole pointing in the transverse direction; see Fig. 22. So the end is free to move transversely but not longitudinally. The pole makes it possible to maintain the tension T in the string. Equivalently, you can consider the string to be infinite, but with a density of $\mu = 0$ to the left of $x = 0$.

As above, the most general form of a leftward-moving sinusoidal wave is

$$\psi_{\rm i}(x,t) = A\cos(\omega t + kx + \phi), \tag{58}$$

Since the massless ring (or equivalently the $\mu = 0$ string) has zero impedance, Eq. (38) gives $R = +1$. Eq. (32) then gives the reflected rightward-moving wave as

$$\psi_{\rm r}(x,t) = R\psi_{\rm i}(-x,t) = +A\cos(\omega t - kx + \phi). \tag{59}$$

The total wave is therefore

$$\begin{aligned} \psi(x,t) = \psi_{\rm i}(x,t) + \psi_{\rm r}(x,t) &= A\cos(\omega t + \phi + kx) + A\cos(\omega t + \phi - kx) \\ &= \boxed{2A\cos(\omega t + \phi)\cos kx} \end{aligned} \tag{60}$$

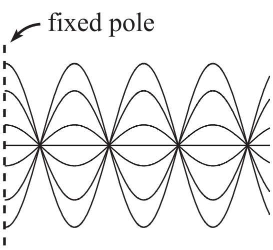


Figure 23

As a double check, this satisfies the boundary condition, $\partial\psi/\partial x|_{x=0} = 0$ for all t. The slope must always be zero at $x = 0$, because otherwise there would be a net transverse force on the massless ring, and hence infinite acceleration. If we choose to construct this setup with a $\mu = 0$ string to the left of $x = 0$, then this string will simply rise and fall, always remaining horizontal. You can assume that the other end is attached to something very far to the left of $x = 0$.

Fig. 23 shows the wave at a number of different times. This wave is similar to the one in Fig. 21 (it has the same amplitude, wavelength, and period), but it is shifted a quarter cycle in both time and space. The time shift isn't of too much importance, but the space shift is critical. The boundary at $x = 0$ now corresponds to an "antinode," that is, a point with the maximum oscillation amplitude.

As in the fixed-end case, if you don't want to have to invoke the $R = 1$ coefficient, you can apply the boundary condition, $\partial\psi/\partial x|_{x=0} = 0$, to the most general form of the wave given in Eq. (8). This allows only the $\cos kx$ terms.

In both this case and the fixed-end case, ω and k can take on a continuous set of values, as long as they are related by $\omega/k = \sqrt{T/\mu} = v$.[7] In the finite-string cases below, we will find that they can take on only discrete values.

[7] Even though the above standing waves don't travel anywhere and thus don't have a speed, it still makes sense to label the quantity $\sqrt{T/\mu}$ as v, because standing waves can be decomposed into two oppositely-moving traveling waves, as shown in Eqs. (56) and (60).

4.5.2 Finite string

We'll now consider the three possible (extreme) cases for the boundary conditions at the two ends of a finite string. We can have both ends fixed, or one fixed and one free, or both free. Let the ends be located at $x = 0$ and $x = L$. In general, the boundary conditions (for all t) are $\psi = 0$ at a fixed end (because the end never moves), and $\partial\psi/\partial x = 0$ at a free end (because the slope must be zero so that there is no transverse force on the massless endpoint).

Two fixed ends

If both ends of the string are fixed, then the two boundary conditions are $\psi(0,t) = 0$ and $\psi(L,t) = 0$ for all t. Eq. (56) gives the most general form of a wave (with particular ω and k values) satisfying the first of these conditions. So we just need to demand that the second one is also true. If we plug $x = L$ into Eq. (56), the only way to have $\psi(L,t) = 0$ for all t is to have $\sin kL = 0$. This implies that kL must equal $n\pi$ for some integer n. So

$$k_n = \frac{n\pi}{L}\,, \tag{61}$$

where we have added the subscript to indicate that k can take on a discrete set of values associated with the integers n. n signifies which "mode" the string is in. The wavelength is $\lambda_n = 2\pi/k_n = 2L/n$. So the possible wavelengths are all integral divisors of $2L$ (which is twice the length of the string). Snapshots of the first few modes are shown below in the first set of waves in Fig. 24. The n values technically start at $n = 0$, but in this case ψ is identically equal to zero since $\sin(0) = 0$. This is certainly a physically possible location of the string, but it is the trivial scenario. So the n values effectively start at 1.

The angular frequency ω is still related to k by $\omega/k = \sqrt{T/\mu} = v$, so we have $\omega_n = vk_n$. Remember that v depends only on T and μ, and not on n (although this won't be true when we get to dispersion in Chapter 6). The frequency in Hertz is then

$$\nu_n = \frac{\omega_n}{2\pi} = \frac{vk_n}{2\pi} = \frac{v(n\pi/L)}{2\pi} = \frac{nv}{2L}\,. \tag{62}$$

The possible frequencies are therefore integer multiples of the "fundamental" frequency, $\nu_1 = v/2L$. In short, the additional boundary condition at the right end constrains the system so that only discrete values of ω and k are allowed. Physically, the wave must undergo an integral number of half oscillations in space, in order to return to the required value of zero at the right end. This makes it clear that λ (and hence k) can take on only discrete values. And since the ratio ω/k is fixed to be $\sqrt{T/\mu}$, this means that ω (and ν) can take on only discrete values too. To summarize:

$$\boxed{\lambda_n = \frac{2L}{n} \qquad \text{and} \qquad \nu_n = \frac{nv}{2L}} \tag{63}$$

The product of these quantities is $\lambda_n\nu_n = v$, as it should be.

Since the wave equation in Eq. (4) is linear, the most general motion of a string with two fixed ends is an arbitrary linear combination of the solutions in Eq. (56), with the restriction that k takes the form $k_n = n\pi/L$ (and ω/k must equal v). So the most general expression for $\psi(x,t)$ is (we'll start the sum at $n = 0$, even though this term doesn't contribute anything)

$$\boxed{\psi(x,t) = \sum_{n=0}^{\infty} B_n \sin(\omega_n t + \phi_n)\sin k_n x} \qquad \text{where } k_n = \frac{n\pi}{L}\,, \quad \text{and} \quad \omega_n = vk_n. \tag{64}$$

The B here equals the $-2A$ from Eq. (56). Note that the amplitudes and phases of the various modes can in general be different.

One fixed end, one free end

Now consider the case where one end is fixed and the other is free. Let's take the fixed end to be the one at $x = 0$. The case where the fixed end is located at $x = L$ gives the same general result; it's just the mirror image of the result we will obtain here.

The two boundary conditions are $\psi(0, t) = 0$ and $\partial\psi/\partial x|_{x=L} = 0$ for all t. Eq. (56) again gives the most general form of a wave satisfying the first of these conditions. So we just need to demand that the second one is also true. From Eq. (56), the slope $\partial\psi/\partial x$ is proportional to $\cos kx$. The only way for this to be zero at $x = L$ is for kL to equal $(n + 1/2)\pi$ for some integer n. So

$$k_n = \frac{(n + 1/2)\pi}{L}\,. \tag{65}$$

n starts at zero here. Unlike in the two-fixed-ends case, the $n = 0$ value now gives a nontrivial wave.

The wavelength is $\lambda_n = 2\pi/k_n = 2L/(n + 1/2)$. These wavelengths are most easily seen in pictures, and snapshots of the first few modes are shown below in the second set of waves in Fig. 24. The easiest way to describe the wavelengths in words is to note that the number of oscillations that fit on the string is $L/\lambda_n = n/2 + 1/4$. So for the lowest mode (the $n = 0$ one), a quarter of a wavelength fits on the string. The higher modes are then obtained by successively adding on half an oscillation (which ensures that $x = L$ is always located at an antinode, with slope zero). The frequency ν_n can be found via Eq. (62), and we can summarize the results:

$$\boxed{\lambda_n = \frac{2L}{n + 1/2} \qquad \text{and} \qquad \nu_n = \frac{(n + 1/2)v}{2L}} \tag{66}$$

Similar to Eq. (64), the most general motion of a string with one fixed end and one free end is a linear combination of the solutions in Eq. (56):

$$\boxed{\psi(x, t) = \sum_{n=0}^{\infty} B_n \sin(\omega_n t + \phi_n) \sin k_n x} \qquad \text{where} \;\; k_n = \frac{(n + 1/2)\pi}{L}\,, \quad \text{and} \quad \omega_n = v k_n. \tag{67}$$

If we instead had the left end as the free one, then Eq. (60) would be the relevant equation, and the $\sin kx$ here would instead be a $\cos kx$. As far as the dependence on time goes, it doesn't matter whether it's a sine or cosine function of t, because a redefinition of the $t = 0$ point yields a phase shift that can turn sines into cosines, and vice versa. We aren't free to redefine the $x = 0$ point, because we have a physical wall there.

Two free ends

Now consider the case with two free ends. The two boundary conditions are $\partial\psi/\partial x|_{x=0} = 0$ and $\partial\psi/\partial x|_{x=L} = 0$ for all t. Eq. (60) gives the most general form of a wave satisfying the first of these conditions. So we just need to demand that the second one is also true. From Eq. (60), the slope $\partial\psi/\partial x$ is proportional to $\sin kx$. The only way for this to be zero at $x = L$ is for kL to equal $n\pi$ for some integer n. So

$$k_n = \frac{n\pi}{L}\,, \tag{68}$$

which is the same as in the case of two fixed ends. The wavelength is $\lambda_n = 2\pi/k_n = 2L/n$. So the possible wavelengths are all integral divisors of $2L$, again the same as in the case of two fixed ends. Snapshots of the first few modes are shown below in the third set of waves in Fig. 24.

The n values technically start at $n = 0$. In this case ψ has no dependence on x, so we simply have a flat line (not necessarily at $\psi = 0$). The line just sits there at rest, because the frequency is again given by Eq. (62) and is therefore zero (the k_n values, and hence ω_n values, are the same as in the two-fixed-ends case, so $\omega_0 = 0$). This case isn't as trivial as the $n = 0$ case for two fixed ends; the resulting constant value of ψ might be necessary to satisfy the initial conditions of the string. This constant value is analogous to the a_0 term in the Fourier-series expression in Eq. (3.1).

As with two fixed ends, an integral number of half oscillations must fit into L, but they now start and end at antinodes instead of nodes. The frequency ν_n is the same as in the two-fixed-ends case, so as in that case we have:

$$\boxed{\lambda_n = \frac{2L}{n} \qquad \text{and} \qquad \nu_n = \frac{nv}{2L}} \tag{69}$$

Similar to Eqs. (64) and (67), the most general motion of a string with two free ends is a linear combination of the solutions in Eq. (60):

$$\boxed{\psi(x,t) = \sum_{n=0}^{\infty} B_n \cos(\omega_n t + \phi_n) \cos k_n x} \qquad \text{where} \quad k_n = \frac{n\pi}{L}, \quad \text{and} \quad \omega_n = vk_n. \tag{70}$$

Fig. 24 summarizes the above results.

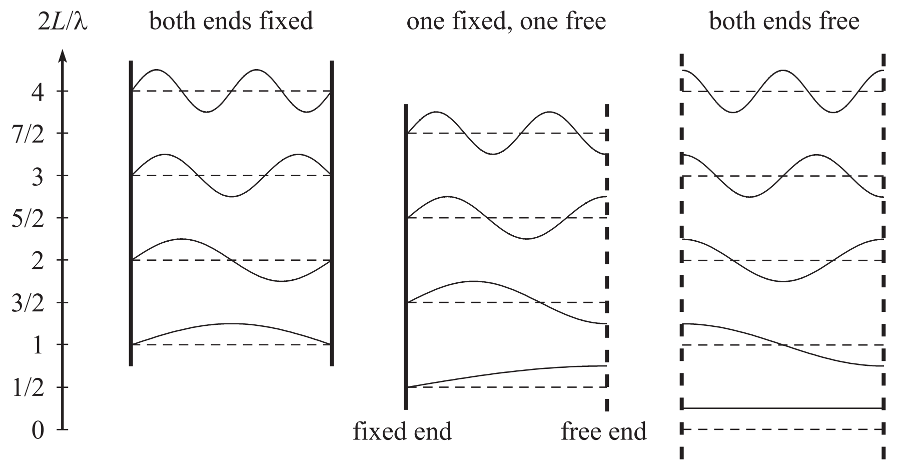


Figure 24

Power in a standing wave

We saw above in Section 4.4 that not only do traveling waves contain energy, they also contain an energy flow along the string. That is, they transmit power. A given point on the string does work (which may be positive or negative, depending on the direction of the wave's velocity) on the part of the string to its right. And it does the opposite amount of work on the string to its left.

A reasonable question to ask now is: Is there energy flow in standing waves? There is certainly an energy *density*, because in general the string both moves and stretches. But is there any energy transfer along the string?

Intuitively, a standing wave is the superposition of two oppositely-moving traveling waves with equal amplitudes. These traveling waves have equal and opposite energy flow, on average, so we expect the net energy flow in a standing wave to be zero, on average. (We'll see below where this "on average" qualification arises.) Alternatively, we can note that there can't be any net energy flow in either direction in a standing wave, due to the left-right symmetry of the system. If you flip the paper over, so that right and left are reversed, then a standing wave looks exactly the same, whereas a traveling wave doesn't, because it's now moving in the opposite direction.

Mathematically, we can calculate the energy flow (that is, the power) as follows. The expression for the power in Eq. (52) is valid for an arbitrary wave. This is the power flow across a given point, with rightward taken to be positive. Let's see what Eq. (52) reduces to for a standing wave. We'll take our standing wave to be $A\sin\omega t\sin kx$. (We could have any combination of sines and cosines here; they all give the same general result.) Eq. (52) becomes

$$\begin{aligned} P(x,t) &= \left(-T\frac{\partial\psi}{\partial x}\right)\left(\frac{\partial\psi}{\partial t}\right) \\ &= -T\big(kA\sin\omega t\cos kx\big)\big(\omega A\cos\omega t\sin kx\big) \\ &= -TA^2\omega k\big(\sin kx\cos kx\big)\big(\sin\omega t\cos\omega t\big). \end{aligned} \tag{71}$$

In general, this is nonzero, so there *is* energy flow across a given point. However, at a given value of x, the average (over one period) of $\sin\omega t\cos\omega t$ (which equals $(1/2)\sin 2\omega t$) is zero. So the *average* power is zero, as we wanted to show.

The difference between a traveling wave and a standing wave is the following. Mathematically: for a traveling wave of the form $A\cos(kx-\omega t)$, the two derivatives in Eq. (71) produce the same $\sin(kx-\omega t)$ function, so we end up with the *square* of a function, which is always positive. There can therefore be no cancelation. But for a standing wave, Eq. (71) yields a bunch of different functions and no squares, and they average out to zero.

Physically: in a traveling wave, the transverse force that a given dot on the string applies to the string on its right is always in phase (or 180° out of phase, depending on the direction of the wave's motion) with the velocity of the dot. This is due to the fact that $\partial\psi/\partial x$ is proportional to $\partial\psi/\partial t$ for a traveling wave. So the power, which is the product of the transverse force and the velocity, always has the same sign. There is therefore never any cancelation between positive and negative amounts of work being done.

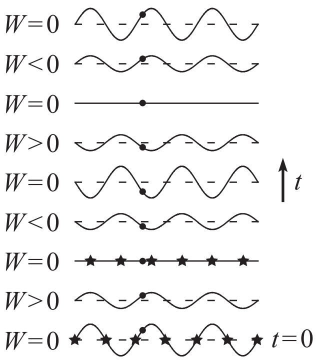


Figure 25

However, for the standing wave $A\sin\omega t\sin kx$, the transverse force is proportional to $-\partial\psi/\partial x = -kA\sin\omega t\cos kx$, while the velocity is proportional to $\partial\psi/\partial t = \omega A\cos\omega t\sin kx$. For a given value of x, the $\sin kx$ and $\cos kx$ functions are constant, so the t dependence tells us that the transverse force is 90° ahead of the velocity. So half the time the force is in the same direction as the velocity, and half the time it is in the opposite direction. The product integrates to zero, as we saw in Eq. (71).

The situation is summarized in Fig. 25, which shows a series of nine snapshots throughout a full cycle of a standing wave. W is the work that the dot on the string does on the string to its *right*. Half the time W is positive, and half the time it is negative. The stars show the points with maximum energy density. When the string is instantaneously at rest at maximal curvature, the nodes have the greatest energy density, in the form of potential energy. The nodes are stretched maximally, and in contrast there is never any stretching at the antinodes. There is no kinetic energy anywhere in the string. A quarter cycle later, when the string is straight and moving quickest, the antinodes have the greatest energy density, in the form of kinetic energy. The antinodes are moving fastest, and in contrast there is never any motion at the nodes. There is no potential energy anywhere in the string.

We see that the energy continually flows back and forth between the nodes and antinodes. Energy never flows across a node (because a node never moves and therefore can do no work), nor across an antinode (because an antinode never applies a transverse force and therefore can do no work). The energy in each "half bump" (a quarter of a wavelength) of the sinusoidal curve is therefore constant. It flows back and forth between one end (a node/antinode) and the other end (an antinode/node). In other words, it flows back and forth across a point such as the dot we have chosen in the figure. This is consistent with the fact that the dot is doing work (positive or negative), except at the quarter-cycle points where $W = 0$.

4.6 Attenuation

What happens if we add some damping to a transverse wave on a string? This damping could arise, for example, by immersing the string in a fluid. As with the drag force in the spring/mass system we discussed in Chapter 1, we'll assume that this drag force is proportional to the (transverse) velocity of the string. Now, we usually idealize a string as having negligible thickness, but such a string wouldn't experience any damping. So for the purposes of the drag force, we'll imagine that the string has some thickness that produces a drag force of $-(\beta\, dx)v_y$ on a length dx of string, where β is the drag coefficient per unit length. The longer the piece, the larger the drag force.

The transverse forces on a little piece of string are the drag force along with the force arising from the tension and the curvature of the string, which we derived in Section 4.1. So the transverse $F = ma$ equation for a little piece is obtained by taking the $F = ma$ equation in Eq. (3) and tacking on the drag force. Since $v_y = \partial\psi/\partial t$, the desired $F = ma$ (or rather, $ma = F$) equation is

$$(\mu\, dx)\frac{\partial^2\psi}{\partial t^2} = T\, dx\frac{\partial^2\psi}{\partial x^2} - (\beta\, dx)\frac{\partial\psi}{\partial t} \quad\Longrightarrow\quad \boxed{\frac{\partial^2\psi}{\partial t^2} + \Gamma\frac{\partial\psi}{\partial t} = v^2\frac{\partial^2\psi}{\partial x^2}} \tag{72}$$

where $\Gamma \equiv \beta/\mu$ and $v^2 = T/\mu$. To solve this equation, we'll use our trusty method of guessing an exponential solution. If we guess

$$\psi(x,t) = De^{i(\omega t - kx)} \tag{73}$$

and plug this into Eq. (72), we obtain, after canceling the factor of $De^{i(\omega t - kx)}$,

$$-\omega^2 + \Gamma(i\omega) = -v^2k^2. \tag{74}$$

This equation tells us how ω and k are related, but it doesn't tell us what the motion looks like. The motion can take various forms, depending on what the given boundary conditions are. To study a concrete example, let's look at the following system.

Consider a setup where the left end of the string is located at $x = 0$ (and it extends rightward to $x = \infty$), and we arrange for that end to be driven up and down sinusoidally with a constant amplitude A. In this scenario, ω must be real, because if it had a complex value of $\omega = a + bi$, then the $e^{i\omega t}$ factor in $\psi(x,t)$ would involve a factor of e^{-bt}, which decays with time. But we're assuming a steady-state solution with constant amplitude A at $x = 0$, so there can be no decay in time. Therefore, ω must be real. The i in Eq. (74) then implies that k must have an imaginary part. Define K and $-i\kappa$ be the real and imaginary parts of k, that is,

$$k = \frac{1}{v}\sqrt{\omega^2 - i\Gamma\omega} \equiv K - i\kappa. \tag{75}$$

If you want to solve for K and κ in terms of ω, Γ, and v, you can square both sides of this equation and solve a quadratic equation in K^2 or κ^2. But we won't need the actual values. You can show however, that if K and ω have the same sign (which they do, since we're looking at a wave that travels rightward from $x = 0$), then κ is positive. Plugging $k \equiv K - i\kappa$ into Eq. (73) gives

$$\psi(x,t) = De^{-\kappa x}e^{i(\omega t - Kx)}. \tag{76}$$

A similar solution exists with the opposite sign in the imaginary exponent (but the $e^{-\kappa x}$ stays the same). The sum of these two solutions gives the actual physical (real) solution,

$$\boxed{\psi(x,t) = Ae^{-\kappa x}\cos(\omega t - Kx + \phi)} \tag{77}$$

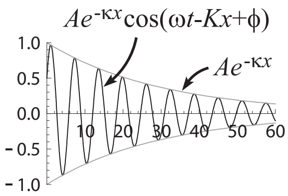


Figure 26

where the phase ϕ comes from the possible phase of D, which may be complex. (Equivalently, you can just take the real part of the solution in Eq. (76).) The coefficient A has been determined by the boundary condition that the amplitude equals A at $x = 0$. Due to the $e^{-\kappa x}$ factor, we see that $\psi(x,t)$ decays with *distance*, and not time. $\psi(x,t)$ is a rightward-traveling wave with the function $Ae^{-\kappa x}$ as its envelope. A snapshot in time is shown in Fig. 26, where we have arbitrarily chosen $A = 1$, $\kappa = 1/30$, $K = 1$, and $\phi = \pi/3$. The snapshot corresponds to $t = 0$. Such a wave is called an *attenuated* wave, because it tapers off as x grows.

Let's consider the case of small damping. If Γ is small (more precisely, if Γ/ω is small), then we can use a Taylor series to write the k in Eq. (75) as

$$k = \frac{\omega}{v}\sqrt{1 - \frac{i\Gamma}{\omega}} \approx \frac{\omega}{v}\left(1 - \frac{i\Gamma}{2\omega}\right) = \frac{\omega}{v} - \frac{i\Gamma}{2v} \equiv K - i\kappa. \tag{78}$$

Therefore, $\kappa = \Gamma/2v$. The envelope therefore takes the form, $Ae^{-\Gamma x/2v}$. So after each distance of $2v/\Gamma$, the amplitude decreases by a factor $1/e$. If Γ is very small, then this distance is very large, which makes sense. Note that this distance $2v/\Gamma$ doesn't depend on ω. No matter how fast or slow the end of the string is wiggled, the envelope dies off on the same distance scale of $2v/\Gamma$ (unless ω is slow enough so that we can't work in the approximation where Γ/ω is small). Note also that $K \approx \omega/v$ in the $\Gamma \to 0$ limit. This must be the case, of course, because we must obtain the undamped result of $k = \omega/v$ when $\Gamma = 0$. The opposite case of large damping (more precisely, large Γ/ω) is the subject of Problem [to be added].

If we instead have a setup with a uniform wave (standing or traveling) on a string, and if we then immerse the whole thing at once in a fluid, then we will have decay in *time*, instead of distance. The relation in Eq. (74) will still be true, but we will now be able to say that k must be real, because all points on the string are immersed at once, so there is no preferred value of x, and hence no decay as a function of x. The i in Eq. (74) then implies that ω must have an imaginary part, which leads to a time-decaying $e^{-\alpha t}$ exponential factor in $\psi(x,t)$.

Chapter 5

Longitudinal waves

David Morin, morin@physics.harvard.edu

In Chapter 4 we discussed transverse waves, in particular transverse waves on a string. We'll now move on to longitudinal waves. Each point in the medium (whatever it consists of) still oscillates back and forth around its equilibrium position, but now in the longitudinal instead of the transverse direction. Longitudinal waves are a bit harder to visualize than transverse waves, partly because everything is taking place along only one dimension, and partly because of the way the forces arise, as we'll see. Most of this chapter will be spent on sound waves, which are the prime example of longitudinal waves.

The outline of this chapter is as follows. As a warm up, in Section 5.1 we take another look at the longitudinal spring/mass system we originally studied in Section 2.4, where we considered at the continuum limit (the $N \to \infty$ limit). In Section 5.2 we study actual sound waves. We derive the wave equation (which takes the same form as all the other wave equations we've seen so far), and then look at the properties of the waves. In Section 5.3 we apply our knowledge of sound waves to musical instruments.

5.1 Springs and masses revisited

Recall that the wave equation for the continuous spring/mass system was given in Eq. (2.80) as

$$\frac{\partial^2 \psi(x,t)}{\partial t^2} = \frac{E}{\mu}\frac{\partial^2 \psi(x,t)}{\partial x^2}\,, \tag{1}$$

where ψ is the longitudinal position relative to equilibrium, μ is the mass density, and E is the elastic modulus. This wave equation is very similar to the one for transverse waves on a string, which was given in Eq. (4.4) as

$$\frac{\partial^2 \psi(x,t)}{\partial t^2} = \frac{T}{\mu}\frac{\partial^2 \psi(x,t)}{\partial x^2}\,, \tag{2}$$

where ψ is the transverse position relative to equilibrium, μ is the mass density, and T is the tension.

These equations take exactly the same form, so all of the same results hold. However, the fact that ψ is a longitudinal position in the former case, whereas it is a transverse position in the latter, makes the former case a little harder to visualize. For example, if we plot ψ for a sinusoidal traveling wave (either transverse or longitudinal), we have the picture shown in Fig. 1. The interpretation of this picture depends on what kind of wave we're talking about.

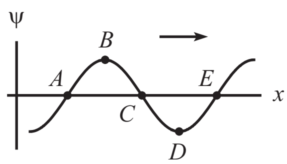


Figure 1

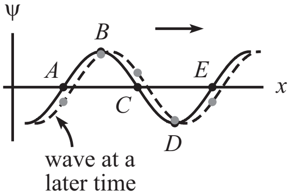


Figure 2

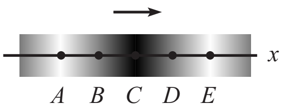


Figure 3

For a *transverse* wave, ψ is the transverse displacement, so Fig. 1 is what the string actually *looks* like from the side. The wave is therefore very easy to visualize – you just need to look at the figure. It's also fairly easy to see what the various points in Fig. 1 are doing as the wave travels to the right. (Imagine that these dots are painted on the string.) Points B and D are instantaneously at rest, points A and E are moving downward, and point C is moving upward. To verify these facts, just draw the wave at a slightly later time. The result is shown in Fig. 2, with the new positions of the dots being represented by gray dots. Remember that the points keep their same longitudinal position and simply move up or down (or not at all). They don't travel longitudinally along with the wave.

However, for a *longitudinal* wave, ψ is the longitudinal displacement, so although Fig. 1 is a perfectly valid plot of ψ, it does *not* indicate what the wave actually looks like. There is no transverse motion, so the system simply lies along a straight line. What changes is the *density* along the line. You could therefore draw the wave by shading it as in Fig. 3, but this is a bit harder to draw than Fig. 1. For a longitudinal wave, the statements in the preceding paragraph about the motion of the various points in Fig. 1 are still true, provided that "downward" is replaced with "leftward," and "upward" is replaced with "rightward. But what do things actually *look* like along the 1-D line? In particular, how does Fig. 3 follow from Fig. 1?

At points B and D in Fig. 3, the density of the masses equals the equilibrium density, because nearby points all have essentially the same displacement (see Fig. 1). But at points A and E, the density is a minimum, because points to the left of them have a negative displacement, while points to the right have a positive displacement (again see Fig. 1). The opposite is true for point C, so the density is maximum there. Various properties of the wave are indicated in Fig. 4. You should stare at this figure for a while and verify all of the stated properties. We'll talk more about the relation among the various quantities when we discuss Fig. 8 later on when we get to sound waves.

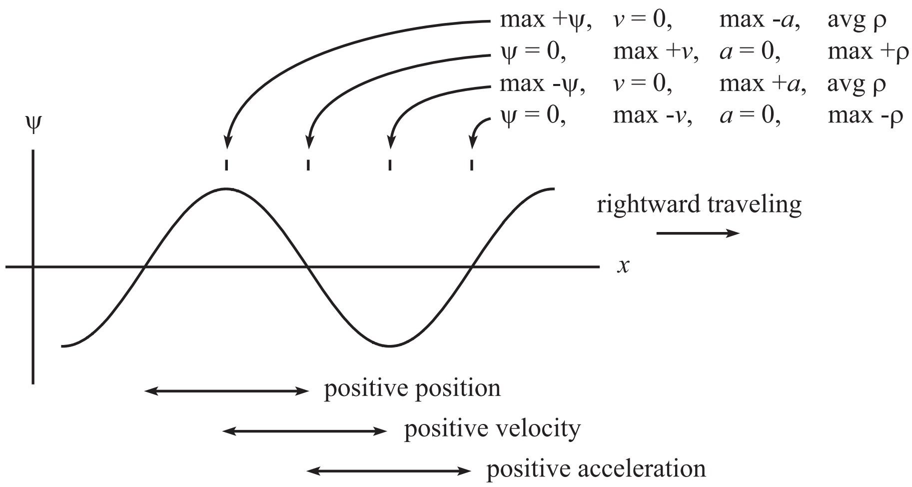


Figure 4

In the Fig. 4, the relation between ψ, v, and a is the same as always, namely, a is 90° ahead of v, and v is 90° ahead of ψ. But you should think about how these relate to the density ρ. For example, from the preceding paragraph, the (excess) ρ is proportional to the negative of the slope (see Problem [to be added] for a rigorous derivation of this fact). But we already know that v is proportional to the negative of the slope; see Eq. (4.44). Therefore, the (excess) ρ is proportional to v. For a *leftward* traveling wave, the same statement about

ρ is still true, but now v is proportional to the slope (with no negative sign). So the (excess) ρ is proportional to $-v$.

We can double check that this result makes sense with the following reasoning. Since the (excess) ρ is proportional to v, we can take the derivative of this statement to say that $\partial\rho/\partial x \propto \partial v/\partial x$ (the word "excess" is now not needed). But since a traveling wave takes the form of $\psi(x,t) = f(x - ct)$, the velocity $v = \partial\psi/\partial t$ also takes the functional form of $g(x - ct)$. Therefore, we have $\partial v/\partial x = -(1/c)\partial v/\partial t$. The righthand side of this is just the acceleration a, so the $\partial\rho/\partial x \propto \partial v/\partial x$ statement becomes

$$\frac{\partial\rho}{\partial x} \propto -a \quad \Longrightarrow \quad a \propto -\frac{\partial\rho}{\partial x}. \tag{3}$$

Does this make sense? It says, for example, that if the density is an increasing function of x at a given point, then the acceleration is negative there. This is indeed correct, because a larger density means that the springs are more compressed (or less stretched), which in turn means that they exert a larger repulsive force (or a smaller attractive force). So if the density is an increasing function of x (that is, if $\partial\rho/\partial x > 0$), then the springs to the right of a given region are pushing leftward more than the springs to the left of the region are pushing rightward. There is therefore a net negative force, which means that the acceleration a is negative, in agreement with Eq. (3).

5.2 Sound waves

5.2.1 Notation

Sound is a longitudinal wave, in both position and pressure/density, as we'll see. Sound can exist in solids, liquids, and gasses, but in this chapter we'll generally work with sound waves in air. In air, molecules push and (effectively, relative to equilibrium) pull on each other, so we have a sort of spring/mass system like the one we discussed above.

The main goal in this section is to derive the wave equation for sound waves in air. We'll find that we obtain exactly the same type of wave equation that we had in Eq. (1) for the spring/mass system. The elastic modulus E appears there, so part of our task below will be to find the analogous quantity for sound waves. We'll consider only one-dimensional waves here. That is, the waves depend only on x. Waves like this that are uniform in the transverse y and z directions are called "plane waves."

To emphasize the 1-D nature of the wave, let's consider a tube of air inside a cylindrical container, with cross-sectional area A. Fig. 5 shows a given section of air at equilibrium, and then also at a later time. Let the ends of this section be located at x and $x + \Delta x$ at equilibrium, and then at $x + \psi(x)$ and $x + \Delta x + \psi(x + \Delta x)$ at a later time, as shown. So the function ψ measures the displacement from equilibrium.

If we define $\Delta\psi$ by $\psi(x + \Delta x) \equiv \psi(x) + \Delta\psi$, then $\Delta\psi$ is how much more the right boundary of the region moves compared with the left boundary. The molecules of air are in thermal motion, of course, so it's not as if the molecules that form the boundary at position x in the first picture correspond to the molecules that form the boundary at position $x+\psi(x)$ in the second picture. But we'll ignore this fact and just pretend that it's the same molecules, for ease of discussion. It doesn't actually matter. Note that $\Delta\psi \approx (\partial\psi/\partial x)\Delta x$ for small Δx, by definition of the derivative. In actual sound waves in air, $\Delta\psi$ is much less than Δx. In other words, $\partial\psi/\partial x$ is very small.

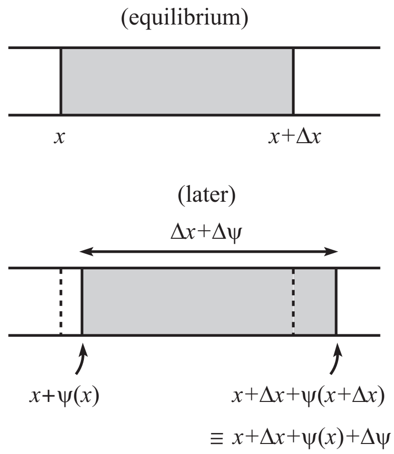


Figure 5

A note on terminology: We're taking x to be the position of a given molecule at equilibrium. So even after the molecule has moved to the position $x + \psi(x)$, it is still associated with the same value of x. So x is analogous to the index n we used in Section 2.3 and the beginning of Section 2.4. The movement of the particle didn't affect its label n there, and it doesn't affect its label x here.

In obtaining the wave equation, we'll need to get a handle on the pressure at the two ends of the given section of air, and then we'll figure out how these pressures cause the section to move. Let the pressure in the tube at equilibrium be p_0. At sea level, the atmospheric pressure happens to be about 14.7 lbs. per square inch. The first picture in Fig. 6 shows the pressures at the two ends at equilibrium; it is simply p_0 at both ends (and everywhere else).

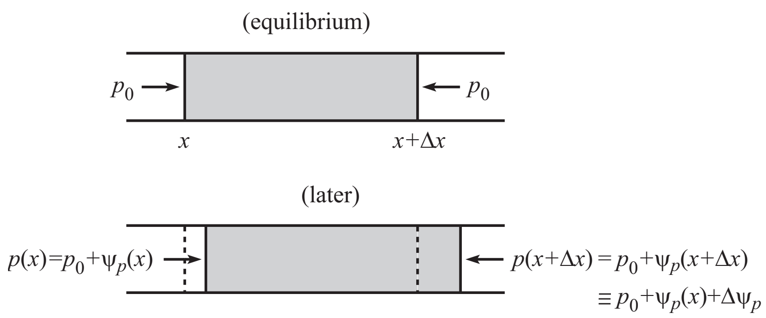


Figure 6

What about at a later time? Let $\psi_p(x)$ be the *excess* pressure (above p_0) as a function of x. (Remember that x labels the equilibrium position of the molecules, not the present position.) The total pressure at the left boundary of the section is then $p_0 + \psi_p(x)$. However, this total pressure won't be too important; the change, $\psi_p(x)$, is what we'll be concerned with. At the right boundary, the total pressure is, by definition, $p_0 + \psi_p(x + \Delta x)$. If we define $\Delta\psi_p$ by $\psi_p(x + \Delta x) \equiv \psi_p(x) + \Delta\psi_p$, then $\Delta\psi_p$ is how much the pressure at the right boundary exceeds the pressure at the left boundary. Note that $\Delta\psi_p \approx (\partial\psi_p/\partial x)\Delta x$ for small Δx. In practice, ψ_p is much smaller than p_0. And $\Delta\psi_p$ is infinitesimally small, assuming that we have picked Δx to be infinitesimally small. The pressures at a later time are summarized in the second picture in Fig. 6.

5.2.2 The wave equation

Having introduced the necessary notation, we can now derive the wave equation for sound waves. The derivation consists of four main steps, so let's go through them systematically. Our strategy will be to find the net force on a given volume of air, and then write down the $F = ma$ equation for that volume.

1. HOW THE VOLUME CHANGES: First, we need to determine how the volume of a gas changes when the pressure is changed. Qualitatively, if we increase the pressure on a given volume, the volume decreases. By how much? The decrease should certainly be proportional to the volume, because if we put two copies of a given volume next to each other, we will obtain twice the decrease. And the decrease should also be proportional to the pressure increase, provided that the increase is small. This is a reasonable claim, but not terribly obvious. We'll derive it in step 2 below. Assuming that it is true, we can write (recalling that ψ_p is defined to be the increase in pressure relative to equilibrium) the change in volume from equilibrium as $\Delta V \propto -V\psi_p$. This correctly incorporates the two proportionality facts above. The minus sign is due to the fact that an increase in pressure causes a decrease in volume. This equation is valid as long as ΔV is small compared with V, as we'll see below.

   Let's define κ to be the constant of proportionality in $\Delta V \propto -V\psi_p$. κ is known as the *compressibility*. The larger κ is, the more the volume is compressed (or expanded), for a given increase (or decrease) in pressure, ψ_p. In terms of κ, we have

   $$\Delta V = -\kappa V \psi_p. \tag{4}$$

   But from the first picture in Fig. 5, we see that the volume of gas is $V = A\Delta x$, where A is the cross-sectional area. And the change in the volume between the two pictures shown is $\Delta V = A\Delta\psi$.[1] Eq. (4) therefore becomes

   $$\frac{\Delta V}{V} = -\kappa\psi_p \implies \frac{A\Delta\psi}{A\Delta x} = -\kappa\psi_p \implies \boxed{\frac{\partial\psi}{\partial x} = -\kappa\psi_p} \tag{5}$$

   where we have taken the infinitesimal limit and changed the Δ's to differentials (partial ones, since ψ is a function of t also). The quantity $\partial\psi/\partial x$ indicates how the displacement from equilibrium grows as a function of x. Equivalently, $\partial\psi/\partial x$ is the "stretching fraction." If the displacement ψ grows by, say, 1 mm over the course of a distance of 10 cm, then the length (and hence the volume) of the region has increased by 1/100, and this equals $\partial\psi/\partial x$. Eq. (5) says that the stretching fraction is proportional to the change in pressure, which is quite reasonable.

2. CALCULATING THE COMPRESSIBILITY, κ: We'll now be rigorous about the above statement that the decrease in volume should be proportional to the pressure increase, provided that the increase is small. In the course of doing this, we'll find the value of κ.

   Let's first give a derivation that isn't quite correct. From the ideal gas law (which we'll accept here; we have to start somewhere), we have $pV = nRT$. If the temperature T is constant, then any changes in p and V must satisfy $(p + dp)(V + dV) = nRT$. Subtracting the original $pV = nRT$ equation from this one, and ignoring the second-order $dp\,dV$ term (this is where the assumption of small changes comes in), we obtain

[1] We can alternatively say that the volume is $V = A(\Delta x + \Delta\psi)$, by looking at the second picture. But we are assuming $\Delta\psi \ll \Delta x$, so to leading order we can ignore the $A\Delta\psi$ term in the volume. However, we can't ignore it in the *change* in volume, because it's the entire change.

$p\,dV + dp\,V = 0 \Longrightarrow dV = -(1/p)V\,dp$. But dp is what we've been calling ψ_p above, so if we compare this result with Eq. (4), we see that the compressibility κ equals $1/p$. Note that what we did here was basically take the differential of the $pV = C$ equation (where C is a constant) to obtain $p\,dV + dp\,V = 0$.

However, this $\kappa = 1/p$ result isn't correct, because it is based on the assumption that T is constant in the relation $pV = nRT$. But T isn't constant in a sound wave (in either space or time). The compressions are actually adiabatic, meaning that heat can't flow quickly enough to redistribute itself and even things out. Basically, the time scale of the heat flow is large compared with the time scale of the wave oscillations. So a region that heats up due to high pressure stays hot, until the pressure decreases. The correct relation (which we'll just accept here) for adiabatic processes turns out to be $pV^\gamma = C$, where γ happens to be about 7/5 for air (γ is 7/5 for a diatomic gas, and air is 99% N_2 and O_2). Taking the differential of our new $pV^\gamma = C$ equation gives

$$p \cdot \gamma V^{\gamma-1}\,dV + dp\,V^\gamma = 0 \quad \Longrightarrow \quad dV = -\left(\frac{1}{\gamma p}\right) V\,dp. \tag{6}$$

Therefore, the correct value of the compressibility κ is

$$\kappa = \frac{1}{\gamma p_0} \approx \frac{5}{7p_0}\,. \tag{7}$$

The above incorrect result wasn't so bad; it was off by only a factor of 7/5.

3. Calculating the difference in pressure, $\Delta\psi_p$: Eq. (5) involves the quantity ψ_p. However, what we're actually concerned with is not ψ_p but the *change* in ψ_p from one end of the volume to the other, because in writing down the $F = ma$ equation for the volume, we're concerned with the *net* force on it, and this involves the difference in the pressures at the ends. So let's see how ψ_p changes with x. Differentiating Eq. (5) with respect to x gives

$$\frac{\partial \psi_p}{\partial x} = -\frac{1}{\kappa}\frac{\partial^2 \psi}{\partial x^2} \quad \Longrightarrow \quad \Delta\psi_p = -\frac{1}{\kappa}\frac{\partial^2 \psi}{\partial x^2}\Delta x, \tag{8}$$

where we have multiplied both sides by ∂x and switched back to the Δ notation. Or equivalently, we have multiplied both sides by Δx, and then used the relation $\Delta\psi_p = (\partial\psi_p/\partial x)\Delta x$. Note that the second derivative $\partial^2\psi/\partial x^2$ appears here. From Eq. (5), a *constant* value of $\partial\psi/\partial x$ corresponds to a constant excess pressure $\psi_p(x)$; the tube of air just stretches uniformly if the pressure is the same everywhere. So to obtain a varying value of ψ_p, we need a varying value of $\partial\psi/\partial x$. That is, we need a nonzero value of $\partial^2\psi/\partial x^2$.

4. The $F = ma$ equation: We can now write down the $F = ma$ equation for a given volume of air. The net rightward force on the tube of air in the second picture in Fig. 6 is the cross-sectional area times the difference in the pressure at the ends. So we have, recalling the definition of $\Delta\psi_p$,

$$F_{\text{net}} = A\big(p(x) - p(x+\Delta x)\big) = A(-\Delta\psi_p). \tag{9}$$

Using the expression for $\Delta\psi_p$ we found in Eq. (8), the $F_{\text{net}} = ma$ equation on a given tube of air is (with ρ being the mass density)

$$\begin{aligned} -A\Delta\psi_p &= (\rho V)\frac{\partial^2\psi}{\partial t^2} \\ \Longrightarrow \quad -A\left(-\frac{1}{\kappa}\frac{\partial^2\psi}{\partial x^2}\Delta x\right) &= \rho(A\Delta x)\frac{\partial^2\psi}{\partial t^2}\,. \end{aligned} \tag{10}$$

Canceling the common factor of $A\Delta x$ and using $\kappa = 1/\gamma p_0$ yields

$$\boxed{\frac{\partial^2\psi}{\partial t^2} = \frac{\gamma p_0}{\rho}\cdot\frac{\partial^2\psi}{\partial x^2}} \qquad (\textit{wave equation}) \tag{11}$$

This is the desired wave equation for sound waves in air. We see that $\gamma p_0/\rho$ is the coefficient that replaces the E/μ coefficient for the longitudinal spring/mass system, or the T/μ coefficient for the transverse string.

The solutions to the wave equation are the usual exponentials,

$$\psi(x,t) = Ae^{i(\pm kx \pm \omega t)}, \tag{12}$$

where ω and k satisfy $\omega/k = \sqrt{\gamma p_0/\rho}$. And since ω/k is the speed c of the wave, we have $c = \sqrt{\gamma p_0/\rho}$. As always, the speed is simply the square root of the factor on the righthand side. As a double check on the units, we have

$$\frac{\gamma p_0}{\rho} = \frac{(\text{unitless})\cdot(\text{force/area})}{\text{mass/volume}} = \frac{(\text{kg m/s}^2)/\text{m}^2}{\text{kg/m}^3} = \frac{\text{m}^2}{\text{s}^2}, \tag{13}$$

which has the correct units of velocity squared. What is the numerical value of the speed of sound waves? For air at sea level, the pressure is $p_0 \approx 10^5$ kg/ms^2 and the density is $\rho \approx 1.3$ kg/m^3. So we have

$$c = \sqrt{\frac{\gamma p_0}{\rho}} \approx \sqrt{\frac{(7/5)(10^5\,\text{kg/ms}^2)}{1.3\,\text{kg/m}^3}} \approx 330\,\text{m/s}. \tag{14}$$

This decreases with ρ, which makes sense because the larger the density, the more inertia the air has, so the harder it is to accelerate it. The speed increases with p_0. This follows from the fact that if p_0 is large, then the compressibility κ is small (meaning the gas is not easily compressed). So for a given value of $\partial^2\psi/\partial x^2$, the force on the left side of Eq. (10) is large, which implies large accelerations.

The two partial derivatives in Eq. (11) come about in the usual way. The second time derivative comes from the "a" in $F = ma$, and the second space derivative comes from the fact that it is the difference in the first derivatives that gives the net force. Eq. (5) tells us that the first space derivative of the displacement gives a measure of the force at a given location (just as with the spring/mass system, the first space derivative told us how much the springs were stretched, which in turn gave the force). The difference in the force at the two ends is therefore proportional to the second space derivative (again as it was with the spring/mass system).

As with the other wave equations we have encountered thus far in this book, the speed of sound waves is independent of ω and k. (This won't be the case for the dispersion-ful waves we discuss in the following chapter.) So all frequencies travel at the same speed. This is fortunate, because if it weren't true, then a music concert would sound like a complete mess!

5.2.3 Pressure waves

Eq. (11) gives the wave equation for the displacement, ψ, from equilibrium for a molecule whose equilibrium position is x. However, it is rather difficult to follow the motion of a single molecule, so it would be nice to obtain the wave equation for the excess pressure ψ_p, because the pressure is much easier to measure. (It is a macroscopic property of the average of a large number of molecules, as opposed to the microscopic position of a particular molecule.) In view of the relation between ψ_p and ψ in Eq. (5), we can generate the wave equation for ψ_p by taking the $\partial/\partial x$ derivative of the wave equation in Eq. (11). Using the fact that partial derivatives commute, we obtain

$$\frac{\partial}{\partial x}\left(\frac{\partial^2\psi}{\partial t^2}\right) = \frac{\gamma p_0}{\rho}\cdot\frac{\partial}{\partial x}\left(\frac{\partial^2\psi}{\partial x^2}\right) \implies \frac{\partial^2}{\partial t^2}\left(\frac{\partial\psi}{\partial x}\right) = \frac{\gamma p_0}{\rho}\cdot\frac{\partial^2}{\partial x^2}\left(\frac{\partial\psi}{\partial x}\right). \tag{15}$$

But from Eq. (5) we know that $\partial\psi/\partial x \propto \psi_p$, so we obtain

$$\boxed{\frac{\partial^2\psi_p}{\partial t^2} = \frac{\gamma p_0}{\rho}\cdot\frac{\partial^2\psi_p}{\partial x^2}} \qquad (\textit{wave equation for pressure}) \tag{16}$$

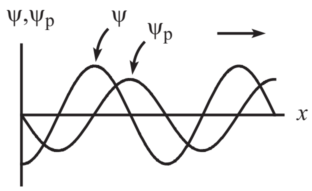


Figure 7

This is the *same* wave equation as the one for the displacement ψ in Eq. (11). So everything that is true about ψ is also true about ψ_p. The only difference is that since $\psi_p \propto -\partial\psi/\partial x$ (the minus sign is important here), the phase of ψ_p is 90° *behind* the phase of ψ. This is shown in Fig. 7. The pressure (and hence also the density) reaches its maximum value a quarter cycle after the displacement does. This is consistent with the values of ψ and ρ given in Fig. 4.

5.2.4 Impedance

What is the impedance of air? In other words, what is the force per velocity that a given region applies to an adjacent region, as a wave propagates? Remember that impedance is a property of the medium and not the wave, even though it is generally easiest to calculate it by considering the properties of a traveling wave. (However, when we discuss dispersion-ful systems in the next chapter, we will find that the impedance depends on the frequency of the wave.)

The velocity of a "sheet" of molecules whose equilibrium position is x is simply $v(x) = \partial\psi(x)/\partial t$. To find the force, consider a cross-sectional area A. We can use Eq. (5) to write the (excess) force that the sheet exerts on the region to its right as

$$F = A\psi_p = A\left(-\frac{1}{\kappa}\frac{\partial\psi}{\partial x}\right). \tag{17}$$

And since we are working with a traveling wave (no need for it to be sinusoidal), we have the usual relationship between $\partial\psi/\partial x$ and $\partial\psi/\partial t$, namely $\partial\psi/\partial x = \mp(1/c)(\partial\psi/\partial t)$ (the minus sign is associated with a rightward traveling wave). So Eq. (17) becomes

$$F = A\left(-\frac{1}{\kappa}\right)\left(\mp\frac{1}{c}\frac{\partial\psi}{\partial t}\right) = \pm\frac{A}{\kappa c}\cdot\frac{\partial\psi}{\partial t}. \tag{18}$$

The force that the sheet *feels* from the region on its right is the negative of this, but the sign isn't important when calculating the impedance Z, because Z is defined to be the magnitude of F/v. Using $v = \partial\psi/\partial t$, Eq. (18) gives the impedance as

$$Z \equiv \frac{F}{v} = \frac{A}{\kappa c}. \tag{19}$$

The impedance per unit area is the more natural thing to talk about, because Z/A is independent of the specific cross section chosen. The force F can be made arbitrarily large by making the area A arbitrarily large, so $Z = F/v$ isn't too meaningful. When people talk about the impedance of air, they usually mean "impedance per area," that is, force per velocity per area. However, we'll stick with the $Z = F/v$ definition of impedance, in which case Eq. (19) tells us that the impedance per unit area is $Z/A = 1/\kappa c$. But $c = 1/\sqrt{\kappa\rho} \Longrightarrow \kappa = 1/\rho c^2$ (this follows from writing the coefficient in Eq. (11) in terms of κ). So we have $Z/A = \rho c$. Using $c = \sqrt{\gamma p_0/\rho}$ from Eq. (14), we can write this alternatively as

$$\boxed{\frac{Z}{A} = \rho c = \sqrt{\gamma\rho p_0}} \tag{20}$$

5.2.5 Energy, Power

Energy

What is the energy density of a (longitudinal) sound wave? The kinetic energy density (per unit volume) is simply $(1/2)\rho(\partial\psi/\partial t)^2$, because the speed of the molecules is $\partial\psi/\partial t$. If we want to consider instead the kinetic energy density *per unit length* along a tube with cross-sectional area A, then this is $(1/2)(A\rho)(\partial\psi/\partial t)^2$, where $A\rho \equiv \mu$ is the mass density per unit length. We are assuming here that ρ and μ are independent of position. This is essentially true for actual sound waves, because $\partial\psi/\partial x$ is small, so the fractional change in the density from its equilibrium value is small.

What about the potential energy density? The task of Problem [to be added] is to show that the potential energy density (or rather, the excess over the equilibrium value) per unit length is $(1/2)A\gamma p_0(\partial\psi/\partial x)^2$. The total energy density per unit length (kinetic plus potential) is therefore

$$\begin{aligned}
\mathcal{E}(x,t) &= \frac{1}{2}A\rho\left(\frac{\partial\psi}{\partial t}\right)^2 + \frac{1}{2}A\gamma p_0\left(\frac{\partial\psi}{\partial x}\right)^2 \\
&= \frac{1}{2}A\rho\left[\left(\frac{\partial\psi}{\partial t}\right)^2 + \frac{\gamma p_0}{\rho}\left(\frac{\partial\psi}{\partial x}\right)^2\right] \\
&= \frac{1}{2}A\rho\left[\left(\frac{\partial\psi}{\partial t}\right)^2 + c^2\left(\frac{\partial\psi}{\partial x}\right)^2\right].
\end{aligned} \tag{21}$$

This is the same as the result we found in Eq. (4.49) for transverse waves, since $A\rho$ is the mass per unit length, μ. As with Eq. (4.49), the present expression for $\mathcal{E}(x,t)$ is valid for an arbitrary wave. But if we consider the special case of a single traveling wave, then we have the usual relation, $\partial\psi/\partial t = \pm c\,\partial\psi/\partial x$. So the two terms in the expression for $\mathcal{E}(x,t)$ are equal at a given point and at a given time. We can therefore write the energy density per unit length as

$$\boxed{\mathcal{E}(x,t) = A\rho\left(\frac{\partial\psi}{\partial t}\right)^2} \qquad \text{(for traveling waves)} \tag{22}$$

The energy density per unit volume is then $\mathcal{E}/A = \rho(\partial\psi/\partial t)^2$.

Power

Consider a cross-sectional "sheet" of molecules. At what rate does the air on the left of the sheet do work on the sheet? (This is the same type of question that we asked in Section 4.4 for a transverse wave: At what rate does the string to the left of a dot do work on the dot?) In a small amount of the time, the work done by the air is $dW = F\,d\psi = (pA)\,d\psi$.

The power on the sheet whose equilibrium position is x is therefore

$$P = \frac{\partial W}{\partial t} = (p_0 + \psi_p)A\frac{\partial \psi}{\partial t}\,. \tag{23}$$

The $Ap_0(\partial\psi/\partial t)$ part of this averages out to zero over time, so we'll ignore it. The $A\psi_p(\partial\psi/\partial t)$ term, however, always has the same sign, for the following reason. From Eq. (5), we have $\psi_p = -(1/\kappa)(\partial\psi/\partial x)$. But as usual, $\partial\psi/\partial x = \mp(1/c)(\partial\psi/\partial t)$ (the minus sign is associated with a rightward traveling wave). So the power is

$$P = A\psi_p\frac{\partial \psi}{\partial t} = A\left(\pm\frac{1}{\kappa c}\cdot\frac{\partial\psi}{\partial t}\right)\frac{\partial\psi}{\partial t} \;= \pm\frac{A}{\kappa c}\left(\frac{\partial\psi}{\partial t}\right)^2. \tag{24}$$

Using $c = 1/\sqrt{\kappa\rho} \Longrightarrow \kappa = 1/\rho c^2$, we have

$$\boxed{P = \pm A\rho c\left(\frac{\partial\psi}{\partial t}\right)^2} \tag{25}$$

Since $Z = A\rho c$ from Eq. (20), we can also write the power as

$$P = \pm Z\left(\frac{\partial\psi}{\partial t}\right)^2, \tag{26}$$

which takes exactly the same form as the result in Eq. (4.53) for transverse waves.

If we compare Eqs. (22) and (25), we see that $P = \pm c\mathcal{E}$. This makes sense, because as with transverse waves, the power must equal the product of the wave velocity and the energy density, because the $\mathcal{E}$ curve moves right along with the wave.

If we want to write P in terms of the pressure ψ_p, we can do this in the following way. Using Eq. (5), Eq. (25) becomes

$$P = \pm A\rho c\left(\mp c\frac{\partial\psi}{\partial x}\right)^2 = \pm A\rho c^3(-\kappa\psi_p)^2 = \pm A\rho c^3\left(\frac{1}{\rho c^2}\right)^2\psi_p^2 = \boxed{\pm\frac{A}{\rho c}\,\psi_p^2} \tag{27}$$

We see that when written in terms of ψ_p, the power decreases with ρ and c. But when written in terms of ψ (or rather $\partial\psi/\partial t$) in Eq. (25), it grows with ρ and c. The latter is fairly clear. For example, a larger ρ means that more matter is moving, so the energy density is larger. But the dependence on ψ_p isn't as obvious. It arises from the fact that there are factors of ρ and c hidden in ψ_p. So, for example, if ρ is increased (for a given function ψ), then ψ_p^2 grows faster than ρ, so the righthand side of Eq. (27) still increases with ρ. However, if ρ is increased for a given function ψ_p, then the power decreases, because the displacement ψ has to decrease to keep ψ_p the same, and this effect wins out over the increase in ρ, thereby decreasing P.

5.2.6 Qualitative description

Let's take a look at what a given molecule in the air is doing at a few different times, as a rightward-traveling wave passes by. A number of snapshots with phase differences of $\pi/4$ are shown in Fig. 8. The darker regions indicate a higher pressure (and density), and the lighter regions indicate a lower pressure (and density). The vertical line, which represents the equilibrium position of the molecule, is drawn for clarity.

The three plots at the top of the figure give the values of the various parameters at $t = 0$, as functions of x. So these plots correspond to the first of the shaded snapshots. The plots for the eight other snapshots are obtained by simply shifting the three plots to the right, by the same amount as the snapshots shift. The three plots at the left of the figure give the values of the various parameters at $x = 0$, as functions of t. So these plots correspond to the molecule in question (the little circle). The leftmost plot (the one for $\psi(0, t)$) is simply a copy of the circles as they appear in the snapshots.[2]

Let's discuss what's happening in each of the snapshots. In doing this, a helpful thing to remember is that for a rightward-traveling wave, the density (and hence pressure) is always in phase with the velocity (as we discussed in Section 5.1). And both the pressure and the velocity are 90° ahead of the position, as functions of time. (For a leftward-traveling wave, the pressure is out of phase with the velocity, which is, as always, 90° ahead of the position.) Again, the displacements are exaggerated in the figures. In reality, the displacements are much smaller than the wavelength. The commentary on the snapshots is as follows.

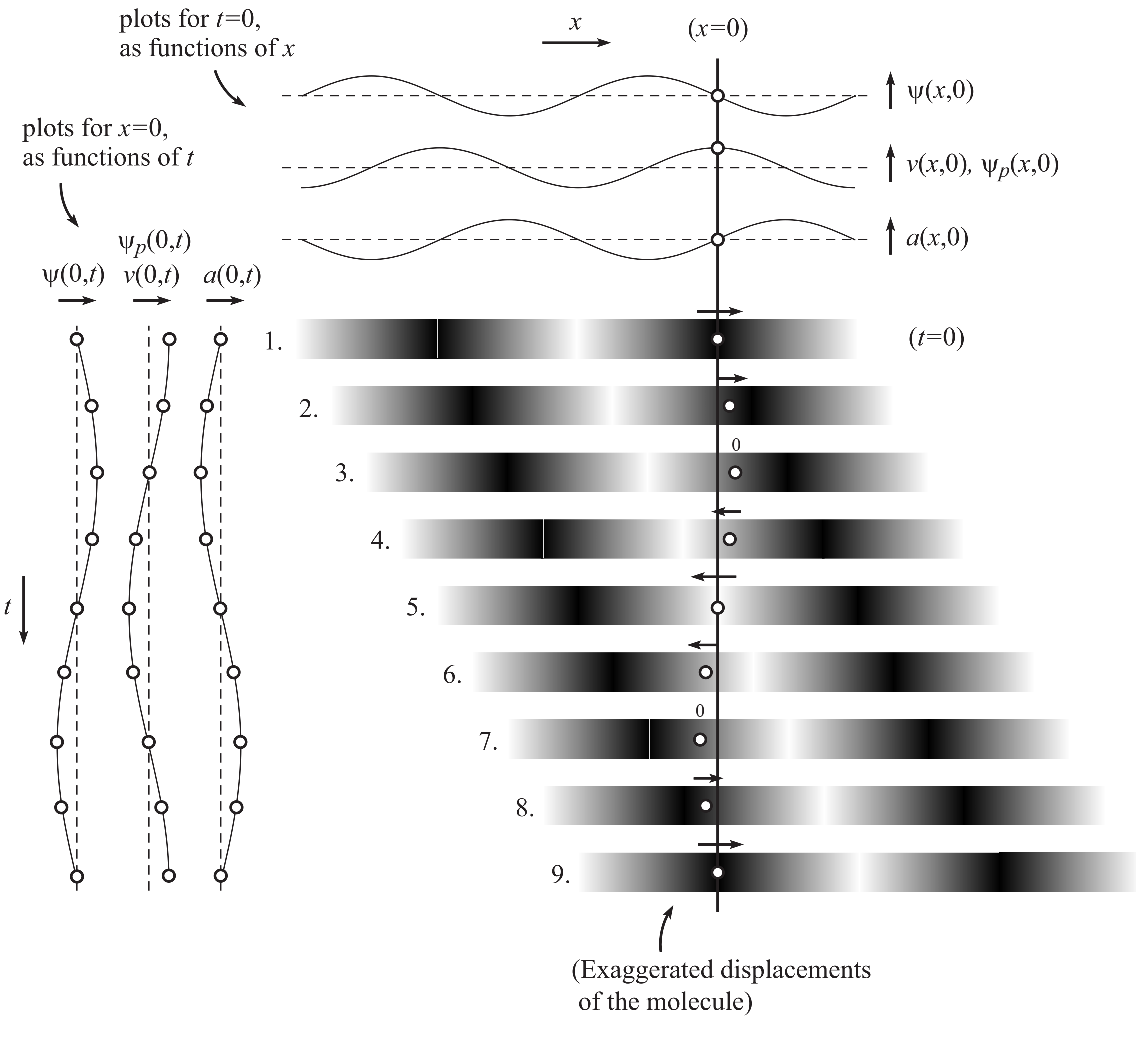


Figure 8

[2]The ψ values are exaggerated for emphasis. In reality, they are much smaller than the wavelength of the wave. But if we drew them to scale, all of the circles would essentially lie on the $x = 0$ line, and we wouldn't be able to tell that they were actually moving.

1. In the first snapshot, the molecule is located at its equilibrium position and is moving to the right with maximum speed.[3] The pressure (and density) is also maximum. The pressure is the same on both sides of the molecule, so there is zero net force, consistent with the fact that it has maximum speed and hence zero acceleration.
2. The molecule is still moving to the right, but it is decelerating ($a < 0$) because there is higher pressure (which goes hand-in-hand with higher density) on its right than on its left.
3. It has now reached its maximum value of ψ and is instantaneously at rest. It has the maximum negative acceleration, because the pressure gradient is largest here; the pressure is changing most rapidly (as a function of x) halfway between the maximum and minimum pressures. The difference between the forces on either side of the molecule is therefore largest here, so the molecule experiences the largest acceleration.
4. It has now started moving leftward and is picking up speed due to the higher pressure on the right.
5. It passes through equilibrium again, but now with the maximum negative velocity. This ends the period of negative acceleration. Up to this time, there was always higher pressure on the molecule's right side. For the next half cycle, the higher pressure will be on the left side, so there will be positive acceleration; see the "$a(0,t)$" plot in the left part of the figure.
6. It is moving to the left but is slowing down due to the higher pressure on the left.
7. It has now reached its maximum negative value of ψ and is instantaneously at rest. As in the third snapshot, the pressure gradient is largest here.
8. It has started moving rightward and is picking up speed due to the higher pressure on the left.
9. We are back to the beginning of the cycle. The molecule is in the equilibrium position and is moving to the right with maximum speed.

5.3 Musical instruments

Musical instruments (at least the wind ones) behave roughly like pipes of various sorts, so let's start our discussion of instruments by considering the simple case of a standing wave in a pipe.

Consider first the case where the pipe is closed at one end, taken to be at $x = 0$. The air molecules at the closed end can't move into the "wall" at the end. And they can't move away from it either, because there would then be a vacuum at the wall which would immediately suck the molecules back to the wall. This boundary condition at $x = 0$ tells us that the wall must be a node of the $\psi(x,t)$ wave. The standing wave for the displacement must therefore be of the form,

$$\psi(x,t) = A \sin kx \cos(\omega t + \phi). \tag{28}$$

What does the pressure wave look like? Since $\psi_p = -(1/\kappa)(\partial\psi/\partial x)$ from Eq. (5), we have

$$\psi_p(x,t) = -\frac{Ak}{\kappa} \cos kx \cos(\omega t + \phi). \tag{29}$$

[3] Or, it is moving to the left if we have a leftward-traveling wave. After reading through this commentary, you should verify that everything works out for a leftward-traveling wave.

Looking at the x dependence of this function, we see that nodes of ψ correspond to antinodes of ψ_p, and vice-versa.

If we instead have an open end at $x = 0$, then the boundary condition isn't as obvious. You might claim that now the air molecules move a maximum amount at the open end, which means that instead of a node in ψ, we have an antinode. This is indeed correct, but it isn't terribly obvious. So let's consider the pressure wave instead. If we have a standing wave inside the pipe, then there is essentially no wave outside the pipe. (Well, there must of course be *some* wave outside, given that there are sound waves hitting your ear.) So the pressure outside must be (essentially) the atmospheric pressure p_0. In other words, $\psi_p = 0$ outside the pipe. And since the pressure must be continuous, the boundary condition at the open end at $x = 0$ is $\psi_p = 0$. So the pressure has a *node* there. The pressure can therefore be written as

$$\psi_p(x,t) = B \sin kx \cos(\omega t + \phi). \tag{30}$$

The $\psi(x,t)$ function that satisfies $\psi_p = -(1/\kappa)(\partial\psi/\partial x)$ is then

$$\psi(x,t) = \frac{B\kappa}{k} \cos kx \cos(\omega t + \phi). \tag{31}$$

(A nonzero constant of integration would just give a redefinition of the equilibrium position.) So ψ does indeed have an antinode at $x = 0$, as we suspected.

The above results hold for any open or closed end, independent of where it is located. It doesn't have to be located at the arbitrarily-chosen position of $x = 0$, of course. So Fig. 9 shows the lowest-frequency (longest-wavelength) modes for the three possible cases of combinations of end types: closed/open, closed/closed, and open/open. In practice, the pressure node is slightly outside the open end, because the air just outside the pipe vibrates a little bit.

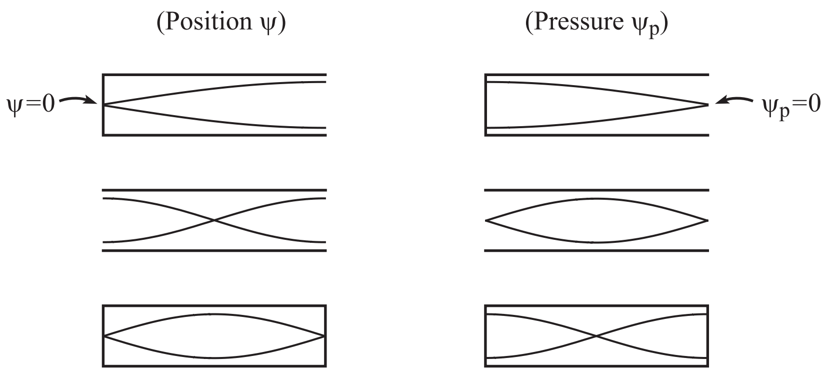


Figure 9

An instrument like a flute is essentially open at both ends (with one end being the mouthpiece). But most other instruments (reeds, brass, etc.) are open at one end and essentially closed at the mouthpiece end. This is due to the fact that the vibrating reed (or the vibrating lips in the mouthpiece) doesn't move much (so $\psi \approx 0$), but it is what is driving the pressure wave (so ψ_p is maximum there). A clarinet therefore corresponds to the first case (closed/open) in Fig. 9, while a flute corresponds to the second case (open/open). In view of this, you can see why a clarinet can play about an octave lower (which means half the frequency) than a flute, even though they have about the same length. The longest

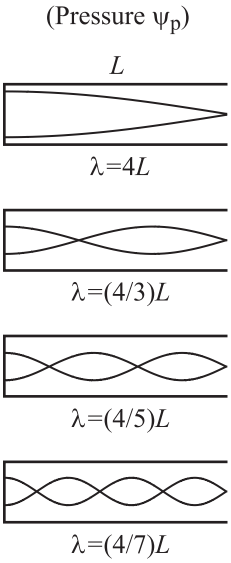


Figure 10

wavelength for a clarinet (which is four times the length of the pipe) is twice a long as the longest wavelength for a flute (which is two times the length of the pipe).[4]

The lowest four standing waves for a clarinet are shown in Fig. 10 (described by the pressure waves, which is customary). The wavelengths are $4L$, $4L/3$, $4L/5$, $4L/7$, etc. The frequencies are inversely proportional to the wavelengths, because $\nu\lambda = c \Longrightarrow \nu \propto 1/\lambda$. So the frequencies are in the ratio of $1 : 3 : 5 : 7 : \cdots$. These notes are very far apart. For an open/open pipe like a flute, the wavelengths are $2L$, $2L/2$, $2L/3$, $2L/4$, etc., which means that the frequencies are in the ratio of $1 : 2 : 3 : 4 : \cdots$. These notes are also very far apart. So if a clarinet or a flute didn't have any keys, you wouldn't be able to play anywhere near all of the notes in a standard scale.

Keys remedy this problem in the following way. If all the keys are closed, then we simply have a pipe. But if a given key is open, then this forces the pressure wave to have a node at that point, because the pressure must match up with the atmospheric pressure there. So we have essentially shortened the pipe by creating an effectively open end at the location of the open key. With many keys, this allows for many different effective pipe lengths, and hence many different notes. And also many different ways to play a given note. If we have a particular standing wave in the instrument, and if we then open a key at the location of a (pressure) node, then this doesn't change anything, so we get the same note.

What about a trumpet, which has only three valves? It's a bit complicated, but the conical shape (at least near the end) has the effect of making the frequencies be closer together (and also higher). And the mouthpiece helps too. The end result (if done properly) is that the frequencies are in the ratio $2 : 3 : 4 : 5 : 6 : \cdots$ (for some reason, the 1 is missing) instead of the $1 : 3 : 5 : 7 : \cdots$ ratios for the closed/open case in Fig. 10. This is indeed the ratio of the frequencies of the notes (C,G,C,E,G,...) that can be be played on a trumpet without pressing down any valves. The valves then change the length of the pipe in a straightforward manner.

The flared bell of a trumpet has the effect (compared with a cylinder of the same length) of raising the low notes more than the high notes, because the long wavelengths (low notes) can't follow the bell as easily, so they're reflected sooner than the short wavelengths.[5] The longer wavelengths therefore effectively see a shorter pipe. However, another effect of the bell is that because the short wavelengths follow it so easily (right out to the outside atmospheric pressure), there isn't much reflection for these waves, so it's harder to get a standing wave. The high notes are therefore less well defined, and thus blend together (which is quite evident if you've ever heard a trumpet player screeching away in the high register).

Everything you ever wanted to know about the physics of musical instruments can be found on this website: http://www.phys.unsw.edu.au/music

[4] The third option in Fig. 9, the closed/closed pipe, isn't too conducive to making music. Such an instrument couldn't have any keys, because they provide openings to the outside world. And furthermore, you couldn't blow into one like in reed or brass instrument, because there would be no place for the air to come out. And you can't blow across an opening like in a flute, because that's an open end.

[5] The fact that the shorter wavelengths (high notes) can follow the bell is the same effect as in the "Gradually changing string density" impedance-matching example we discussed in Section 4.3.2.

Chapter 6

Dispersion

David Morin, morin@physics.harvard.edu

The waves we've looked at so far in this book have been "dispersionless" waves, that is, waves whose speed is independent of ω and k. In all of the systems we've studied (longitudinal spring/mass, transverse string, longitudinal sound), we ended up with a wave equation of the form,

$$\frac{\partial^2\psi}{\partial t^2} = c^2\frac{\partial^2\psi}{\partial x^2}\,, \tag{1}$$

where c depends on various parameters in the setup. The solutions to this equation can be built up from exponential functions, $\psi(x,t) = Ae^{i(kx-\omega t)}$. Plugging this function into Eq. (1) gives

$$\omega^2 = c^2k^2. \tag{2}$$

This is the so-called *dispersion relation* for the above wave equation. But as we'll see, it is somewhat of a trivial dispersion relation, in the sense that there is no dispersion. We'll explain what we mean by this below.

The velocity of the wave is $\omega/k = \pm c$, which is independent of ω and k. More precisely, this is the *phase velocity* of the wave, to distinguish it from the *group velocity* which we'll define below. The qualifier "phase" is used here, because the speed of a sinusoidal wave $\sin(kx-\omega t)$ is found by seeing how fast a point with constant phase, $kx-\omega t$, moves. So the phase velocity is given by

$$kx - \omega t = \text{Constant} \quad\Longrightarrow\quad \frac{d(kx-\omega t)}{dt} = 0 \quad\Longrightarrow\quad k\frac{dx}{dt} - \omega = 0 \quad\Longrightarrow\quad \frac{dx}{dt} = \frac{\omega}{k}\,, \tag{3}$$

as desired.

As we've noted many times, a more general solution to the wave equation in Eq. (1) is *any* function of the form $f(x-ct)$; see Eq. (2.97). So the phase velocity could reasonably be called the "argument velocity," because c is the speed with which a point with constant argument, $x-ct$, of the function f moves.

However, not all systems have the property that the phase velocity ω/k is constant (that is, independent of ω and k). It's just that we've been lucky so far. We'll now look at a so-called *dispersive* system, in which the phase velocity isn't constant. We'll see that things get more complicated for a number of reasons. In particular, a new feature that arises is the *group velocity*.

The outline of this chapter is as follows. In Section 6.1 we discuss a classic example of a dispersive system: transverse waves in a setup consisting of a massless string with discrete point masses attached to it. We will find that ω/k is not constant. That is, the speed of a wave depends on its ω (or k) value. In Section 6.2 we discuss *evanescent waves*. Certain dispersive systems support sinusoidal waves only if the frequency is above or below a certain cutoff value. We will determine what happens when these bounds are crossed. In Section 6.3 we discuss the *group velocity*, which is the speed with which a wave packet (basically, a bump in the wave) moves. We will find that this speed is *not* the same as the phase velocity. The fact that these two velocities are different is a consequence of the fact that in a dispersive system, waves with different frequencies move with different speeds. The two velocities are the same in a non-dispersive system, which is why there was never any need to introduce the group velocity in earlier chapters.

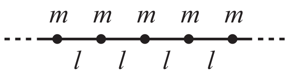


Figure 1

6.1 Beads on a string

Consider a system that is made up of beads on a massless string. The beads have mass m and are glued to the string with separation ℓ, as shown in Fig. 1. The tension is T. We'll assume for now that the system extends infinitely in both directions. The goal of this section is to determine what transverse waves on this string look like. We'll find that they behave fundamentally different from the waves on the continuous string that we discussed in Chapter 4. However, we'll see that in a certain limit they behave the same.

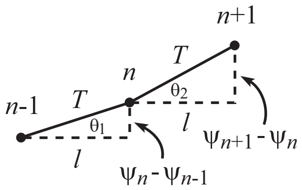


Figure 2

We'll derive the wave equation for the beaded string by writing down the transverse $F = ma$ equation on a given bead. Consider three adjacent beads, label by $n-1$, n, and $n+1$, as shown in Fig. 2. For small transverse displacements, ψ, we can assume (as we did in Chapter 4) that the beads move essentially perpendicular to the equilibrium line of the string. And as in Chapter 4, the tension is essentially constant, for small displacements. So the transverse $F = ma$ equation on the middle mass in Fig. 2 is (using $\sin\theta \approx \tan\theta$ for small angles)

$$\begin{aligned} m\ddot{\psi}_n &= -T\sin\theta_1 + T\sin\theta_2 \\ &= -T\left(\frac{\psi_n - \psi_{n-1}}{\ell}\right) + T\left(\frac{\psi_{n+1} - \psi_n}{\ell}\right) \\ \Longrightarrow \quad \ddot{\psi}_n &= \omega_0^2(\psi_{n+1} - 2\psi_n + \psi_{n-1}), \qquad \text{where } \omega_0^2 = \frac{T}{m\ell}\,. \end{aligned} \tag{4}$$

This has *exactly* the same form as the $F = ma$ equation for the longitudinal spring/mass system we discussed in Section 2.3; see Eq. (2.41). The only difference is that ω_0^2 now equals $T/m\ell$ instead of k/m. We can therefore carry over *all* of the results from Section 2.3. You should therefore reread that section before continuing onward here.

We will, however, make one change in notation. The results in Section 2.3 were written in terms of n, which labeled the different beads. But for various reasons, we'll now find it more convenient to work in terms of the position, x, along the string (as we did in Section 2.4). The solutions in Eq. (2.55) are linear combinations of functions of the form,

$$\psi_n(t) = \text{trig}(n\theta)\,\text{trig}(\omega t), \tag{5}$$

where each "trig" means either sine or cosine, and where we are now using ψ to label the displacement (which is now transverse). θ can take on a continuous set of values, because we're assuming for now that the string extends infinitely in both directions, so there's aren't any boundary conditions that restrict θ. ω can also take on a continuous set of values, but it must be related to θ by Eq. (2.56):

$$2\cos\theta \equiv \frac{2\omega_0^2 - \omega^2}{\omega_0^2} \quad \Longrightarrow \quad \omega^2 = 4\omega_0^2\left(\frac{1-\cos\theta}{2}\right) \quad \Longrightarrow \quad \omega = 2\omega_0\sin\left(\frac{\theta}{2}\right). \tag{6}$$

Let's now switch from the $n\theta$ notation in Eq. (5) to the more common kx notation. But remember that we only care about x when it is a multiple of ℓ, because these are the locations of the beads. We define k by

$$kx \equiv n\theta \quad \Longrightarrow \quad k(n\ell) = n\theta \quad \Longrightarrow \quad k\ell = \theta. \tag{7}$$

We have chosen the $x = 0$ point on the string to correspond to the $n = 0$ bead. The $\psi_n(t)$ in Eq. (5) now becomes

$$\psi(x,t) = \text{trig}(kx)\,\text{trig}(\omega t). \tag{8}$$

In the old notation, θ gave a measure of how fast the wave oscillated as a function of n, In the new notation, k gives a measure of how fast the wave oscillates as a function of x. k and θ differ simply by a factor of the bead spacing, ℓ. Plugging $\theta = k\ell$ into Eq. (6) gives the relation between ω and k:

$$\boxed{\omega(k) = 2\omega_0 \sin\left(\frac{k\ell}{2}\right)} \qquad \text{(dispersion relation)} \tag{9}$$

where $\omega_0 = \sqrt{T/m\ell}$. This is known as the *dispersion relation* for our beaded-string system. It tells us how ω and k are related. It looks quite different from the $\omega(k) = ck$ dispersion relation for a continuous string (technically $\omega(k) = \pm ck$, but we generally don't bother with the sign). However, we'll see below that they agree in a certain limit.

What is the velocity of a wave with wavenumber k? (Just the phase velocity for now. We'll introduce the group velocity in Section 6.3.) The velocity is still ω/k (the reasoning in Eq. (3) is still valid), so we have

$$c(k) = \frac{\omega}{k} = \frac{2\omega_0 \sin(k\ell/2)}{k}\,. \tag{10}$$

The main point to note here is that this velocity depends on k, unlike in the dispersionless systems in earlier chapters. In the present system, ω isn't proportional to k.

We can perform a double check on the velocity $c(k)$. In the limit of very small ℓ (technically, in the limit of very small $k\ell$), we essentially have a continuous string. So Eq. (10) had better reduce to the $c = \sqrt{T/\mu}$ result we found in Eq. (4.5) for transverse waves on a continuous string. Or said in another way, the velocity in Eq. (10) had better not depend on k in this limit. And indeed, using $\sin\epsilon \approx \epsilon$, we have

$$c(k) = \frac{2\omega_0 \sin(k\ell/2)}{k} \approx \frac{2\omega_0 (k\ell/2)}{k} = \omega_0 \ell = \sqrt{\frac{T}{m\ell}}\,\ell = \sqrt{\frac{T}{m/\ell}} \equiv \sqrt{\frac{T}{\mu}}\,, \tag{11}$$

where μ is the mass density per unit length. So it does reduce properly to the constant value of $\sqrt{T/\mu}$. Note that the condition $k\ell \ll 1$ can be written as $(2\pi/\lambda)\ell \ll 1 \Longrightarrow \ell \ll \lambda$. In other words, if the spacing between the beads is much shorter than the wavelength of the wave in question, then the string acts like a continuous string. This makes sense. And it makes sense that the condition should involve these two lengths, because they are the only two length scales in the system.

If the $\ell \ll \lambda$ condition doesn't hold, then the value of ω/k in Eq. (10) isn't independent of k, so the beaded string apparently doesn't behave like a continuous string. What does it behave like? Well, the exact expression for ω in terms of k given in Eq. (9) yields the plot shown in Fig. 3.

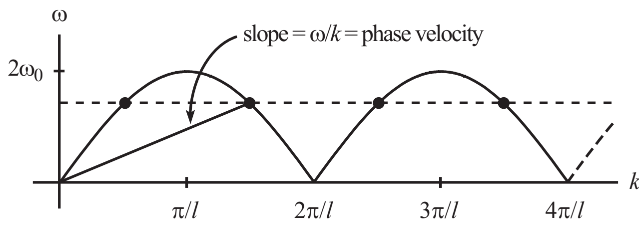


Figure 3

There are various things to note about this figure:

1. Given a value of k and its associated value of ω, the phase velocity of the wave is ω/k. But ω/k is the slope from the origin to the point (k, ω) in the figure, as shown. So the phase velocity has this very simple graphical interpretation. As we saw in Eq. (11), the slope starts off with a value of $\omega_0 \ell = \sqrt{T/\mu}$ near the origin, but then it decreases. It then repeatedly increases and decreases as the point (k, ω) runs over the successive bumps in the figure. As $k \to \infty$, the slope ω/k goes to zero (we'll talk more about this in the third comment below).

2. Note that $\omega_0 \equiv \sqrt{T/m\ell}$ can be written as $\omega_0 = (1/\ell)\sqrt{T/\mu}$, where $\mu = m/\ell$ is the mass density. So if we hold T and μ constant and decrease ℓ (for example, if we keep subdividing the masses, thereby making the string more and more continuous), then ω_0 grows like $1/\ell$. So the maximum height of the bumps in Fig. 3, which is $2\omega_0$, behaves like $1/\ell$. But the width of the bumps also behaves like $1/\ell$. So if we decrease ℓ while keeping T and μ constant, the whole figure simply expands uniformly. The linear approximation in Eq. (11) near the origin is therefore relevant for a larger range of k values. This means that the string behaves like a continuous string for more k values, which makes sense.

3. From Fig. 3, we see that many different values of k give the same ω value. In particular, k_1, $2\pi/\ell - k_1$, $2\pi/\ell + k_1$, etc., all give the same ω. However, it turns out that only the first half-bump of the curve (between $k = 0$ and $k = \pi/\ell$) is necessary for describing any possible motion of the beads. The rest of the curve gives repetitions of the first half-bump. The reason for this is that we care only about the movement of the beads, which are located at positions of the form $x = n\ell$. We don't care about the rest of the string. Consider, for example, the case where the wavenumber is $k_2 = 2\pi/\ell - k_1$. A rightward-traveling wave with this wavenumber takes the form,

$$\begin{aligned} A\cos(k_2 x - \omega t) &= A\cos\left(\left(\frac{2\pi}{\ell} - k_1\right)(n\ell) - \omega t\right) \\ &= A\cos\Big(2n\pi - k_1(n\ell) - \omega t\Big) \\ &= A\cos(-k_1 x - \omega t). \end{aligned} \tag{12}$$

We therefore conclude that a rightward-moving wave with wavenumber $2\pi/\ell - k_1$ and frequency ω gives exactly the same positions of the beads as a *leftward*-moving wave with wavenumber k_1 and frequency ω. (A similar statement holds with "right" and "left" reversed.) If we had instead picked the wavenumber to be $2\pi/\ell + k_1$, then a quick sign change in Eq. (12) shows that this would yield the same positions as a *rightward*-moving wave with wavenumber k_1 and frequency ω. The rightward/leftward correspondence alternates as we run through the class of equivalent k's.

It is worth emphasizing that although the waves have the same values at the positions of the beads, the waves look quite different at other locations on the string. Fig. 4 shows the case where $k_1 = \pi/2\ell$, and so $k_2 = 2\pi/\ell - k_1 = 3\pi/2\ell$. The two waves have common values at positions of the form $x = n\ell$ (we have arbitrarily chosen $\ell = 1$). The k values are in the ratio of 1 to 3, so the speeds ω/k are in the ratio of 3 to 1 (because the ω values are the same). The k_2 wave moves slower. From the previous paragraph, if the k_2 wave has speed v to the right, then the k_1 wave has speed $3v$ to the left. If we look at slightly later times when the waves have moved distances $3d$ to the left and d to the right, we see that they still have have common values at positions of the form $x = n\ell$. This is what Eq. (12) says in equations. The redundancy of the k values is simply the Nyquist effect we discussed at the end of Section 2.3, so you should reread that subsection now if you haven't already done so.

In comment "1" above, we mentioned that as $k \to \infty$, the phase velocity ω/k goes to zero. It is easy to see this graphically. Fig. 5 shows waves with wavenumbers $k_1 = \pi/2\ell$, and $k_2 = 6\pi/\ell - k_1 = 11\pi/2\ell$. The wave speed of the latter is small; it is only 1/11 times the speed of the former. This makes sense, because the latter wave (the very wiggly one) has to move only a small distance horizontally in order for the dots (which always have integral values of x here) to move appreciable distances vertically. A small movement in the wiggly wave will cause a dot to undergo, say, a full oscillation cycle. At the location of any of the dots, the slope of the k_2 wave is always -11 times the slope of the k_1 wave. So the k_2 wave has to move only 1/11 as far as the k_1 wave, to give the same change in height of a given dot. This slope ratio of -11 (at x values of the form $n\ell$) is evident from taking the $\partial/\partial x$ derivative of Eq. (12); the derivatives (the slopes) are in the ratio of the k's.

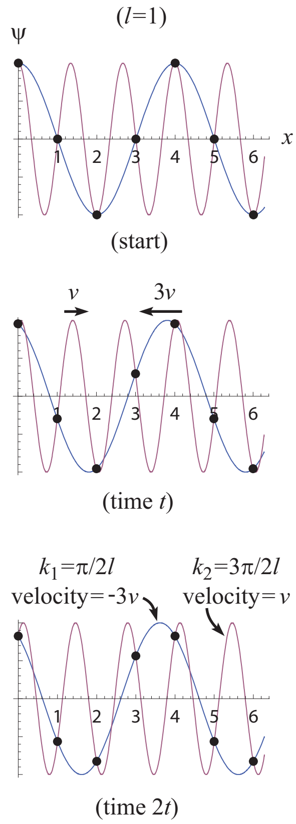


Figure 4

4. How would the waves in Fig. 4 behave if we were instead dealing with the dispersionless system of transverse waves on a continuous string, which we discussed in Chapter 4? (The length ℓ now doesn't exist, so we'll just consider two waves with wavenumbers k and $3k$, for some value of k.) In the dispersionless case, all waves move with the *same* speed. (We would have to be given more information to determine the direction, because $\omega/k = c$ only up to a $\pm$ sign.) The transverse-oscillation frequency of a given point described by the k_2 wave is therefore 3 times what the frequency would be if the point were instead described by the k_1 wave, as indicated in the straight-line dispersion relation in Fig. 6. Physically, this fact is evident if you imagine shifting both of the waves horizontally by, say, 1 cm on the paper. Since the k_2 wavelength is 1/3 the k_1 wavelength, a point on the k_2 curve goes through 3 times as much oscillation phase as a point on the k_2 wave. In contrast, in a dispersion*ful* system, the speeds of the waves don't have to all be equal. And furthermore, for the k values associated with the points on the horizontal line in Fig. 3, the speeds work out in just the right way so that the oscillation frequencies of the points don't depend on which wave they're considered to be on.

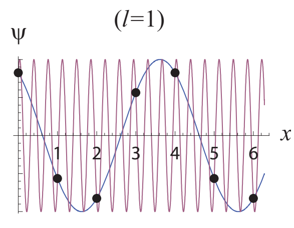


Figure 5

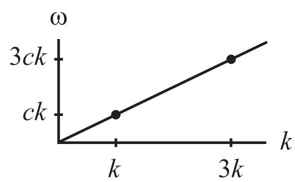


Figure 6

6.2 Evanescent waves

6.2.1 High-frequency cutoff

The dispersion relation in Eq. (9) follows from Eq. (6), which we derived back in Section 2.3.1 (although we didn't use the term "dispersion" there). The bulk of the derivation was contained in Claim 2.1. But recall that this claim assumed that $\omega \le 2\omega_0$. This is consistent with the fact that the largest value of ω in Fig. 3 is $2\omega_0$? However, what if we grab the end of a string and wiggle it sinusoidally with a frequency ω that is larger than $2\omega_0$. We're free to pick any ω we want, and the string will certainly undergo *some* kind of motion. But apparently this motion, whatever it is, isn't described by the above sinusoidal waves that we found for the $\omega \le 2\omega_0$ case.

If $\omega > 2\omega_0$, then the math in Section 2.3.1 that eventually led to the $\omega = 2\omega_0 \sin(k\ell/2)$ result in Eq. (9) is still perfectly valid. So if $\omega > 2\omega_0$, we conclude that $\sin(k\ell/2)$ must be greater than 1. This isn't possible if k is real, but it *is* possible if k is complex. So let's plug $k \equiv K + i\kappa$ into $\omega = 2\omega_0 \sin(k\ell/2)$, and see what we get. We obtain (the trig sum formula works fine for imaginary arguments)

$$\begin{aligned} \frac{\omega}{2\omega_0} &= \sin\left(\frac{K\ell}{2} - \frac{i\kappa\ell}{2}\right) \\ &= \sin\left(\frac{K\ell}{2}\right)\cos\left(\frac{i\kappa\ell}{2}\right) - \cos\left(\frac{K\ell}{2}\right)\sin\left(\frac{i\kappa\ell}{2}\right). \end{aligned} \tag{13}$$

By looking at the Taylor series for cosine and sine, the $\cos(i\kappa\ell/2)$ function is real because the series has only even exponents, while the $\sin(i\kappa\ell/2)$ function is imaginary (and nonzero) because the series has only odd exponents. But we need the righthand side of Eq. (13) to be real, because it equals the real quantity $\omega/2\omega_0$. The only way for this to be the case is for the $\cos(K\ell/2)$ coefficient of $\sin(i\kappa\ell/2)$ to be zero. Therefore, we must have $K\ell = \pi, 3\pi, 5\pi, \ldots$. However, along the same lines as the redundancies in the k values we discussed in the third comment in the previous section, the $3\pi, 5\pi, \ldots$ values for $K\ell$ simply reproduce the motions (at least at the locations of the beads) that are already described by the $Ka = \pi$ value. So we need only consider the $K\ell = \pi \Longrightarrow K = \pi/\ell$ value. Said in another way, if we're ignoring all the Nyquist redundancies, then we know that $k = \pi/\ell$ when $\omega = 2\omega_0$ (see Fig. 3). And since the real part of k should be continuous at $\omega = 2\omega_0$ (imagine increasing ω gradually across this threshold), we conclude that $K = \pi/\ell$ for all $\omega > 2\omega_0$. So $k \equiv K + i\kappa$ becomes

$$k = \frac{\pi}{\ell} + i\kappa. \tag{14}$$

Plugging $K = \pi/\ell$ into Eq. (13) yields

$$\frac{\omega}{2\omega_0} = (1)\cos\left(\frac{i\kappa\ell}{2}\right) - 0 \quad \Longrightarrow \quad \cosh\left(\frac{\kappa\ell}{2}\right) = \frac{\omega}{2\omega_0}. \tag{15}$$

This equation determines κ. You van verify the conversion to the hyperbolic cosh function by writing out the Taylor series for both $\cos(iy)$ and $\cosh(y)$. We'll keep writing things in terms of κ, but it's understood that we can solve for it by using Eq. (15).

What does the general exponential solution, $Be^{i(kx-\omega t)}$, for ψ look like when k takes on the value in Eq. (14)? (We could work in terms of trig functions, but it's *much* easier to use exponentials; we'll take the real part in the end.) Remembering that we care only about the position of the string at the locations of the masses, the exponential solution at positions of the form $x = n\ell$ becomes

$$\begin{aligned} \psi(x,t) &= Be^{i(kx-\omega t)} \\ \Longrightarrow \quad \psi(n\ell,t) &= Be^{i((\pi/\ell+i\kappa)(n\ell)-\omega t)} \\ &= Be^{-\kappa n\ell}e^{i(n\pi-\omega t)} \\ &= Be^{-\kappa n\ell}(-1)^n e^{-i\omega t} \\ &\to \boxed{Ae^{-\kappa n\ell}(-1)^n\cos(\omega t+\phi)} \end{aligned} \tag{16}$$

where we have taken the real part. The phase ϕ comes from a possible phase in B.[1] If we

[1] If you're worried about the legality of going from real k values to complex ones, and if you have your doubts that this ψ function actually does satisfy Eq. (4), you should plug it in and explicitly verify that it works, provided that κ is given by Eq. (15). This is the task of Problem [to be added].

want to write this as a function of $x = n\ell$, then it equals $\psi(x,t) = Ae^{-\kappa x}(-1)^{x/\ell}\cos(\omega t+\phi)$. But it is understood that this is valid only for x values that are integral multiples of ℓ. Adjacent beads are 180° out of phase, due to the $(-1)^n$ factor in ψ. As a function of position, the wave is an alternating function that is suppressed by an exponential. A wave that dies out exponentially like this is called an *evanescent wave*. Two snapshots of the wave at a time of maximal displacement (that is, when $\cos(\omega t+\phi) = 1$) are shown in Fig. 7, for the values $A = 1$, $\ell = 1$. The first plot has $\kappa = 0.03$, and the second has $\kappa = 0.3$. From Eq. (15), we then have $\omega \approx (2.02)(2\omega_0)$ in the latter, and ω is essentially equal to $2\omega_0$ in the former.

Since the time and position dependences in the wave appear in separate factors (and not in the form of a $kx - \omega t$ argument), the wave is a standing wave, not a traveling wave. As time goes on, each wave in Fig. 7 simply expands and contracts (and inverts), with frequency ω. All the beads in each wave pass through equilibrium at the same time.

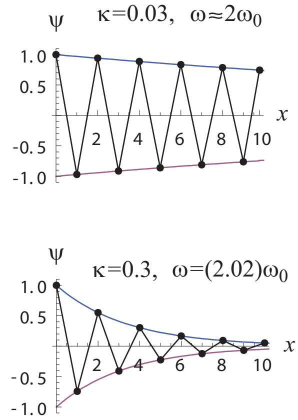


Figure 7

REMARK: A note on terminology. In Section 4.6 we discussed *attenuated* waves, which are also waves that die out. The word "attenuation" is used in the case of actual sinusoidal waves that decrease to zero due to an envelope curve (which is usually an exponential); see Fig. 4.26. The word "evanescent" is used when there is no oscillatory motion; the function simply decreases to zero. The alternating signs in Eq. (16) make things a little ambiguous. It's semantics as to whether Eq. (16) represents two separate exponential curves going to zero, or a very fast oscillation within an exponential envelope. But we'll choose to call it an evanescent wave. At any rate, the system in Section 6.2.2 below will support waves that are unambiguously evanescent. ♣

We see that $2\omega_0$ is the frequency above which the system doesn't support traveling waves. Hence the "High-frequency cutoff" name of this subsection. We can have nice traveling-wave motion below this cutoff, but not above. Fig. 8 shows the $\psi(x,t) = A\cos(kx-\omega t+\phi)$ waves that arise if we wiggle the end of the string with frequencies ω equal to $(0.1)\omega_0$, ω_0, $(1.995)\omega_0$, and $2\omega_0$. The corresponding values of k are determined from Eq. (9). We have arbitrarily picked $A = 1$ and $\ell = 1$. The sine waves are shown for convenience, but they aren't really there. We haven't shown the actual straight-line string segments between the masses. At the $\omega = 2\omega_0$ cutoff between the traveling waves in Fig. 8 and the evanescent waves in Fig. 7, the masses form two horizontal lines that simply rise and fall.

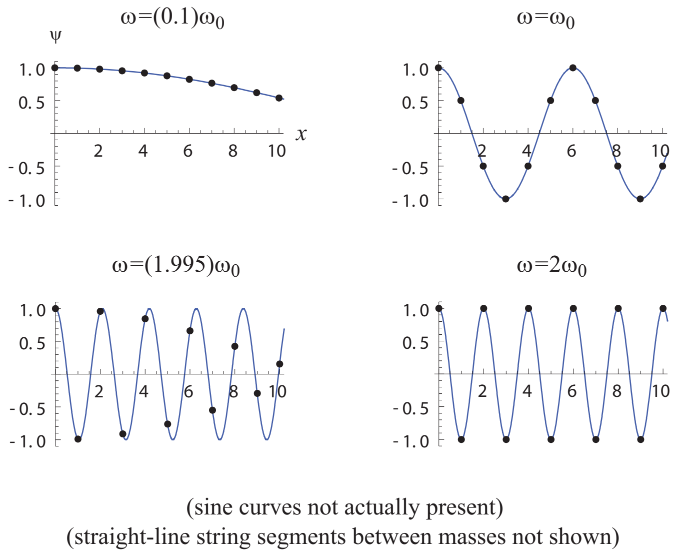


Figure 8

The cutoff case of $\omega = 2\omega_0$ can be considered to be either an evanescent wave or a traveling wave. So both of these cases must reduce to the same motion when $\omega = \omega_0$. Let's verify this. If we consider the wave to be an evanescent wave, then with $\omega = 2\omega_0$, Eq. (15) gives $\kappa = 0$. So there is no exponential decay, and the beads' positions in Eq. (16) simply alternate indefinitely to the right, in two horizontal lines. If we instead consider the wave to be a traveling wave, then with $\omega = 2\omega_0$, Eq. (9) gives $k\ell = \pi \Longrightarrow k = \pi/\ell$, which means that the wavelength approaches 2ℓ. The traveling wave at positions of the form $x = n\ell$ looks like

$$\begin{aligned} A\cos(kx - \omega t + \phi) &= A\cos\left(\frac{\pi}{\ell}\cdot n\ell - \omega t + \phi\right) \\ &= A\cos(n\pi - \omega t + \phi) \\ &= A\cos(n\pi)\cos(\omega t - \phi) \\ &= A(-1)^n\cos(\omega t - \phi), \end{aligned} \tag{17}$$

which agrees with Eq. (16) when $\kappa = 0$ (with a different definition of ϕ).

In the extreme case where $\omega \gg \omega_0$, Eq. (15) tells us that κ is large, which means that the exponential factor $e^{-\kappa n\ell}$ goes to zero very quickly. This makes sense. If you wiggle the end of the string very quickly (as always, we're assuming that the amplitude is small, in particular much smaller than the bead spacing), then the bead that you're holding onto will move with your hand, but all the other beads will hardly move at all. This is because they have essentially zero time to accelerate in one direction before you move your hand to the other side and change the direction of the force.

In practice, ω doesn't have to be much larger than $2\omega_0$ for this lack of motion to arise. Even if ω is only $4\omega_0$, then Eq. (15) gives the value of $\kappa\ell$ as 2.6. The amplitude of the $n = 1$ and $n = 2$ masses are then suppressed by a factors of $e^{-\kappa n\ell} = e^{-1(2.6)} \approx 1/14$, and $e^{-\kappa n\ell} = e^{-2(2.6)} \approx 1/200$. So if you grab the $n = 0$ mass at the end of the string and move it back and forth at frequency $4\omega_0$, you'll end up moving the $n = 1$ mass a little bit, but all the other masses will essentially not move.

Remark: For given values of T and μ, the relation $\omega_0 = (1/\ell)\sqrt{T/\mu}$ (see Eq. (11)) implies that if ℓ is very small, you need to wiggle the string very fast to get into the $\omega > 2\omega_0$ evanescent regime. In the limit of a continuous string ($\ell \to 0$), ω_0 is infinite, so you can never get to the evanescent regime. In other words, any wiggling that you do will produce a normal traveling wave. This makes intuitive sense. It also makes dimensional-analysis sense, for the following reason. Since a continuous string is completely defined in terms of the two parameters T (with units of kg m/s^2) and μ (with units of kg/m), there is no way to combine these parameters to form a quantity with the dimensions of frequency (that is, s^{-1}). So for a continuous string, there is therefore no possible frequency value that can provide the cutoff between traveling and evanescent waves. All waves must therefore fall into one of these two categories, and it happens to be the traveling waves. If, however, a length scale ℓ is introduced, then it *is* possible to generate a frequency (namely ω_0), which can provide the scale of the cutoff (which happens to be $2\omega_0$). ♣

Power

If you wiggle the end of a beaded string with a frequency larger than $2\omega_0$, so that an evanescent wave of the form in Eq. (16) arises, are you transmitting any net power to the string? Since the wave dies off exponentially, there is essentially no motion far to the right. Therefore, there is essentially no power being transmitted across a given point far to the right. So it had better be the case that you are transmitting zero net power to the string, because otherwise energy would be piling up indefinitely somewhere between your hand and the given point far to the right. This is impossible, because there is no place for this energy to go.[2]

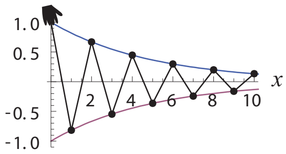


Figure 9

It is easy to see directly why you transmit zero net power over a full cycle (actually over each half cycle). Let's start with the position shown in Fig. 9, which shows a snapshot when the masses all have maximal displacement. The $n = 0$ mass is removed, and you grab the end of the string where that mass was. As you move your hand downward, you do negative work on the string, because you are pulling upward (and also horizontally, but there is no work associated with this force because your hand is moving only vertically), but your hand is moving in the opposite direction, downward. However, after your hand passes through equilibrium, it is still moving downward but now pulling downward too (because the string you're holding is now angled up to the right), so you are doing positive work on the string. The situation is symmetric (except for signs) on each side of equilibrium, so the positive and negative works cancel, yielding zero net work, as expected.

In short, your force is in "quadrature" with the velocity of your hand. That is, it is 90° out of phase with the velocity (behind it). So the product of the force and the velocity (which is the power) cancels over each half cycle. This is exactly the same situation that arises in a simple harmonic oscillator with a mass on a spring. You can verify that the force and velocity are in quadrature (the force is ahead now), and there is no net work done by the spring (consistent with the fact that the average motion of the mass does't change over time).

How do traveling waves ($\omega < 2\omega_0$) differ from evanescent waves ($\omega > 2\omega_0$), with regard to power? For traveling waves, when you wiggle the end, your force isn't in quadrature with the velocity of your hand, so you end up doing net positive work. In the small-ω limit (equivalently, the continuous-string limit), your force is exactly *in phase* with the velocity, so you're always doing positive work. However, as ω increases (with a beaded string), your force gradually shifts from being in phase with the velocity at $\omega \approx 0$, to being in quadrature with it at $\omega = 2\omega_0$, at which point no net work is being done. The task of Problem [to be added] is to be quantitative about this.

6.2.2 Low-frequency cutoff

The above beaded string supported traveling waves below a certain cutoff frequency but not above it. Let's now consider the opposite situation. We'll look at a system that supports traveling waves *above* a certain cutoff frequency but not below it. Hence the "Low-frequency cutoff" name of this subsection.

Consider a *continuous* string (not a beaded one) with tension T and density μ. Let the (infinite) string be connected to a wall by an essentially continuous set of springs, initially at their relaxed length, as shown in Fig. 10.

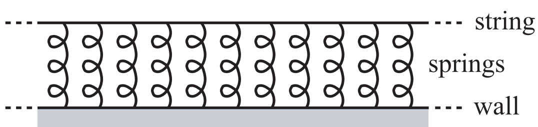


Figure 10

If the springs weren't present, then the return force (for transverse waves) on a little piece of the string with length Δx would be $T\Delta x(\partial^2\psi/\partial x^2)$; see Eq. (4.2). But we now also

[2]We are assuming steady-state motion. At the start, when you get the string going, you *are* doing net work on the string. Energy piles up at the start, because the string goes from being in equilibrium to moving back and forth. But in steady state, the average motion of the string doesn't change in time.

have the spring force, $-(\sigma\Delta x)\psi$, where σ is the spring constant per unit length. (The larger the piece, the more springs that touch it, so the larger the force.) The transverse $F = ma$ equation on the little piece is therefore modified from Eq. (4.3) to

$$\begin{aligned}(\mu\Delta x)\frac{\partial^2\psi}{\partial t^2} &= T\Delta x\frac{\partial^2\psi}{\partial x^2} - (\sigma\Delta x)\psi \\ \Longrightarrow \frac{\partial^2\psi}{\partial t^2} &= c^2\frac{\partial^2\psi}{\partial x^2} - \omega_s^2\psi, \qquad \text{where} \quad c^2 \equiv \frac{T}{\mu} \quad \text{and} \quad \omega_s^2 \equiv \frac{\sigma}{\mu}. \end{aligned} \tag{18}$$

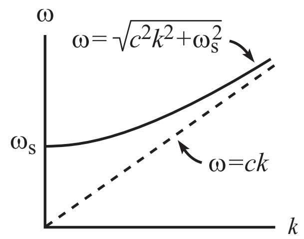


Figure 11

We will find that c is *not* the wave speed, as it was in the simple string system with no springs. To determine the dispersion relation associated with Eq. (18), we can plug in our standard exponential solution, $\psi(x,t) = Ae^{i(kx-\omega t)}$. This tells us that ω and k are related by

$$\boxed{\omega^2 = c^2k^2 + \omega_s^2} \qquad \text{(dispersion relation)} \tag{19}$$

This is the dispersion relation for the string-spring system. The plot of ω vs. k is shown in Fig. 11. There is no (real) value of k that yields a ω smaller than ω_s. However, there is an *imaginary* value of k that does. If $\omega < \omega_s$, then Eq. (19) gives

$$k = \frac{\sqrt{\omega^2 - \omega_s^2}}{c} \equiv i\kappa, \qquad \text{where} \quad \kappa \equiv \frac{\sqrt{\omega_s^2 - \omega^2}}{c}. \tag{20}$$

Another solution for κ is the negative of this one, but we'll be considering below the case where the string extends to $x = +\infty$, so this other solution would cause ψ to diverge, given our sign convention in the exponent of $e^{i(kx-\omega t)}$. Substituting $k \equiv i\kappa$ into $\psi(x,t) = Ae^{i(kx-\omega t)}$ gives

$$\psi(x,t) = Ae^{i((i\kappa)x-\omega t)} = Ae^{-\kappa x}e^{-i\omega t} \to \boxed{Be^{-\kappa x}\cos(\omega t + \phi)} \tag{21}$$

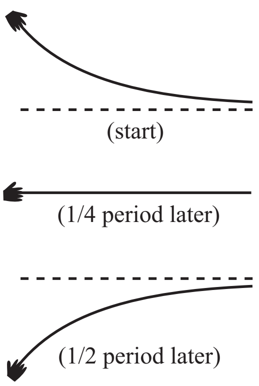


Figure 12

where as usual we have taken the real part. We see that $\psi(x,t)$ decays as a function of x, and that all points on the string oscillate with the *same* phase as a function of t. (This is in contrast with adjacent points having opposite phases in the above beaded-string setup. Opposite phases wouldn't make any sense here, because we don't have discrete adjacent points on a continuous string.) So we have an evanescent standing wave. If we wiggle the left end up and down sinusoidally with a frequency $\omega < \omega_s$, then snapshots of the motion take the general form shown in Fig. 12. The rate of the exponential decrease (as a function of x) depends on ω. If ω is only slightly smaller than ω_s, then the κ in Eq. (20) is small, so the exponential curve dies out very slowly to zero. In the limit where $\omega \approx 0$, we're basically holding the string at rest in the first position in Fig. 12, and you can show from scratch by balancing transverse forces in this static setup that the string does indeed take the shape of a decreasing exponential with $\kappa \approx \omega_s/c = \sqrt{\sigma/T}$; see Problem [to be added]. This static case yields the quickest spatial decay to zero.

It makes sense that the system doesn't support traveling waves for $\omega < \omega_s$, because even without any tension force, the springs would still make the atoms in the string oscillate with a frequency of at least $\sqrt{\sigma\Delta x/\mu\Delta x} = \omega_s$. Adding on a tension force will only increase the restoring force (at least in a traveling wave, where the curvature is always toward the x axis), and thus also the frequency. In the cutoff case where $\omega = \omega_s$, the string remains straight, and it oscillates back and forth as a whole, just as a infinite set of independent adjacent masses on springs would oscillate.

We have been talking about evanescent waves where $\omega < \omega_s$, but we can still have normal traveling waves if $\omega > \omega_s$, because the k in Eq. (20) is then real. If $\omega \gg \omega_s$, then Eq. (20) tells us that $\omega \approx ck$. In other words, the system acts like a simple string with no springs attached to it. This makes sense; on the short time scale of the oscillations, the spring doesn't have time to act, so it's effectively like it's not there. Equivalently, the transverse force from the springs completely dominates the spring force. If on the other hand ω is only slightly larger than ω_s, then Eq. (20) says that k is very small, which means that the wavelength is very large. In the $\omega \to \omega_s$ limit, we have $\lambda \to \infty$. The speed of the wave is then $\omega/k \approx \omega_s/k \approx \infty$. This can be seen graphically in Fig. 13, where the slope from the origin is ω/k, which is the phase velocity (just as it was in Fig. 3). This slope can be made arbitrarily large by making k be arbitrarily small. We'll talk more about excessively large phase velocities (in particular, larger than the speed of light) in Section 6.3.3.

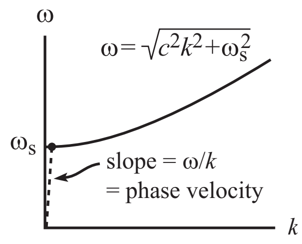


Figure 13

Note that the straight-line shape of the string in the $\omega = \omega_s$ case that we mentioned above can be considered to be the limit of a traveling wave with an infinitely long wavelength, and also an evanescent wave with an infinitely slow decay.

Power

As with the evanescent wave on the beaded string in Section 6.2.1, no net power can be transmitted in the present evanescent wave, because otherwise there would be energy piling up somewhere (because the wave dies out). But there is no place for it to pile up, because we are assuming steady-state motion. This can be verified with the same reasoning as in the beaded-string case; the net power you transmit to the string as you wiggle the left end alternates sign each quarter cycle, so there is complete cancelation over a full cycle.

Let's now consider the modified setup shown in Fig. 14. To the left of $x = 0$, we have a normal string with no springs. What happens if we have a rightward-traveling wave that comes in from the left, with a frequency $\omega < \omega_s$. (Or there could even be weak springs in the left region, as long as we have $\omega_{s,\text{left}} < \omega < \omega_{s,\text{right}}$. This would still allow a traveling wave in the left region.)

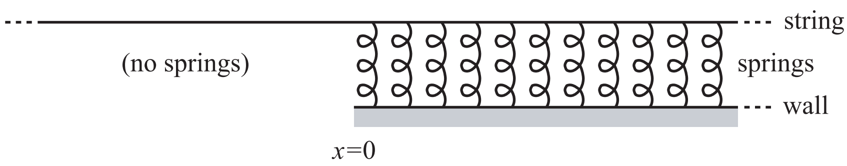


Figure 14

From the same reasoning as above, the fact that the wave dies out on the right side implies that no net power is transmitted along the string (in the steady state). However, there certainly *is* power transmitted in the incoming traveling wave. Where does it go? Apparently, there must be *complete* reflection at $x = 0$, so that all the power gets reflected back. The spring region therefore behaves effectively like a brick wall, as far as reflection goes. But the behavior isn't exactly like a brick wall, because ψ isn't constrained to be zero at the boundary in the present case.

To figure out what the complete wave looks like, we must apply the boundary conditions (continuity of the function and the slope) at $x = 0$.[3] If we work with exponential solutions, then the incoming, reflected, and transmitted waves take the form of $Ae^{i(\omega t-kx)}$, $Be^{i(\omega t+kx)}$, and $Ce^{i\omega t}e^{-\kappa x}$, respectively. The goal is to solve for B and C (which may be complex) in terms of A, that is, to solve for the ratios B/A and C/A. For ease of computation, it is

[3]As usual, the continuity of the slope follows from the fact that there can be no net force on the essentially massless atom at the boundary. The existence of the springs in the right region doesn't affect this fact.

customary to divide all of the functions by A, in which case the total waves on the two sides of the boundary can be written as

$$\begin{aligned}\psi_{\rm L} &= (e^{-ikx} + Re^{ikx})e^{i\omega t},\\ \psi_{\rm R} &= De^{-\kappa x}e^{i\omega t},\end{aligned} \tag{22}$$

where $R \equiv B/A$ and $D \equiv C/A$. R is the complex reflection coefficient. Its magnitude is what we normally call the "reflection coefficient," and its phase yields the phase of the reflected wave when we eventually take the real part to obtain the physical solution. The task of Problem [to be added] is to apply the boundary conditions and solve for R and D.

Tunneling

What happens if we have a setup in which the region with strings is finite, as shown in Fig. 15? If a rightward-traveling wave with $\omega < \omega_{\rm s}$ comes in from the left, some of it will make it through, and some of it will be reflected. That is, there will be a nonzero rightward-traveling wave in the right region, and there will be the usual combination of a traveling and standing wave in the left region that arises from partial reflection.

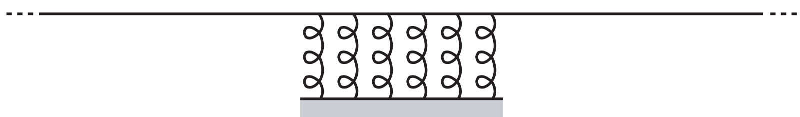


Figure 15

The nonzero wave in the right region implies that power is transmitted. This is consistent with the fact that we *cannot* use the reasoning that zero power follows from the fact that the wave dies out to zero; the wave doesn't die out to zero, because the middle region has finite length. We can't rule out the $e^{+\kappa x}$ solution, because the $x = +\infty$ boundary condition isn't relevant now. So the general (steady state) solution in the left, middle, and right regions takes the form,

$$\begin{aligned}\psi_{\rm L} &= (e^{-ikx} + Re^{ikx})e^{i\omega t},\\ \psi_{\rm M} &= (Be^{-\kappa x} + Ce^{\kappa x})e^{i\omega t},\\ \psi_{\rm R} &= Te^{-ikx}e^{i\omega t},\end{aligned} \tag{23}$$

where R and T are the complex reflection and transmission coefficients, defined to be the ratio of the reflected and transmitted (complex) amplitudes to the incident (complex) amplitude. As in Eq. (22), we have written the waves in their complex forms, with the understanding that we will take the real part to find the actual wave. The four boundary conditions (value and slope at each of the two boundaries) allow us to solve for the four unknowns (R, T, B, C). This is the task of Problem [to be added]; the math gets a bit messy.

The effect where some of the wave makes it through the "forbidden" region where traveling waves don't exist is known as *tunneling*. The calculation in Problem [to be added] is exactly the same as in a *quantum mechanical* system involving tunneling through a classically forbidden region (a region where the total energy is less than the potential energy). In quantum mechanics, the waves are probability waves instead of transverse string waves, so the interpretation of the waves is different. But all the math is exactly the same as in the above string-spring system. We'll talk much more about quantum-mechanical waves in Chapter 11.

6.3 Group velocity

6.3.1 Derivation

Whether a system is dispersionless (with a linear relationship between ω and k) or dispersionful (with a nonlinear relationship between ω and k), the phase velocity in both cases is $v_\text{p} = \omega/k$. The phase velocity gives the speed of a *single* sinusoidal traveling wave. But what if we have, say, a wave in the form of a lone bump, which (from Fourier analysis) can be written as the sum of many (or an infinite number of) sinusoidal waves. How fast does this bump move? If the system is dispersionless, then all of the wave components move with the same speed v_p, so the bump also moves with this speed. We discussed this effect in Section 2.4 (see Eq. (2.97)), where we noted that any function of the form $f(x-ct)$ is a solution to a *dispersionless* wave equation, that is, an equation of the form $\partial^2\psi/\partial t^2 = c^2(\partial^2\psi/\partial x^2)$. This equation leads to the relation $\omega = ck$, where c takes on a single value that is independent of ω and k.

However, if the system is dispersionful, then ω/k depends on k (and ω), so the different sinusoidal waves that make up the bump travel at different speeds. So it's unclear what the speed of the bump is, or even if the bump *has* a well-defined speed. It turns out that it does in fact have a well-defined speed, and it is given by the slope of the $\omega(k)$ curve:

$$\boxed{v_\text{g} = \frac{d\omega}{dk}} \tag{24}$$

This is called the *group velocity*, which is a sensible name considering that a bump is made up of a group of Fourier components, as opposed to a single sinusoidal wave. Although the components travel at different speeds, we will find below that they conspire in such a way as to make their sum (the bump) move with speed $v_\text{g} = d\omega/dk$. However, an unavoidable consequence of the differing speeds of the components is the fact that as time goes on, the bump will shrink in height and spread out in width, until you can hardly tell that it's a bump. In other words, it will disperse. Hence the name dispersion.

Since the bump consists of wave components with many different values of k, there is an ambiguity about which value of k is the one where we should evaluate $v_\text{g} = d\omega/dk$. The general rule is that it is evaluated at the value of k that dominates the bump. That is, it is evaluated at the peak of the Fourier transform of the bump.

We'll now derive the result for the group velocity in Eq. (24). And because it is so important, we derive it in three ways.

First derivation

Although we just introduced the group velocity by talking about the speed of a bump, which consists of many Fourier components, we can actually understand what's going on by considering just two waves. Such a system has all the properties needed to derive the group velocity. So consider the two waves:

$$\begin{aligned} \psi_1(x,t) &= A\cos(\omega_1 t - k_1 x), \\ \psi_2(x,t) &= A\cos(\omega_2 t - k_2 x). \end{aligned} \tag{25}$$

It isn't necessary that they have equal amplitudes, but it simplifies the discussion. Let's see what the sum of these two waves looks like. It will be advantageous to write the ω's and

k's in terms of their averages and differences.[4] So let's define:

$$\omega_+ \equiv \frac{\omega_1+\omega_2}{2}, \qquad \omega_- \equiv \frac{\omega_1-\omega_2}{2}, \qquad k_+ \equiv \frac{k_1+k_2}{2}, \qquad k_- \equiv \frac{k_1-k_2}{2}. \tag{26}$$

Then $\omega_1 = \omega_+ + \omega_-$ and $\omega_2 = \omega_+ - \omega_-$. And likewise for the k's. So the sum of the two waves can be written as

$$\begin{aligned}\psi_1(x,t)+\psi_2(x,t) &= A\cos\Big((\omega_+ + \omega_-)t - (k_+ + k_-)x\Big)\\ &\quad + A\cos\Big((\omega_+ - \omega_-)t - (k_+ - k_-)x\Big)\\ &= A\cos\Big((\omega_+ t - k_+ x) + (\omega_- t - k_- x)\Big)\\ &\quad + A\cos\Big((\omega_+ t - k_+ x) - (\omega_- t - k_- x)\Big)\\ &= 2A\cos\big(\omega_+ t - k_+ x\big)\cos\big(\omega_- t - k_- x\big),\end{aligned} \tag{27}$$

where we have used $\cos(\alpha\pm\beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$. We see that the sum of the two original traveling waves can be written as the product of two other traveling waves.

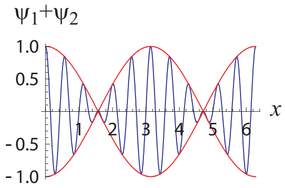


Figure 16

This is a general result for any values of ω_1, ω_2, k_1, and k_2. But let's now assume that ω_1 is very close to ω_2 (more precisely, that their difference is small compared with their sum). And likewise that k_1 is very close to k_2. We then have $\omega_- \ll \omega_+$ and $k_- \ll k_+$. Under these conditions, the sum $\psi_1+\psi_2$ in Eq. (27) is the product of the quickly-varying (in both space and time) wave, $\cos\big(\omega_+ t - k_+ x\big)$, and the slowly-varying wave, $\cos\big(\omega_- t - k_- x\big)$. A snapshot (at the arbitrarily-chosen time of $t = 0$) of $\psi_1+\psi_2$ is shown in Fig. 16. The quickly-varying wave is the actual sum, while the slowly-varying envelope is the function $2A\cos\big(\omega_- t - k_- x\big)$. We have arbitrarily picked $2A = 1$ in the figure. And we have chosen $k_1 = 10$ and $k_2 = 12$, which yield $k_+ = 11$ and $k_- = 1$. So the envelope function is $\cos(x)$, and the wiggly function (which equals $\psi_1+\psi_2$) is $\cos(11x)\cos(x)$.

At t increases, the quickly- and slowly-varying waves will move horizontally. What are the velocities of these two waves? The velocity of the quickly-wiggling wave is ω_+/k_+, which is essentially equal to either of ω_1/k_1 and ω_2/k_2, because we are assuming $\omega_1 \approx \omega_2$ and $k_1 \approx k_2$. So the phase velocity of the quickly-wiggling wave is essentially equal to the phase velocity of either wave.

The velocity of the slowly-varying wave (the envelope) is

$$\frac{\omega_-}{k_-} = \frac{\omega_1-\omega_2}{k_1-k_2}. \tag{28}$$

(Note that this may be negative, even if the phase velocities of the original two waves are both positive.) If we have a linear dispersion relation, $\omega = ck$, then this speed equals $c(k_1-k_2)/(k_1-k_2) = c$. So the group velocity equals the phase velocity, and this common velocity is constant, independent of k. But what if ω and k aren't related linearly? Well, if ω is given by the function $\omega(k)$, and if k_1 is close to k_2, then $(\omega_1-\omega_2)/(k_1-k_2)$ is essentially equal to the derivative, $d\omega/dk$. This is the velocity of the envelope formed by the two waves, and it is called the the *group velocity*. To summarize:

$$\boxed{v_\text{p} = \frac{\omega}{k} \qquad \text{and} \qquad v_\text{g} = \frac{d\omega}{dk}} \tag{29}$$

In general, both of these velocities are functions of k. And in general they are not equal. (The exception to both of these statements occurs in the case of linear dispersion.) In the

[4] We did something similar to this when we talked about beats in Section 2.1.4. But things are a little different here because the functions are now functions of both x and t, as opposed to just t.

general case where $v_g \neq v_p$, the fast wiggles in Fig. 16 move with respect to the envelope. If $v_p > v_g$, then the little wiggles pop into existence at the left end of an envelope bump (or the right end if $v_p < v_g$). They grow and then shrink as they move through the bump, until finally they disappear when they reach the right end of the bump.

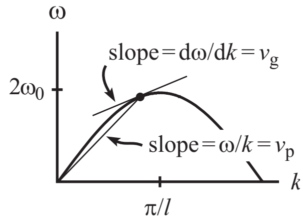


Figure 17

In the case of the beaded-string system discussed in Section 6.1, the plot of $\omega(k)$ was shown in Fig. 3. So the phase and group velocities are shown graphically in Fig. 17. For $0 < k < \pi/\ell$ (which is generally the part of the graph we're concerned with), the slope of the curve at any point is less than the slope of the line from the origin to the point, so we see that v_g is always less than v_p.

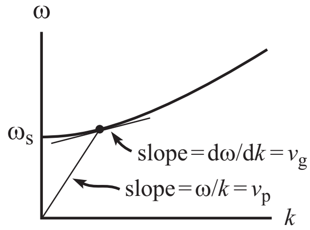


Figure 18

In the case of the string/spring system discussed in Section 6.2.2, the plot of $\omega(k)$ was shown in Fig. 11. So the phase and group velocities are shown graphically in Fig. 18. We again see that v_g is always less than v_p. However, this need not be true in general, as we'll see in the examples in Section 6.3.2 below. For now, we'll just note that in the particular case where the plot of $\omega(k)$ passes through the origin, there are two basic possibilities of what the $\omega(k)$ curve can look like, depending on whether it is concave up or down. These are shown in Fig. 19. The first case always has $v_g > v_p$, while the second always has $v_g < v_p$. In the first case, sinusoidal waves with small k (large λ) travel slower (that is, they have a smaller v_p) than waves with large k (small λ). The opposite is true in the second case.

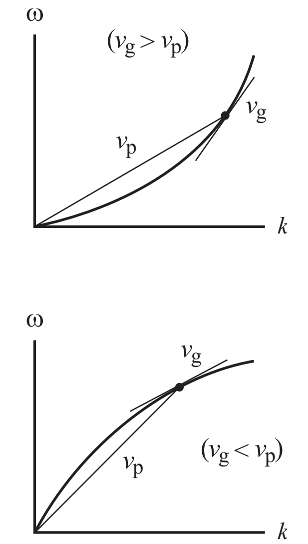


Figure 19

Second derivation

Consider two waves with different values of ω and k, as shown in the first pair of waves in Fig. 20. These two waves constructively interfere at the dots, so there will be a bump there. When and where does the next bump occur? If we can answer these questions, then we can find the effective velocity of the bump.

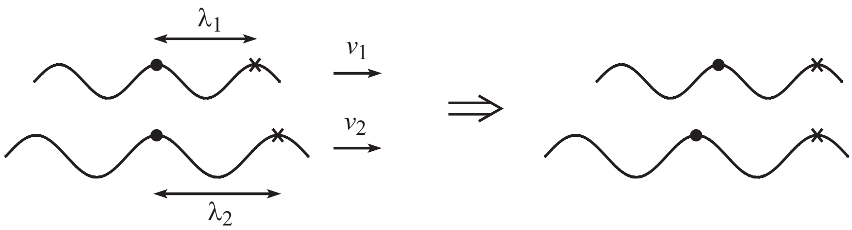


Figure 20

If $v_1 = v_2$ (that is, if $\omega_1/k_1 = \omega_2/k_2$), then both waves travel at the same speed, so the bump simply travels along with the waves, at their common speed. But if $v_1 \neq v_2$, then the dots will become unaligned. If we assume that $v_1 > v_2$ (the $v_1 < v_2$ case is similar), then at some later time the next two peaks will line up, as shown in the second pair of waves in Fig. 20. These peaks are marked with x's. There will then be a bump at this new location. (If $v_1 < v_2$, the next alignment will occur to the left of the initial one.)

When do these next peaks line up? The initial distance between the x's is $\lambda_2 - \lambda_1$, and the top wave must close this gap at a relative speed of $v_1 - v_2$, so $t = (\lambda_2 - \lambda_1)/(v_1 - v_2)$. Equivalently, just set $\lambda_1 + v_1 t = \lambda_2 + v_2 t$, because these two quantities represent the positions of the two x's, relative to the initial dots. Having found the time t, the position of the next alignment is given by $x = \lambda_1 + v_1 t$ (and also $\lambda_2 + v_2 t$). The velocity at which the bump effectively travels is therefore

$$\frac{x}{t} = \frac{\lambda_1 + v_1 t}{t} = \frac{\lambda_1}{t} + v_1 = \lambda_1\left(\frac{v_1 - v_2}{\lambda_2 - \lambda_1}\right) + v_1 = \frac{\lambda_1 v_1 - \lambda_1 v_2}{\lambda_2 - \lambda_1} + \frac{\lambda_2 v_1 - \lambda_1 v_1}{\lambda_2 - \lambda_1}$$

$$= \frac{\lambda_2 v_1 - \lambda_1 v_2}{\lambda_2 - \lambda_1} = \frac{\frac{2\pi}{k_2}\frac{\omega_1}{k_1} - \frac{2\pi}{k_1}\frac{\omega_2}{k_2}}{\frac{2\pi}{k_2} - \frac{2\pi}{k_1}} = \frac{\frac{\omega_1 - \omega_2}{k_1 k_2}}{\frac{k_1 - k_2}{k_1 k_2}} = \frac{\omega_1 - \omega_2}{k_1 - k_2} \equiv v_{\rm g}. \tag{30}$$

This is the same speed we obtained in Eq. (28). This is no surprise, because we basically did the same calculation here. In the previous derivation, we assumed that the waves were nearly identical, whereas we didn't assume that here. This assumption isn't needed for the $v_{\rm g} = (\omega_1 - \omega_2)/(k_1 - k_2)$ result. Whenever and wherever a bump in the present derivation exists, it touches the top of the envelope curve (if we had drawn it). So what we effectively did in this derivation is find the speed of the envelope curve. But this is exactly what we did in the previous derivation.

In between the alignments of the dot and the x in Fig. 20, the bump disappears, then appears in the negative direction, then disappears again before reappearing at the x. This is consistent with the fact that the wiggly wave in Fig. 16 doesn't always (in fact, rarely does) touch the midpoint (the highest point) of the envelope bump. But on average, the bump effectively moves with velocity $v_{\rm g} = (\omega_1 - \omega_2)/(k_1 - k_2)$.

Note that if k_1 is very close to k_2, and if ω_1 is *not* very close to ω_2, then $v_{\rm g} = (\omega_1 - \omega_2)/(k_1 - k_2)$ is large. It is easy to see intuitively why this is true. We may equivalently describe this scenario by saying that λ_1 is very close to λ_2, and that v_1 is not very close to v_2 (because $v = \omega k$). The nearly equal wavelengths imply that in Fig. 20 the two x's are very close together. This means that it takes essentially no time for them to align (because the velocities aren't close to each other). The location of the alignment therefore jumps ahead by a distance of one wavelength in essentially no time, which means that the effective speed is large (at least as large as the λ_1/t term in Eq. (30)).

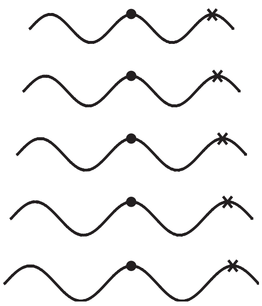


Figure 21

What if we have a large number of waves with roughly the same values of k (and hence ω), with a peak of each wave lining up, as shown by the dots in Fig. 21? Since the plot of $\omega(k)$ at any point is locally approximately a straight line, the quotient $(\omega_1 - \omega_2)/(k_1 - k_2)$, which is essentially equal to the derivative $d\omega/dk$, is the same for all nearby points, as shown in Fig. 22. This means that the next bumps (the x's in Fig. 21) will all line up at the same time and at the same place, because the location of all of the alignments is given by $x = v_{\rm g}t$, by Eq. (30). In other words, the group velocity $v_{\rm g}$ is well defined. The various waves all travel with different phase velocities $v_{\rm p} = \omega/k$, but this is irrelevant as far as the group velocity goes, because $v_{\rm g}$ depends on the *differences* in ω and k through Eq. (30), and not on the actual values of ω and k.

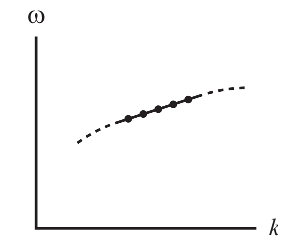


Figure 22

Third derivation

By definition, $v_{\rm g}$ is the velocity at which a bump in a wave travels. From Fourier analysis, we know that in general a wave consists of components with many different frequencies. If these components are to "work together" to form a bump at a certain location, then the phases $\omega_i t - k_i t + \phi_i$ of the different components (or at least many of them) must be equal at the bump, if they are to add constructively to form the bump.

Assume that we have a bump at a particular value of x and t. We are free to pick the origins of x and t to be wherever and whenever we want, so let's pick the bump to be located at $x = 0$ and $t = 0$. Since the phases $\omega_i t - k_i t + \phi_i$ are all equal (or at least many of them) in a region around some particular k value, we conclude that the ϕ_i are all equal, because $x = t = 0$. In other words, ϕ is independent of k.[5]

At what other values of x and t, besides $(x, t) = (0, 0)$, is there a bump? That is, at what other values of x and t are the phases still all equal? Well, we want $\omega t - kx + \phi$ to be independent of k near some particular k value, because then the phases of the waves for

[5]You can check that the following derivation still works in the case of general initial coordinates (x_0, t_0); see Problem [to be added]. But it's less messy if we choose $(0, 0)$.

all the different k values will be equal, which means that the waves will add constructively and therefore produce another bump. (We have dropped the index i on ω and k, and it is understood that ω is a function $\omega(k)$ of k.) Demanding that the phase be independent of k gives

$$0 = \frac{d(\omega t - kx + \phi)}{dk} \quad \Longrightarrow \quad 0 = \frac{d\omega}{dk}t - x \quad \Longrightarrow \quad \frac{x}{t} = \frac{d\omega}{dk}\,, \tag{31}$$

where we have used $d\phi/dk = 0$. So we have a bump at any values of x and t satisfying this relation. In other words, the speed of the bump is

$$v_{\text{g}} = \frac{d\omega}{dk}\,, \tag{32}$$

in agreement with the result from the previous derivations.

Note that the *phase* velocity (of single traveling wave) is obtained by demanding that the phase $\omega t - kx + \phi$ of the wave be independent of *time*:

$$0 = \frac{d(\omega t - kx + \phi)}{dt} \quad \Longrightarrow \quad 0 = \omega - k\frac{dx}{dt} \quad \Longrightarrow \quad v_{\text{p}} \equiv \frac{dx}{dt} = \frac{\omega}{k}\,. \tag{33}$$

But the *group* velocity (of a group of traveling waves) is obtained by demanding that the phase $\omega t - kx + \phi$ of all the different waves be independent of the *wavenumber* k:

$$0 = \frac{d(\omega t - kx + \phi)}{dk} \quad \Longrightarrow \quad 0 = \frac{d\omega}{dk}t - x \quad \Longrightarrow \quad v_{\text{g}} \equiv \frac{x}{t} = \frac{d\omega}{dk}\,. \tag{34}$$

Just because the quantity $d\omega/dk$ exists, there's no guarantee that there actually *will* be a noticeable bump traveling at the group velocity v_{g}. It's quite possible (and highly likely if things are random) that there is no constructive interference anywhere. But what we showed above was that *if* there is a bump at a given time and location, then it travels with velocity $v_{\text{g}} = d\omega/dk$, evaluated at the k value that dominates the bump. This value can be found by calculating the Fourier transform of the bump.

6.3.2 Examples

Beaded string

We discussed the beaded string in Section 6.1. Eq. (9) gives the dispersion relation as $\omega(k) = 2\omega_0 \sin(k\ell/2)$, where $\omega_0 \equiv \sqrt{T/m\ell}$. Therefore,

$$v_{\text{p}} = \frac{\omega}{k} = \frac{2\omega_0 \sin(k\ell/2)}{k}\,, \qquad \text{and} \qquad v_{\text{g}} = \frac{d\omega}{dk} = \omega_0 \ell \cos(k\ell/2). \tag{35}$$

For small k (more precisely, for $k\ell \ll 1$), we can use $\sin\epsilon \approx \epsilon$ and $\cos\epsilon \approx 1$, to quickly show that

$$\begin{aligned} v_{\text{p}} &\approx \frac{2\omega_0(k\ell/2)}{k} = \omega_0 \ell = \sqrt{\frac{T}{m\ell}}\,\ell = \sqrt{\frac{T}{m/\ell}} = \sqrt{\frac{T}{\mu}}\,, \\ v_{\text{g}} &\approx \omega_0 \ell(1) = \sqrt{\frac{T}{\mu}}\,. \end{aligned} \tag{36}$$

As expected, these both agree with the (equal) phase and group velocities for a continuous string, because $k\ell \ll 1$ implies $\ell \ll \lambda$, which means that the string is essentially continuous on a length scale of the wavelength.

String/spring system

We discussed the string/spring system in Section 6.2.2. Eq. (19) gives the dispersion relation as $\omega^2 = c^2k^2 + \omega_s^2$, where $c^2 \equiv T/\mu$ and $\omega_s^2 \equiv \sigma/\mu$. Therefore,

$$v_p = \frac{\omega}{k} = \frac{\sqrt{c^2k^2 + \omega_s^2}}{k}, \qquad \text{and} \qquad v_g = \frac{d\omega}{dk} = \frac{c^2k}{\sqrt{c^2k^2 + \omega_s^2}}. \tag{37}$$

If $\omega_s \approx 0$, then these reduce to $v_p \approx c$ and $v_g \approx c$, as expected. If ω_s is large (more precisely, if $\omega_s \gg ck$), then v_p is large and v_g is small (more precisely, $v_p \gg c$ and $v_g \ll c$). These facts are consistent with the slopes in Fig. 18.

Stiff string

When dealing with uniform strings, we generally assume that they are perfectly flexible. That is, we assume that they don't bounce back when they are bent. But if we have a "stiff string" that offers resistance when bent, it can be shown that the wave equation picks up an extra term and now takes the form,

$$\frac{\partial^2\psi}{\partial t^2} = c^2\left[\frac{\partial^2\psi}{\partial x^2} - \alpha\left(\frac{\partial^4\psi}{\partial x^4}\right)\right],$$

where α depends on various things (the cross-sectional area, Young's modulus, etc.).[6] Plugging in $\psi(x,t) = Ae^{i(\omega t - kx)}$ yields the dispersion relation,

$$\omega^2 = c^2k^2 + \alpha c^2k^4 \quad \Longrightarrow \quad \omega = ck\sqrt{1 + \alpha k^2}. \tag{38}$$

This yields

$$v_p = \frac{\omega}{k} = c\sqrt{1 + \alpha k^2}, \qquad \text{and} \qquad v_g = \frac{d\omega}{dk} = \frac{c(1 + 2\alpha k^2)}{\sqrt{1 + \alpha k^2}}. \tag{39}$$

The dispersion relation in Eq. (38) has implications in piano tuning, because although the strings in a piano are reasonably flexible, they aren't perfectly so. They are slightly stiff, with a small value of α. If they were perfectly flexible ($\alpha = 0$), then the linear dispersion relation, $\omega = ck$, would imply that the standing-wave frequencies are simply proportional to the mode number, n, because the wavenumbers take the usual form of $k = n\pi/L$. So the "first harmonic" mode ($n = 2$) would have twice the frequency of the fundamental mode ($n = 1$). In other words, it would be an octave higher.

However, for a stiff string ($\alpha \neq 0$), Eq. (38) tells us that the frequency of the first harmonic is larger than twice the frequency of the fundamental. (The k values still take the form of $k = n\pi/L$. This is a consequence of the boundary conditions and is independent of the dispersion relation.)

Consider two notes that are an "octave" apart on the piano (the reason for the quotes will soon be clear). These notes are in tune with each other if the *first harmonic* of the lower string equals the *fundamental* of the higher string. Your ear then won't hear any beats between these two modes when the strings are played simultaneously, so things will sound nice.[7] A piano is therefore tuned to make the first harmonic of the lower string equal to

[6] In a nutshell, the fourth derivative comes from the facts that (1) the resistance to bending (the so-called "bending moment") is proportional to the curvature, which is the second derivative of ψ, and (2) the resulting net transverse force can be shown to be proportional to the second derivative of the bending moment.

[7] Your ear only cares about beats between nearby frequencies. The relation between the two fundamentals is irrelevant because they are so far apart. Beats don't result from widely-different frequencies.

the fundamental of the higher string. But since the dispersion relation tells us that the first harmonic (of any string) has more than twice the frequency of the fundamental, we conclude that the spacing between the fundamentals of the two strings is larger than an octave. But this is fine because it's what your ear wants to hear. The (equal) relation between the first harmonic of the lower string and the fundamental of the higher string is what's important. The relation between the two fundamentals doesn't matter.[8]

Power law

If a dispersion relation takes the form of a power law, $\omega = Ak^r$, then

$$v_{\rm p} = \frac{\omega}{k} = Ak^{r-1}, \qquad \text{and} \qquad v_{\rm g} = \frac{d\omega}{dk} = rAk^{r-1}. \tag{40}$$

We see that $v_{\rm g} = rv_{\rm p}$ for any value of k. If $r = 1$, then we have a dispersionless system. If $r > 1$, then the dispersion curve is concave up, so it looks like the first plot we showed in Fig. 19, with $v_{\rm g} > v_{\rm p}$. Sinusoidal waves with small k travel slower than waves with large k. If $r < 1$, then we have the second plot in Fig. 19, and these statements are reversed.

Quantum mechanics

In nonrelativistic quantum mechanics, particles are replaced by probability waves. The wave equation (known as the Schrodinger equation) for a free particle moving in one dimension happens to be

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2}\,, \tag{41}$$

where $\hbar = 1.05 \cdot 16^{-34}\text{J}\cdot\text{s}$ is Planck's constant. Plugging in $\psi(x,t) = Ae^{i(\omega t - kx)}$ yields the dispersion relation,

$$\omega = \frac{\hbar k^2}{2m}\,. \tag{42}$$

We'll give an introduction to quantum mechanics in Chapter 12, but for now we'll just note that the motivation for the dispersion relation (and hence the wave equation) comes from the substitutions of $E = \hbar\omega$ and $p = \hbar k$ into the standard classical relation, $E = p^2/2m$. We'll discuss the origins of these forms of E and p in Chapter 12.

The dispersion relation gives

$$v_{\rm p} = \frac{\omega}{k} = \frac{\hbar k}{2m}\,, \qquad \text{and} \qquad v_{\rm g} = \frac{d\omega}{dk} = \frac{\hbar k}{m}\,. \tag{43}$$

Classically, the velocity of a particle is given by $v = p/m$. So if $p = \hbar k$, then we see that $v_{\rm g}$, and not $v_{\rm p}$, corresponds to the classical velocity of a particle. This is consistent with the fact that a particle can be thought of as a localized bump in the probability wave, and this bump moves with the group velocity $v_{\rm g}$. A single sinusoidal wave moving with velocity $v_{\rm p}$ doesn't correspond to a localized particle, because the wave (which represents the probability) extends over all space. So we shouldn't expect $v_{\rm p}$ to correspond to the standard classical velocity of p/m.

[8]A nice article on piano tuning is: *Physics Today*, December 2009, pp 46-49. It's based on a letter from Richard Feynman to his piano tuner. See in particular the "How to tune a piano" box on page 48.

Water waves

We'll discuss water waves in detail in Chapter 11, but we'll invoke some results here so that we can see what a few phase and group velocities look like. There are three common types of waves:

- *Small ripples*: If the wavelength is short enough so that the effects of surface tension dominate the effects of gravity, then the dispersion relation takes the form, $\omega = \sqrt{\sigma k^3/\rho}$, where σ is the surface tension and ρ is the mass density. The surface tension dominates if the wavelength be small compared with about 2 cm. The dispersion relation then gives

$$v_{\text{p}} = \frac{\omega}{k} = \sqrt{\frac{\sigma k}{\rho}}\,, \qquad \text{and} \qquad v_{\text{g}} = \frac{d\omega}{dk} = \frac{3}{2}\sqrt{\frac{\sigma k}{\rho}}\,. \tag{44}$$

  So $v_{\text{g}} = 3v_{\text{p}}/2$. The smaller the wavelength (the larger the k), then the larger the v_{p}. Very small ripples travel fast.

- *Long wavelengths in deep water*: If the wavelength is large compared with 2 cm, then the effects of gravity dominate. If we further assume that the wavelength is small compared with the depth of the water, then the dispersion relation takes the form, $\omega = \sqrt{gk}$. This gives

$$v_{\text{p}} = \frac{\omega}{k} = \sqrt{\frac{g}{k}}\,, \qquad \text{and} \qquad v_{\text{g}} = \frac{d\omega}{dk} = \frac{1}{2}\sqrt{\frac{g}{k}}\,. \tag{45}$$

  So $v_{\text{g}} = v_{\text{p}}/2$. The larger the wavelength (the smaller the k), then the larger the v_{p}. Long waves travel fast.

- *Long wavelengths, compared with depth*: If the wavelength is large compared with the depth of the water, then the dispersion relation takes the form, $\omega = \sqrt{gH}\,k$, where H is the depth. This is a dispersionless system, with

$$v_{\text{p}} = v_{\text{g}} = \sqrt{gH}. \tag{46}$$

  So all waves travel with the same speed (provided that the wavelength is large compared with H). This has dramatic consequences with regard to tsunamis.

  Consider a huge wave that is created in the ocean, most commonly by an earthquake. If the wave has, say, an amplitude of 100 ft and a wavelength of half a mile (which is indeed a huge wave), what will happen as it travels across the ocean? The depth of the ocean is on the order of a few miles, so we're in the regime of "long wavelengths in deep water." From above, this is a *dispersive* system. Different wavelengths therefore travel with different speeds, and the wave disperses. It grows shallower and wider, until there is hardly anything left. When it reaches the other side of the ocean, it will be barely distinguishable from no wave at all. The fast Fourier components of the initial bump (the ones with long wavelengths) will arrive much sooner than the slower components, so the energy of the wave will be diluted over a long period of time.

  However, consider instead a wave with an amplitude of only 5 ft, but with a wavelength of 10 miles. (Assuming roughly the same shape, this wave has the same volume as the one above.) What will happen to this wave as it travels across the ocean? We're now in the "long wavelengths, compared with depth" regime. This is a *nondispersive* system, so all of the different Fourier components of the initial "bump" travel with the same speed. The wave therefore keeps the same shape as it travels across the ocean.

Now, a 5 ft wave might not seem severe, but when the wave reaches shallower water near the shore, its energy gets concentrated into a smaller region, so the amplitude grows. If the boundary between the ocean and land were a hypothetical vertical wall extending miles downward, then the waves would simply reflect off the wall and travel back out to sea. But in reality the boundary is sloped. In short, the very long wavelength allows the wave to travel intact all the way across the ocean, and the sloped shore causes the amplitude to grow once the wave arrives.

What is the speed of a tsunami wave in deep water? The average depth of the Pacific Ocean is about 4000 m, so we have $v = \sqrt{gH} \approx 200\,\text{m/s} \approx 450\,\text{mph}$, which is quite fast. It takes only a little over a minute for all of the 10-mile wave to hit the shore. So the energy is deposited in a short amount of time. It isn't diluted over a large time as it was with the half-mile wave above. Note that in contrast with the dramatic effects at the shore, the wave is quite unremarkable far out to sea. It rises to a height of 5 ft over the course of many miles, so the slope at any point is extremely small. It is impossible to spot such a wave visually, but fortunately deep-sea pressure sensors on the ocean floor can measure changes in the water level with extreme precision.

6.3.3 Faster than c?

Group velocity

In the first derivation of the group velocity in Section 6.3.1, we found the velocity of the envelope to be $v_g = (\omega_2 - \omega_1)/(k_2 - k_1)$. If $\omega_1 \neq \omega_2$, and if the k values are close together, then v_g is large. In fact, v_g can be made arbitrarily large by making k_2 be arbitrarily close to k_1. This means that it is possible for v_g to exceed the speed of light. Is this bad? Does it mean that we can send a signal faster than the speed of light (commonly denoted by "c"), which would violate the theory of relativity? Answers: No, No.

In order for this scenario to be possible, the two individual waves we used in the derivation of Eq. (28) already needed to be in existence over a very wide range of x values. So the envelope in Fig. 16 is going to travel with velocity v_g, independent of what anyone does. Two people therefore can't use this effect to send information. To communicate something, you need to *change* the wave, and it can be demonstrated that the *leading edge* of this change can never travel faster than c.

Demonstrating this fact requires invoking some facts about relativity. It certainly can't be demonstrated *without* invoking anything about relativity, because there is nothing at all special about the speed of light if you haven't invoked the postulates of relativity. One line of reasoning is to say that if the leading edge travels faster than c, then there exists a frame in which causality is violated. This in turn violates innumerable laws of physics ($F = ma$ type laws, for example). The fact of the matter is that a line of atoms can't talk to each other faster than c, independent of whether they're part of a wave.

At any rate, the point of the present discussion isn't so much to say what's right about the relativistic reasoning (because we haven't introduced relativity in this book), but rather to say what's wrong about the reasoning that $v_g > c$ implies a contradiction with relativity. As we stated above, the error is that no information is contained in a wave that is formed from the superposition of two waves that already existed over a wide range of x values. If you had set up these waves, you must have set them up a while ago.

Another way to have $v_g > c$ is shown in Fig. 23. Because the bump creeps forward, it might be possible for v_g to exceed c even if the speed of the leading edge doesn't. But that's fine. The leading edge is what is telling someone that something happened, and this speed never exceeds c.

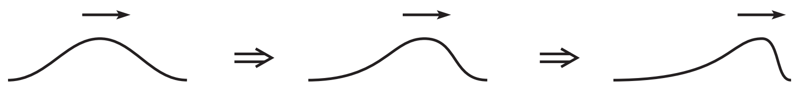


Figure 23

Phase velocity

The phase velocity can also exceed c. For the string/spring system in Section 6.2.2, we derived the dispersion relation, $\omega^2 = v^2k^2 + \omega_\mathrm{s}^2$, where $v^2 \equiv T/\mu$ and $\omega_\mathrm{s}^2 \equiv \sigma/\mu$. (We're using v here, to save c for the speed of light.) The phase velocity, $v_\mathrm{p} = \omega/k$, is the slope of the line from the origin to a point on the $\omega(k)$ curve. By making k as small as we want, we can make the slope be as large as we want, as we noted in Fig. 13. Is this bad? No. Again, we need to make a *change* in the wave if we want to convey information. And any signal can travel only as fast as the leading edge of the change.

It is quite easy to create a system whose phase velocity v_p is arbitrarily large. Just put a bunch of people in a long line, and have them stand up and sit down at *prearranged* times. If the person a zillion meters away stands up 1 second after the person at the front of the line stands up, then the phase velocity is $v_\mathrm{p} = 1\,\text{zillion m/s}$. But no information is contained in this "wave" because the actions of the people were already decided.

This is sort of like scissors. If you have a huge pair of scissors held at a very small angle, and if you close them, then it seems like the intersection point can travel faster than c. You might argue that this doesn't cause a conflict with relativity, because there is no actual object in this system that is traveling faster than c (the intersection point isn't an actual object). However, although it is correct that there isn't a conflict, this reasoning isn't valid. Information need not be carried by an actual object.

The correct reasoning is that the intersection point will travel faster than c only if you *prearrange* for the blades to move at a given instant very far away. If you were to simply apply forces at the handles, then the parts of the blades very far away wouldn't know right away that they should start moving. So the blades would bend, even if they were made out of the most rigid material possible.

Said in another way, when we guess a solution of the form $e^{i(\omega t - kx)}$ in our various wave equations, it is assumed that this is the solution for *all* space and time. These waves always were there, and they always will be there, so they don't convey any information by themselves. We have to make a *change* in them to send a signal. And the leading edge of the change can travel no faster than c.

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

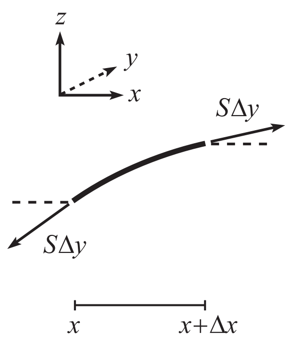


Figure 1

[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,

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

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.

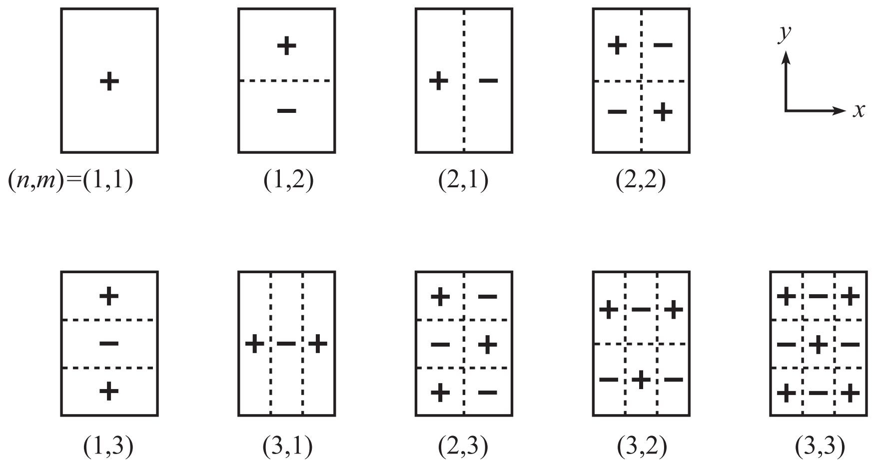


Figure 2

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.

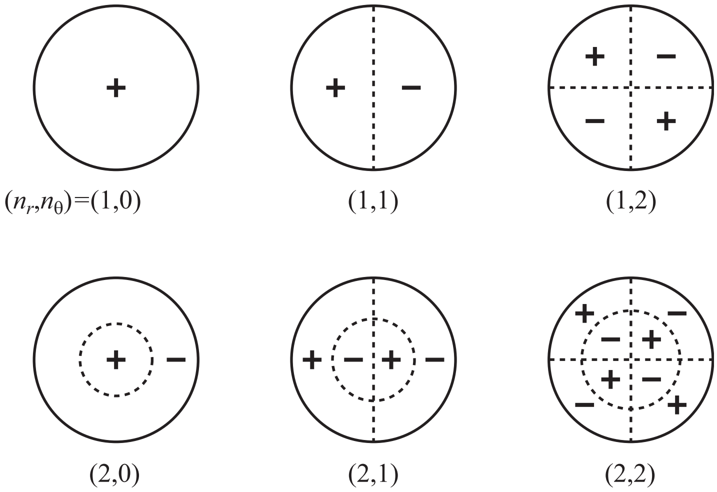


Figure 3

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.

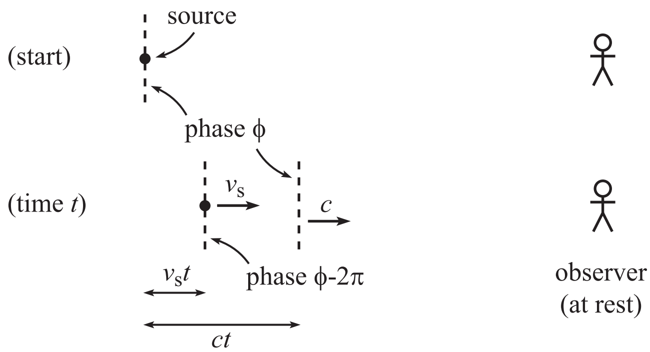


Figure 4

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)

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

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 $t = 1/f$.

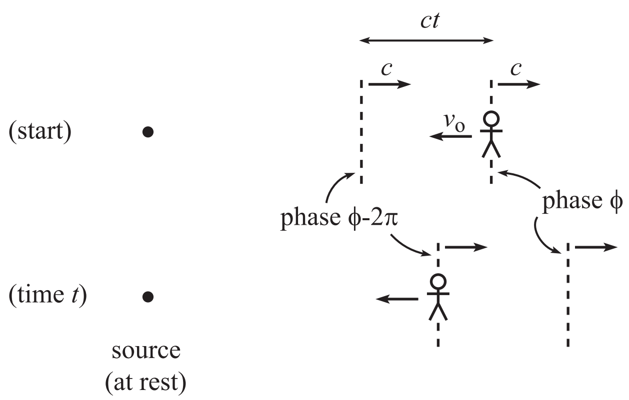


Figure 5

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_{s} and v_{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

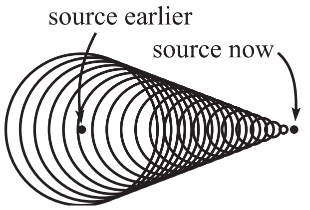


Figure 6

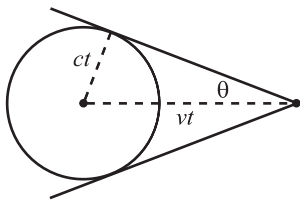


Figure 7

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.

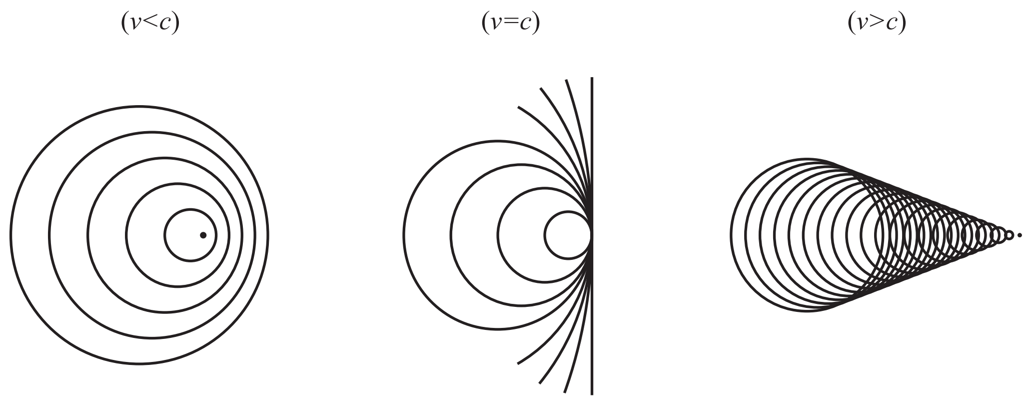


Figure 8

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).

Chapter 8

Electromagnetic waves

David Morin, morin@physics.harvard.edu

The waves we've dealt with so far in this book have been fairly easy to visualize. Waves involving springs/masses, strings, and air molecules are things we can apply our intuition to. But we'll now switch gears and talk about electromagnetic waves. These are harder to get a handle on, for a number of reasons. First, the things that are oscillating are electric and magnetic fields, which are much harder to see (which is an ironic statement, considering that we see with light, which is an electromagnetic wave). Second, the fields can have components in various directions, and there can be relative phases between these components (this will be important when we discuss polarization). And third, unlike all the other waves we've dealt with, electromagnetic waves don't need a medium to propagate in. They work just fine in vacuum. In the late 1800's, it was generally assumed that electromagnetic waves required a medium, and this hypothesized medium was called the "ether." However, no one was ever able to observe the ether. And for good reason, because it doesn't exist.

This chapter is a bit long. The outline is as follows. In Section 8.1 we talk about waves in an extended LC circuit, which is basically what a coaxial cable is. We find that the system supports waves, and that these waves travel at the speed of light. This section serves as motivation for the fact that light is an electromagnetic wave. In Section 8.2 we show how the wave equation for electromagnetic waves follows from Maxwell's equations. Maxwell's equations govern all of electricity and magnetism, so it is no surprise that they yield the wave equation. In Section 8.3 we see how Maxwell's equations constrain the form of the waves. There is more information contained in Maxwell's equations than there is in the wave equation. In Section 8.4 we talk about the energy contained in an electromagnetic wave, and in particular the energy flow which is described by the *Poynting vetor*. In Section 8.5 we talk about the momentum of an electromagnetic wave. We saw in Section 4.4 that the waves we've discussed so far carry energy but not momentum. Electromagnetic waves carry both.[1] In Section 8.6 we discuss polarization, which deals with the relative phases of the different components of the electric (and magnetic) field. In Section 8.7 we show how an electromagnetic wave can be produced by an oscillating (and hence accelerating) charge. Finally, in Section 8.8 we discuss the reflection and transmission that occurs when an electromagnetic wave encounters the boundary between two different regions, such as air

[1]Technically, all waves carry momentum, but this momentum is suppressed by a factor of v/c, where v is the speed of the wave and c is the speed of light. This follows from the relativity fact that energy is equivalent to mass. So a flow of energy implies a flow of mass, which in turn implies nonzero momentum. However, the factor of v/c causes the momentum to be negligible unless we're dealing with relativistic speeds.

and glass. We deal with both normal and non-normal angles of incidence. The latter is a bit more involved due to the effects of polarization.

8.1 Cable waves

Before getting into Maxwell's equations and the wave equation for light, let's do a warmup example and study the electromagnetic waves that propagate down a coaxial cable. This example should help convince you that light is in fact an electromagnetic wave.

To get a handle on the coaxial cable, let's first look at the idealized circuit shown in Fig. 1. All the inductors are L, and all the capacitors are C. There are no resistors in the circuit. With the charges, currents, and voltages labeled as shown, we have three facts:

1. The charge on a capacitor is $Q = CV \Longrightarrow q_n = CV_n$.
2. The voltage across an inductor is $V = L(dI/dt) \Longrightarrow V_{n-1} - V_n = L(dI_n/dt)$.
3. Conservation of charge gives $I_n - I_{n+1} = dq_n/dt$.

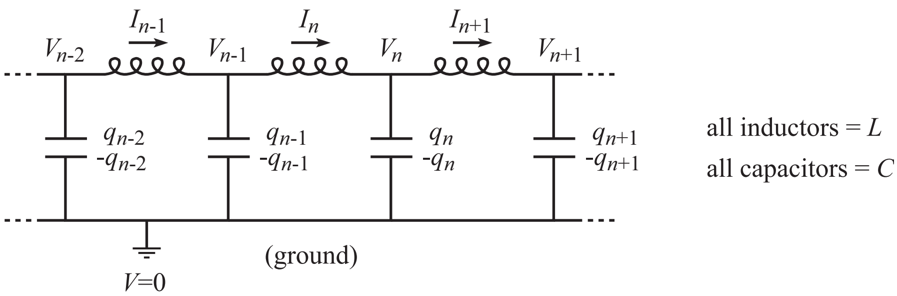


Figure 1

Our goal is to produce an equation, which will end up being a wave equation, for one of the three variables, q, I, and V (the wave equations for all of them will turn out to be the same). Let's eliminate q and I, in favor of V. We could manipulate the above equations in their present form in terms of discrete quantities, and then take the continuum limit (see Problem [to be added]). But it is much simpler to first take the continuum limit and then do the manipulation. If we let the grid size in Fig. 1 be Δx, then by using the definition of the derivative, the above three facts become

$$\begin{aligned} q &= CV, \\ -\Delta x \frac{\partial V}{\partial x} &= L\frac{\partial I}{\partial t}, \\ -\Delta x \frac{\partial I}{\partial x} &= \frac{\partial q}{\partial t}. \end{aligned} \tag{1}$$

Substituting $q = CV$ from the first equation into the third, and defining the inductance and capacitance per unit length as $L_0 \equiv L/\Delta x$ and $C_0 \equiv C/\Delta x$, the last two equations become

$$-\frac{\partial V}{\partial x} = L_0\frac{\partial I}{\partial t}, \qquad \text{and} \qquad -\frac{\partial I}{\partial x} = C_0\frac{\partial V}{\partial t}. \tag{2}$$

If we take $\partial/\partial x$ of the first of these equations and $\partial/\partial t$ of the second, and then equate the results for $\partial^2 I/\partial x\,\partial t$, we obtain

$$-\frac{1}{L_0}\frac{\partial^2 V}{\partial x^2} = -C_0\frac{\partial^2 V}{\partial t^2} \quad \Longrightarrow \quad \boxed{\frac{\partial^2 V(x,t)}{\partial t^2} = \frac{1}{L_0 C_0}\frac{\partial^2 V(x,t)}{\partial x^2}} \tag{3}$$

This is the desired wave equation, and it happens to be dispersionless. We can quickly read off the speed of the waves, which is

$$v = \frac{1}{\sqrt{L_0 C_0}} \,. \tag{4}$$

If we were to subdivide the circuit in Fig. 1 into smaller and smaller cells, L and C would depend on Δx (and would go to zero as $\Delta x \to 0$), so it makes sense to work with the quantities L_0 and C_0. This is especially true in the case of the actual cable we'll discuss below, for which the choice of Δx is arbitrary. L_0 and C_0 are the meaningful quantities that are determined by the nature of the cable.

Note that since the first fact above says that $q \propto V$, the exact same wave equation holds for q. Furthermore, if we had eliminated V instead of I in Eq. (2) by taking $\partial/\partial t$ of the first equation and $\partial/\partial x$ of the second, we would have obtained the same wave equation for I, too. So V, q, and I all satisfy the same wave equation.

Figure 2

Let's now look at an actual coaxial cable. Consider a conducting wire inside a conducting cylinder, with vacuum in the region between them, as shown in Fig. 2. Assume that the wire is somehow constrained to be in the middle of the cylinder. (In reality, the inbetween region is filled with an insulator which keeps the wire in place, but let's keep things simple here with a vacuum.) The cable has an inductance L_0 per unit length, in the same way that two parallel wires have a mutual inductance per unit length. (The cylinder can be considered to be made up of a large number wires parallel to its axis.) It also has a capacitance C_0 per unit length, because a charge difference between the wire and the cylinder will create a voltage difference.

It can be shown that (see Problem [to be added], although it's perfectly fine to just accept this)

$$L_0 = \frac{\mu_0}{2\pi}\ln(r_2/r_1) \qquad \text{and} \qquad C_0 = \frac{2\pi\epsilon_0}{\ln(r_2/r_1)} \,, \tag{5}$$

where r_2 is the radius of the cylinder, and r_1 is the radius of the wire. The two physical constants in these equations are the *permeability of free space*, μ_0, and the *permittivity of free space*, ϵ_0. Their values are (μ takes on this value by definition)

$$\mu_0 = 4\pi \cdot 10^{-7}\,\text{H/m} \,, \qquad \text{and} \qquad \epsilon_0 \approx 8.85 \cdot 10^{-12}\,\text{F/m} \,. \tag{6}$$

H and F are the Henry and Farad units of inductance and capacitance. Using Eq. (5), the wave speed in Eq. (4) equals

$$v = \frac{1}{\sqrt{L_0 C_0}} = \frac{1}{\sqrt{\mu_0\epsilon_0}} \approx \frac{1}{\sqrt{(4\pi \cdot 10^{-7}\,\text{H/m})(8.85 \cdot 10^{-12}\,\text{F/m})}} \approx 3 \cdot 10^8\,\text{m/s}. \tag{7}$$

This is the speed of light! We see that the voltage (and charge, and current) wave that travels down the cable travels at the speed of light. And because there are electric and magnetic fields in the cable (due to the capacitance and inductance), these fields also undergo wave motion. Since the waves of these fields travel with the same speed as the original voltage wave, it is a good bet that electromagnetic waves have something to do with light. The reasoning here is that there probably aren't too many things in the world that travel with the speed of light. So if we find something that travels with this speed, then it's probably light (loosely speaking, at least; it need not be in the visible range). Let's now be rigorous and show from scratch that all electromagnetic waves travel at the speed of light (in vacuum).

8.2 The wave equation

By "from scratch" we mean by starting with Maxwell's equations. We have to start somewhere, and Maxwell's equations govern all of (classical) electricity and magnetism. There are four of these equations, although when Maxwell first wrote them down, there were 22 of them. But they were gradually rewritten in a more compact form over the years. Maxwell's equations in vacuum in SI units are (in perhaps overly-general form):

Differential form	Integrated form
$\nabla \cdot \mathbf{E} = \frac{\rho_E}{\epsilon_0}$	$\int \mathbf{E} \cdot d\mathbf{A} = \frac{Q_E}{\epsilon_0}$
$\nabla \cdot \mathbf{B} = \rho_B$	$\int \mathbf{B} \cdot d\mathbf{A} = Q_B$
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} + \mathbf{J}_B$	$\int \mathbf{E} \cdot d\mathbf{l} = -\frac{d\Phi_B}{dt} + I_B$
$\nabla \times \mathbf{B} = \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J}_E$	$\int \mathbf{B} \cdot d\mathbf{l} = \mu_0\epsilon_0 \frac{d\Phi_E}{dt} + \mu_0 I_E$

Table 1

If you erase the μ_0's and ϵ_0's here (which arise from the arbitrary definitions of the various units), then these equations are symmetric in **E** and **B**, except for a couple minus signs. The ρ's are the electric and (hypothetical) magnetic charge densities, and the **J**'s are the current densities. The Q's are the charges enclosed by the surfaces that define the $d\mathbf{A}$ integrals, the Φ's are the field fluxes through the loops that define the $d\mathbf{l}$ integrals, and the I's are the currents through these loops.

No one has ever found an isolated magnetic charge (a magnetic monopole), and there are various theoretical considerations that suggest (but do not yet prove) that magnetic monopoles can't exist, at least in our universe. So we'll set ρ_B, $\mathbf{J}_B$, and I_B equal to zero from here on. This will make Maxwell's equations appear non-symmetrical, but we'll soon be setting the analogous electric quantities equal to zero too, since we'll be dealing with vacuum. So in the end, the equations for our purposes will be symmetric (except for the μ_0, the ϵ_0, and a minus sign). Maxwell's equations with no magnetic charges (or currents) are:

Differential form	Integrated form	Known as
$\nabla \cdot \mathbf{E} = \frac{\rho_E}{\epsilon_0}$	$\int \mathbf{E} \cdot d\mathbf{A} = \frac{Q_E}{\epsilon_0}$	Gauss' Law
$\nabla \cdot \mathbf{B} = 0$	$\int \mathbf{B} \cdot d\mathbf{A} = 0$	No magnetic monopoles
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	$\int \mathbf{E} \cdot d\mathbf{l} = -\frac{d\Phi_B}{dt}$	Faraday's Law
$\nabla \times \mathbf{B} = \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J}_E$	$\int \mathbf{B} \cdot d\mathbf{l} = \mu_0\epsilon_0 \frac{d\Phi_E}{dt} + \mu_0 I_E$	Ampere's Law

Table 2

The last of these, Ampere's Law, includes the so-called "displacement current," $d\Phi_E/dt$.

Our goal is to derive the wave equation for the **E** and **B** fields in vacuum. Since there are no charges of any kind in vacuum, we'll set ρ_E and $\mathbf{J}_E = 0$ from here on. And we'll only need the differential form of the equations, which are now

$$\nabla \cdot \mathbf{E} = 0, \quad (8)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (9)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (10)$$

$$\nabla \times \mathbf{B} = \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \quad (11)$$

These equations are symmetric in **E** and **B** except for the factor of $\mu_0\epsilon_0$ and a minus sign. Let's eliminate **B** in favor of **E** and see what we get. If we take the curl of Eq. (10) and then use Eq. (11) to get ride of **B**, we obtain

$$\begin{aligned}\nabla \times (\nabla \times \mathbf{E}) &= -\nabla \times \frac{\partial \mathbf{B}}{\partial t} \\ &= -\frac{\partial(\nabla \times \mathbf{B})}{\partial t} \\ &= -\frac{\partial}{\partial t}\left(\mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t}\right) \\ &= -\mu_0\epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}\,. \end{aligned} \tag{12}$$

On the left side, we can use the handy "BAC-CAB" formula,

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}). \tag{13}$$

See Problem [to be added] for a derivation of this. This formula holds even if we have differential operators (such as ∇) instead of normal vectors, but we have to be careful to keep the ordering of the letters the same (this is evident if you go through the calculation in Problem [to be added]). Since both **A** and **B** are equal to ∇ in the present application, the ordering of **A** and **B** in the $\mathbf{B}(\mathbf{A} \cdot \mathbf{C})$ term doesn't matter. But the $\mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ must correctly be written as $(\mathbf{A} \cdot \mathbf{B})\mathbf{C}$. The lefthand side of Eq. (12) then becomes

$$\begin{aligned}\nabla \times (\nabla \times \mathbf{E}) &= \nabla(\nabla \cdot \mathbf{E}) - (\nabla \cdot \nabla)\mathbf{E} \\ &= 0 - \nabla^2 \mathbf{E}, \end{aligned} \tag{14}$$

where the zero follows from Eq. (8). Plugging this into Eq. (12) finally gives

$$-\nabla^2 \mathbf{E} = -\mu_0\epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad \Longrightarrow \quad \boxed{\frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{1}{\mu_0\epsilon_0}\nabla^2 \mathbf{E}} \qquad (\textit{wave equation}) \tag{15}$$

Note that we didn't need to use the second of Maxwell's equations to derive this.

In the above derivation, we could have instead eliminated **E** in favor of **B**. The same steps hold; the minus signs end up canceling again, as you should check, and the first equation is now not needed. So we end up with exactly the same wave equation for **B**:

$$\boxed{\frac{\partial^2 \mathbf{B}}{\partial t^2} = \frac{1}{\mu_0\epsilon_0}\nabla^2 \mathbf{B}} \qquad (\textit{wave equation}) \tag{16}$$

The speed of the waves (both **E** and **B**) is given by the square root of the coefficient on the righthand side of the wave equation. (This isn't completely obvious, since we're now working in three dimensions instead of one, but we'll justify this in Section 8.3.1 below.) The speed is therefore

$$c = \frac{1}{\sqrt{\mu_0\epsilon_0}} \approx 3 \cdot 10^8\,\text{m/s}. \tag{17}$$

This agrees with the result in Eq. (7). But we now see that we don't need a cable to support the propagation of electromagnetic waves. They can propagate just fine in vacuum! This is a fundamentally new feature, because every wave we've studied so far in this book (longitudinal spring/mass waves, transverse waves on a string, longitudinal sound waves, etc.), needs a medium to propagate in/on. But not so with electromagnetic waves.

Eq. (15), and likewise Eq. (16), is a vector equation. So it is actually shorthand for three separate equations for each of the components:

$$\frac{\partial^2 E_x}{\partial t^2} = \frac{1}{\mu_0\epsilon_0}\left(\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2}\right), \tag{18}$$

and likewise for E_y and E_z. Each component undergoes wave motion. As far as the wave equation in Eq. (15) is concerned, the waves for the three components are completely independent. Their amplitudes, frequencies, and phases need not have anything to do with each other. *However*, there is more information contained in Maxwell's equations than in the wave equation. The latter follows from the former, but not the other way around. There is no reason why one equation that follows from four equations (or actually just three of them) should contain as much information as the original four. In fact, it is highly unlikely. And as we will see in Section 8.3, Maxwell's equations do indeed further constrain the form of the waves. In other words, although the wave equation in Eq. (15) gives us information about the electric-field wave, it doesn't give us *all* the information.

Index of refraction

In a dielectric (equivalently, an insulator), the vacuum quantities μ_0 and ϵ_0 in Maxwell's equations are replaced by new values, μ and ϵ. (We'll give some justification of this below, but see Sections 10.11 and 11.10 in Purcell's book for the full treatment.) Our derivation of the wave equation for electromagnetic waves in a dielectric proceeds in exactly the same way as for the vacuum case above, except with $\mu_0 \to \mu$ and $\epsilon_0 \to \epsilon$. We therefore end up with a wave velocity equal to

$$v = \frac{1}{\sqrt{\mu\epsilon}}\,. \tag{19}$$

The *index of refraction*, n, of a dielectric is defined by $v \equiv c/n$, where c is the speed of light in vacuum. We therefore have

$$v = \frac{c}{n} \quad \Longrightarrow \quad n = \frac{c}{v} = \sqrt{\frac{\mu\epsilon}{\mu_0\epsilon_0}}\,. \tag{20}$$

Since it happens to be the case that that $\mu \approx \mu_0$ for most dielectrics, we have the approximate result that

$$n \approx \sqrt{\frac{\epsilon}{\epsilon_0}} \qquad (\text{if} \quad \mu \approx \mu_0). \tag{21}$$

And since we must always have $v \le c$, this implies $n \ge 1 \Longrightarrow \epsilon \ge \epsilon_0$.

Strictly speaking, Maxwell's equations with μ_0 and ϵ_0 work in any medium. But the point is that if we don't have a vacuum, then induced charges and currents may arise. In particular there are two types of charges. There are so-called *free* charges, which are additional charges that we can plop down in a material. This is normally what we think of when we think of charge. (The term "free" is probably not the best term, because the charges need not be free to move. We can bolt them down if we wish.) But additionally, there are *bound* charges. These are the effective charges that get produced when polar molecules align themselves in certain ways to "shield" the bound charges.

For example, if we place a positive free charge q_{free} in a material, then the nearby polar molecules will align themselves so that their negative ends form a negative layer around the free charge. The net charge inside a Gaussian surface around the charge is therefore less than q. Call it q_{net}. Maxwell's first equation is then $\nabla \cdot \mathbf{E} = \rho_{\text{net}}/\epsilon_0$. However, it is generally much easier to deal with ρ_{free} than ρ_{net}, so let's define ϵ by $\rho_{\text{net}}/\rho_{\text{free}} \equiv \epsilon_0/\epsilon < 1$.[2] Maxwell's first equation can then be written as

$$\nabla \cdot \mathbf{E} = \frac{\rho_{\text{free}}}{\epsilon}\,. \tag{22}$$

[2]The fact that the shielding is always proportional to q_{free} (at least in non-extreme cases) implies that there is a unique value of ϵ that works for all values of q_{free}.

The electric field in the material around the point charge is less than what it would be in vacuum, by a factor of ϵ_0/ϵ (and ϵ is always greater than or equal to ϵ_0, because it isn't possible to have "anti-shielding"). In a dielectric, the fact that ϵ is greater than ϵ_0 is consistent with the fact that the index of refraction n in Eq. (21) is always greater than 1, which in turn is consistent with the fact that v is always less than c.

A similar occurrence happens with currents. There can be *free* currents, which are the normal ones we think about. But there can also be *bound* currents, which arise from tiny current loops of electrons spinning around within their atoms. This is a little harder to visualize than the case with the charges, but let's just accept here that the fourth of Maxwell's equations becomes $\nabla \times \mathbf{B} = \mu\epsilon\partial\mathbf{E}/\partial t + \mu\mathbf{J}_{\text{free}}$. But as mentioned above, μ is generally close to μ_0 for most dielectrics, so this distinction usually isn't so important.

To sum up, we can ignore all the details about what's going on at the atomic level by pretending that we have a vacuum with modified μ and ϵ values. Although there certainly exist *bound* charges and currents in the material, we can sweep them under the rug and consider only the *free* charges and currents, by using the modified μ and ϵ values.

The above modified expressions for Maxwell's equations are correct if we're dealing with a single medium. But if we have two or more mediums, the correct way to write the equations is to multiply the first equation by ϵ and divide the fourth equation by μ (see Problem [to be added] for an explanation of this). The collection of all four Maxwell's equations is then

$$\begin{aligned} \nabla \cdot \mathbf{D} &= \rho_{\text{free}}, \\ \nabla \cdot \mathbf{B} &= 0, \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \times \mathbf{H} &= \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}_{\text{free}}, \end{aligned} \tag{23}$$

where $\mathbf{D} \equiv \epsilon\mathbf{E}$ and $\mathbf{H} \equiv \mathbf{B}/\mu$. $\mathbf{D}$ is called the *electric displacement* vector, and $\mathbf{H}$ goes by various names, including simply the "magnetic field." But you can avoid confusing it with $\mathbf{B}$ if you use the letter $\mathbf{H}$ and not the name "magnetic field."

8.3 The form of the waves

8.3.1 The wavevector k

What is the dispersion relation associated with the wave equation in Eq. (15)? That is, what is the relation between the frequency and wavenumber? Or more precisely, what is the dispersion relation for each component of $\mathbf{E}$, for example the E_x that satisfies Eq. (18)? All of the components are in general functions of four coordinates: the three spatial coordinates x, y, z, and the time t. So by the same reasoning as in the two-coordinate case we discussed at the end of Section 4.1, we know that we can Fourier-decompose the function $E_x(x, y, z, t)$ into exponentials of the form,

$$Ae^{i(k_x x + k_y y + k_z z - \omega t)} \equiv Ae^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}, \qquad \text{where} \quad \mathbf{k} \equiv (k_x, k_y, k_z). \tag{24}$$

Likewise for E_y and E_z. And likewise for the three components of $\mathbf{B}$. These are traveling waves, although we can form combinations of them to produce standing waves.

$\mathbf{k}$ is known as the *wavevector*. As we'll see below, the magnitude $k \equiv |\mathbf{k}|$ plays exactly the same role that k played in the 1-D case. That is, k is the wavenumber. It equals 2π times the number of wavelengths that fit into a unit length. So $k = 2\pi/\lambda$. We'll also see below that the direction of $\mathbf{k}$ is the direction of the propagation of the wave. In the 1-D case, the wave had no choice but to propagate in the $\pm\mathbf{x}$ direction. But now it can propagate in any direction in 3-D space.

Plugging the exponential solution in Eq. (24) into Eq. (18) gives

$$-\omega^2 = \frac{1}{\mu_0\epsilon_0}(-k_x^2 - k_y^2 - k_z^2) \quad\Longrightarrow\quad \omega^2 = \frac{|\mathbf{k}|^2}{\mu_0\epsilon_0} \quad\Longrightarrow\quad \boxed{\omega = c|\mathbf{k}|} \qquad (25)$$

where $c = 1/\sqrt{\mu_0\epsilon_0}$, and where we are using the convention that ω is positive. Eq. (25) is the desired dispersion relation. It is a trivial relation, in the sense that electromagnetic waves in vacuum are dispersionless.

When we go through the same procedure for the other components of $\mathbf{E}$ and $\mathbf{B}$, the "A" coefficient in Eq. (24) can be different for the $2 \cdot 3 = 6$ different components of the fields. And technically $\mathbf{k}$ and ω can be different for the six components too (as long as they satisfy the same dispersion relation). However, although we would have solutions to the six different waves equations, we wouldn't have solutions to Maxwell's equations. This is one of the cases where the extra information contained in Maxwell's equations is important. You can verify (see Problem [to be added]) that if you want Maxwell's equations to hold for *all* $\mathbf{r}$ and t, then $\mathbf{k}$ and ω must be the same for all six components of $\mathbf{E}$ and $\mathbf{B}$. If we then collect the various "A" components into the two vectors $\mathbf{E}_0$ and $\mathbf{B}_0$ (which are constants, independent of $\mathbf{r}$ and t), we can write the six components of $\mathbf{E}$ and $\mathbf{B}$ in vector form as

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \qquad \text{and} \qquad \mathbf{B} = \mathbf{B}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \qquad (26)$$

where the $\mathbf{k}$ vector and the ω frequency are the same in both fields. The vectors $\mathbf{E}_0$ and $\mathbf{B}_0$ can be complex. If they do have an imaginary part, it will produce a phase in the cosine function when we take the real part of the above exponentials. This will be important when we discuss polarization.

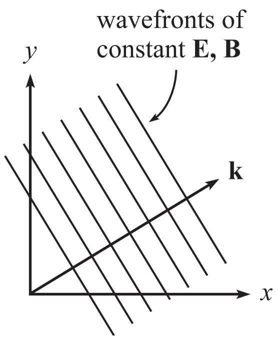


Figure 3

From Eq. (26), we see that $\mathbf{E}$ (and likewise $\mathbf{B}$) depends on $\mathbf{r}$ through the dot product $\mathbf{k}\cdot\mathbf{r}$. So $\mathbf{E}$ has the same value everywhere on the surface defined by $\mathbf{k}\cdot\mathbf{r} = C$, where C is some constant. This surface is a plane that is perpendicular to $\mathbf{k}$. This follows from the fact that if $\mathbf{r}_1$ and $\mathbf{r}_2$ are two points on the surface, then $\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2) = C - C = 0$. Therefore, the vector $\mathbf{r}_1 - \mathbf{r}_2$ is perpendicular to $\mathbf{k}$. Since this holds for any $\mathbf{r}_1$ and $\mathbf{r}_2$ on the surface, the surface must be a plane perpendicular to $\mathbf{k}$. If we suppress the z dependence of $\mathbf{E}$ and $\mathbf{B}$ for the sake of drawing a picture on a page, then for a given wavevector $\mathbf{k}$, Fig. 3 shows some "wavefronts" with common phases $\mathbf{k}\cdot\mathbf{r} - \omega t$, and hence common values of $\mathbf{E}$ and $\mathbf{B}$. The planes perpendicular to $\mathbf{k}$ in the 3-D case become lines perpendicular to $\mathbf{k}$ in the 2-D case. Every point on a given plane is equivalent, as far as $\mathbf{E}$ and $\mathbf{B}$ are concerned.

How do these wavefronts move as time goes by? Well, they must always be perpendicular to $\mathbf{k}$, so all they can do is move in the direction of $\mathbf{k}$. How fast do they move? The dot product $\mathbf{k}\cdot\mathbf{r}$ equals $kr\cos\theta$, where θ is the angle between $\mathbf{k}$ and a given position $\mathbf{r}$, and where $k \equiv |\mathbf{k}|$ and $r = |\mathbf{r}|$. If we group the product as $k(r\cos\theta)$, we see that it equals k times the projection of $\mathbf{r}$ along $\mathbf{k}$. If we rotate our coordinate system so that a new x' axis points in the $\mathbf{k}$ direction, then the projection $r\cos\theta$ simply equals the x' value of the position. So the phase $\mathbf{k}\cdot\mathbf{r} - \omega t$ equals $kx' - \omega t$. We have therefore reduced the problem to a 1-D problem (at least as far as the phase is concerned), so we can carry over all of our 1-D results. In particular, the phase velocity (and group velocity too, since the wave equation in Eq. (15) is dispersionless) is $v = \omega/k$, which we see from Eq. (25) equals $c = 1/\sqrt{\mu_0\epsilon_0}$.

Remark: We just found that the phase velocity has magnitude

$$v = \frac{\omega}{k} \equiv \frac{\omega}{|\mathbf{k}|} = \frac{\omega}{\sqrt{k_x^2 + k_y^2 + k_z^2}}\,, \qquad (27)$$

and it points in the $\hat{\mathbf{k}}$ direction. You might wonder if the simpler expression ω/k_x has any meaning. And likewise for y and z. It does, but it isn't a terribly useful quantity. It is the velocity at which a point with constant phase moves in the x direction, with y and z held constant. This follows from the fact that if we let the constant y and z values be y_0 and z_0, then the phase equals $k_x x + k_y y_0 + k_z z_0 - \omega t = k_x x - \omega t + C$, where C is a constant. So we effectively have a 1-D problem for which the phase velocity is ω/k_x.

But note that this velocity can be made arbitrarily large, or even infinite if $k_x = 0$. Fig. 4 shows a situation where $\mathbf{k}$ points mainly in the y direction, so k_x is small. Two wavefronts are shown, and they move upward along the direction of $\mathbf{k}$. In the time during which the lower wavefront moves to the position of the higher one, a point on the x axis with a particular constant phase moves from one dot to the other. This means that it is moving very fast (much faster than the wavefronts), consistent with the fact the ω/k_x is very large if k_x is very small. In the limit where the wavefronts are horizontal ($k_x = 0$), a point of constant phase moves infinitely fast along the x axis. The quantities ω/k_x, ω/k_y, and ω/k_z therefore *cannot* be thought of as components of the phase velocity in Eq. (27). The component of a vector should be smaller than the vector itself, after all.

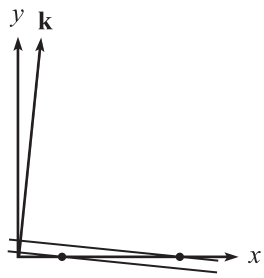
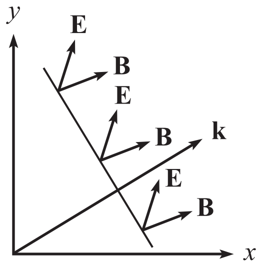


Figure 4

The vector that *does* correctly break up into components is the wavevector $\mathbf{k} = (k_x, k_y, k_z)$. Its magnitude $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$ represents how much the phase of the wave increases in each unit distance along the $\hat{\mathbf{k}}$ direction. (In other words, it equals 2π times the number of wavelengths that fit into a unit distance.) And k_x represents how much the phase of the wave increases in each unit distance along the x direction. This is less than k, as it should be, in view of Fig. 4. For a given distance along the x axis, the phase advances by only a small amount, compared with along the $\mathbf{k}$ vector. The phase needs the entire distance between the two dots to increase by 2π along the x axis, whereas it needs only the distance between the wavefronts to increase by 2π along the $\mathbf{k}$ vector. ♣

8.3.2 Further constraints due to Maxwell's equations

Fig. 3 tells us only that points along a given line have common values of $\mathbf{E}$ and $\mathbf{B}$. It doesn't tell us what these values actually are, or if they are constrained in other ways. For all we know, $\mathbf{E}$ and $\mathbf{B}$ on a particular wavefront might look like the vectors shown in Fig. 5 (we have ignored any possible z components). But it turns out the these vectors aren't actually possible. Although they satisfy the wave equation, they don't satisfy Maxwell's equations. So let's now see how Maxwell's equations further constrain the form of the waves. Later on in Section 8.8, we'll see that the waves are even further constrained by any boundary conditions that might exist. We'll look at Maxwell's equations in order and see what each of them implies.



(impossible **E, B** vectors)

Figure 5

- Using the expression for $\mathbf{E}$ in Eq. (26), the first of Maxwell's equations, Eq. (8), gives

$$\begin{aligned} \nabla \cdot \mathbf{E} = 0 \quad &\Longrightarrow \quad \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0 \\ &\Longrightarrow \quad ik_x E_x + ik_y E_y + ik_z E_z = 0 \\ &\Longrightarrow \quad \boxed{\mathbf{k} \cdot \mathbf{E} = 0} \end{aligned} \tag{28}$$

  This says that $\mathbf{E}$ is always perpendicular to $\mathbf{k}$. As we see from the second line here, each partial derivative simply turns into a factor if ik_x, etc.

- The second of Maxwell's equations, Eq. (9), gives the analogous result for $\mathbf{B}$, namely,

$$\boxed{\mathbf{k} \cdot \mathbf{B} = 0} \tag{29}$$

  So $\mathbf{B}$ is also perpendicular to $\mathbf{k}$.

- Again using the expression for $\mathbf{E}$ in Eq. (26), the third of Maxwell's equations, Eq. (10), gives

$$\begin{aligned}\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad &\Longrightarrow \quad \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \\ &\Longrightarrow \quad \big(ik_x, ik_y, ik_z\big) \times \mathbf{E} = -(-i\omega)\mathbf{B} \\ &\Longrightarrow \quad \boxed{\mathbf{k} \times \mathbf{E} = \omega \mathbf{B}} \end{aligned} \tag{30}$$

Since the cross product of two vectors is perpendicular to each of them, this result says that $\mathbf{B}$ is perpendicular to $\mathbf{E}$. And we already know that $\mathbf{B}$ is perpendicular to $\mathbf{k}$, from the second of Maxwell's equations. But technically we didn't need to use that equation, because the $\mathbf{B} \perp \mathbf{k}$ result is also contained in this $\mathbf{k} \times \mathbf{E} = \omega\mathbf{B}$ result. Note that as above with the divergences, each partial derivative in the curl simply turns into a factor if ik_x, etc.

We know from the first of Maxwell's equations that $\mathbf{E}$ is perpendicular to $\mathbf{k}$, so the magnitude of $\mathbf{k} \times \mathbf{E}$ is simply $|\mathbf{k}||\mathbf{E}| \equiv kE$. The magnitude of the $\mathbf{k} \times \mathbf{E} = \omega\mathbf{B}$ relation then tells us that

$$kE = \omega B \quad \Longrightarrow \quad E = \frac{\omega}{k} B \quad \Longrightarrow \quad \boxed{E = cB} \tag{31}$$

Therefore, the magnitudes of $\mathbf{E}$ and $\mathbf{B}$ are related by a factor of the wave speed, c. Eq. (31) is very useful, but its validity is limited to a single traveling wave, because the derivation of Eq. (30) assumed a unique $\mathbf{k}$ vector. If we form the sum of two waves with different $\mathbf{k}$ vectors, then the sum doesn't satisfy Eq. (30) for any particular vector $\mathbf{k}$. There isn't a unique $\mathbf{k}$ vector associated with the wave. Likewise for Eqs. (28) and (29).

- The fourth of Maxwell's equations, Eq. (11), can be written as $\nabla \times \mathbf{B} = (1/c^2)\partial\mathbf{E}/\partial t$, so the same procedure as above yields

$$\boxed{\mathbf{k} \times \mathbf{B} = -\frac{\omega}{c^2}\mathbf{E}} \quad \Longrightarrow \quad B = \frac{\omega}{kc^2} E \quad \Longrightarrow \quad \boxed{B = \frac{E}{c}} \tag{32}$$

This doesn't tell us anything new, because we already know that $\mathbf{E}$, $\mathbf{B}$, and $\mathbf{k}$ are all mutually perpendicular, and also that $E = cB$. In retrospect, the first and third (or alternatively the second and fourth) of Maxwell's equations are sufficient to derive all of the above results, which can be summarized as

$$\boxed{\mathbf{E} \perp \mathbf{k}, \qquad \mathbf{B} \perp \mathbf{k}, \qquad \mathbf{E} \perp \mathbf{B}, \qquad E = cB} \tag{33}$$

If three vectors are mutually perpendicular, there are two possibilities for how they are oriented. With the conventions of $\mathbf{E}$, $\mathbf{B}$, and $\mathbf{k}$ that we have used in Maxwell's equations and in the exponential solution in Eq. (24) (where the $\mathbf{k} \cdot \mathbf{r}$ term comes in with a plus sign), the orientation is such that $\mathbf{E}$, $\mathbf{B}$, and $\mathbf{k}$ form a "righthanded" triplet. That is, $\mathbf{E} \times \mathbf{B}$ points in the same direction as $\mathbf{k}$ (assuming, of course, that you're defining the cross product with the righthand rule!). You can show that this follows from the $\mathbf{k} \times \mathbf{E} = \omega\mathbf{B}$ relation in Eq. (30) by either simply drawing three vectors that satisfy Eq. (30), or by using the determinant definition of the cross product to show that a cyclic permutation of the vectors maintains the sign of the cross product.

A snapshot (for an arbitrary value of t) of a possible electromagnetic wave is shown in Fig. 6. We have chosen $\mathbf{k}$ to point along the z axis, and we have drawn the field only for points on the z axis. But for a given value z_0, all points of the form (x, y, z_0), which is a plane perpendicular to the z axis, have common values of **E** and **B**. **E** points in the $\pm x$ direction, and **B** points in the $\pm y$ direction. As time goes by, the whole figure simply slides along the z axis at speed c. Note that **E** and **B** reach their maximum and minimum values at the same locations. We will find below that this isn't the case for standing waves.

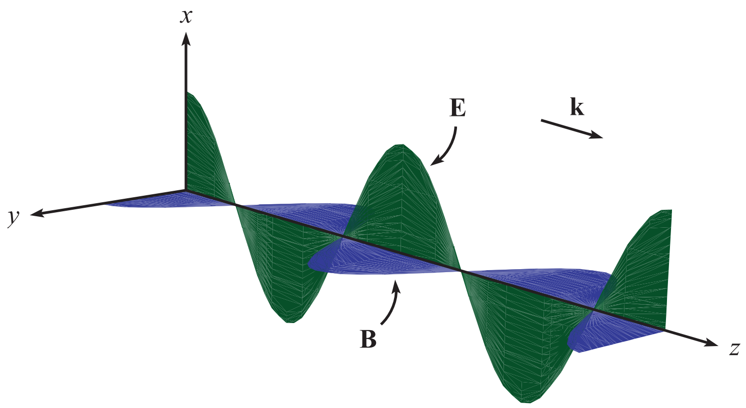


Figure 6

What are the mathematical expressions for the **E** and **B** fields in Fig. 6? We've chosen **k** to point along the z axis, so we have $\mathbf{k} = k\hat{\mathbf{z}}$, which gives $\mathbf{k}\cdot\mathbf{r} = kz$. And since **E** points in the x direction, its amplitude takes the form of $E_0 e^{i\phi}\hat{\mathbf{x}}$. (The coefficient can be complex, and we have written it as a magnitude times a phase.) This then implies that **B** points in the y direction (as drawn), because it must be perpendicular to both **E** and **k**. So its amplitude takes the form of $B_0 e^{i\phi}\hat{\mathbf{y}} = (E_0 e^{i\phi}/c)\hat{\mathbf{y}}$. This is the same phase ϕ, due to Eq. (30) and the fact that **k** is real, at least for simple traveling waves. The desired expressions for **E** and **B** are obtained by taking the real part of Eq. (26), so we arrive at

$$\mathbf{E} = \hat{\mathbf{x}}E_0\cos(kz - \omega t + \phi), \qquad \text{and} \qquad \mathbf{B} = \hat{\mathbf{y}}\frac{E_0}{c}\cos(kz - \omega t + \phi), \tag{34}$$

These two vectors are in phase with each other, consistent with Fig. 6. And **E**, **B**, and **k** form a righthanded triple of vectors, as required.

REMARKS:

1. When we talk about polarization in Section 8.6, we will see that **E** and **B** don't have to point in specific directions, as they do in Fig. 6, where **E** points only along $\hat{\mathbf{x}}$ and **B** points only along $\hat{\mathbf{y}}$. Fig. 6 happens to show the special case of "linear polarization."
2. The **E** and **B** waves don't have to be sinusoidal, of course. Because the wave equation is linear, we can build up other solutions from sinusoidal ones. And because the wave equation is dispersionless, we know (as we saw at the end of Section 2.4) that any function of the form $f(z - vt)$, or equivalently $f(kz - \omega t)$, satisfies the wave equation. But the restrictions placed by Maxwell's equations still hold. In particular, the **E** field determines the **B** field.
3. A static solution, where **E** and **B** are constant, can technically be thought of as a sinusoidal solution in the limit where $\omega = k = 0$. In vacuum, we can always add on a constant field to **E** or **B**, and it won't affect Maxwell's equations (and therefore the wave equation either), because all of the terms in Maxwell's equations in vacuum involve derivatives (either space or time). But we'll ignore any such fields, because they're boring for the purposes we'll be concerned with. ♣

8.3.3 Standing waves

Let's combine two waves (with equal amplitudes) traveling in opposite directions, to form a standing wave. If we add $\mathbf{E}_1 = \hat{\mathbf{x}}E_0\cos(kz - \omega t)$ and $\mathbf{E}_2 = \hat{\mathbf{x}}E_0\cos(-kz - \omega t)$, we obtain

$$\mathbf{E} = \hat{\mathbf{x}}(2E_0)\cos kz\cos\omega t. \tag{35}$$

This is indeed a standing wave, because all z values have the same phase with respect to time.

There are various ways to find the associated $\mathbf{B}$ wave. Actually, there are (at least) two right ways and one wrong way. The wrong way is to use the result in Eq. (30) to say that $\omega\mathbf{B} = \mathbf{k}\times\mathbf{E}$. This would yield the result that $\mathbf{B}$ is proportional to $\cos kz\cos\omega t$, which we will find below is incorrect. The error (as we mentioned above after Eq. (31)) is that there isn't a unique $\mathbf{k}$ vector associated with the wave in Eq. (35), because it is generated by two waves with opposite $\mathbf{k}$ vectors. If we insisted on using Eq. (30), we'd be hard pressed to decide if we wanted to use $k\hat{\mathbf{z}}$ or $-k\hat{\mathbf{z}}$ as the $\mathbf{k}$ vector.

A valid method for finding $\mathbf{B}$ is the following. We can find the traveling $\mathbf{B}_1$ and $\mathbf{B}_2$ waves associated with each of the traveling $\mathbf{E}_1$ and $\mathbf{E}_2$ waves, and then add them. You can quickly show (using $\mathbf{B} = (1/\omega)\mathbf{k}\times\mathbf{E}$ for each traveling wave separately) that $\mathbf{B}_1 = \hat{\mathbf{y}}(E_0/c)\cos(kz - \omega t)$ and $\mathbf{B}_2 = -\hat{\mathbf{y}}(E_0/c)\cos(-kz - \omega t)$. The sum of these waves give the desired associated $\mathbf{B}$ field,

$$\mathbf{B} = \hat{\mathbf{y}}(2E_0/c)\sin kz\sin\omega t. \tag{36}$$

Another method is to use the third of Maxwell's equations, Eq. (10), which says that $\nabla\times\mathbf{E} = -\partial\mathbf{B}/\partial t$. Maxwell's equations hold for *any* $\mathbf{E}$ and $\mathbf{B}$ fields. We don't have to worry about the uniqueness of $\mathbf{k}$ here. Using the $\mathbf{E}$ in Eq. (35), the cross product $\nabla\times\mathbf{E}$ can be calculated with the determinant:

$$\begin{aligned}\nabla\times\mathbf{E} = \begin{vmatrix}\hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}}\\ \partial/\partial x & \partial/\partial y & \partial/\partial z\\ E_x & 0 & 0\end{vmatrix} &= \hat{\mathbf{y}}\frac{\partial E_x}{\partial z} - \hat{\mathbf{z}}\frac{\partial E_x}{\partial y}\\ &= -\hat{\mathbf{y}}(2E_0)k\sin kz\cos\omega t - 0.\end{aligned} \tag{37}$$

Eq. (30) tells us that this must equal $-\partial\mathbf{B}/\partial t$, so we conclude that

$$\mathbf{B} = \hat{\mathbf{y}}(2E_0)(k/\omega)\sin kz\sin\omega t = \hat{\mathbf{y}}(2E_0/c)\sin kz\sin\omega t. \tag{38}$$

in agreement with Eq. (36). We have ignored any possible additive constant in $\mathbf{B}$.

Having derived the associated $\mathbf{B}$ field in two different ways, we can look at what we've found. $\mathbf{E}$ and $\mathbf{B}$ are still perpendicular to each other, which makes since, because $\mathbf{E}$ is the superposition of two vectors that point in the $\pm\hat{\mathbf{x}}$ direction, and $\mathbf{E}$ is the superposition of two vectors that point in the $\pm\hat{\mathbf{y}}$ direction. But there is a major difference between standing waves and traveling waves. In traveling waves, $\mathbf{E}$ and $\mathbf{B}$ run along in step with each other, as shown above in Fig. 6. They reach their maximum and minimum values at the same times and positions. However, in standing waves $\mathbf{E}$ is maximum *when* $\mathbf{B}$ is zero, and also *where* $\mathbf{B}$ is zero (and vice versa). $\mathbf{E}$ and $\mathbf{B}$ are 90° out of phase with each other in both time and space. That is, the $\mathbf{B}$ in Eq. (36) can be written as

$$\mathbf{B} = \hat{\mathbf{y}}\frac{2E_0}{c}\cos\left(kz - \frac{\pi}{2}\right)\cos\left(\omega t - \frac{\pi}{2}\right), \tag{39}$$

which you can compare with the $\mathbf{E}$ in Eq. (35). A few snapshots of the $\mathbf{E}$ and $\mathbf{B}$ waves are shown in Fig. 7.

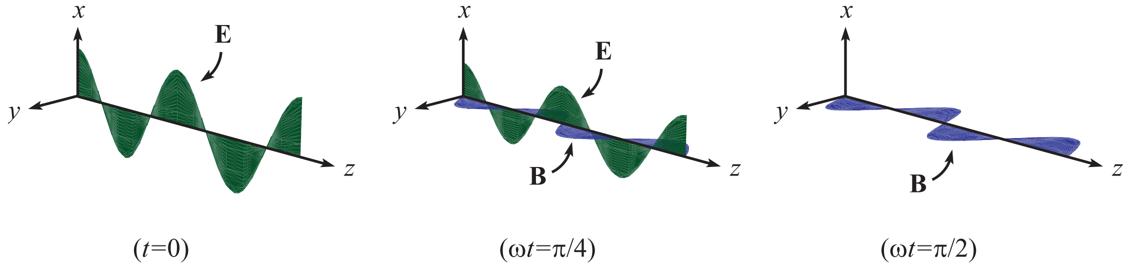


Figure 7

8.4 Energy

8.4.1 The Poynting vector

The energy density of an electromagnetic field is

$$\mathcal{E} = \frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}B^2, \tag{40}$$

where $E \equiv |\mathbf{E}|$ and $B \equiv |\mathbf{B}|$ are the magnitudes of the fields at a given location in space and time. We have suppressed the (x, y, z, t) arguments of $\mathcal{E}$, E, and B. This energy density can be derived in various ways (see Problem [to be added]), but we'll just accept it here. The goal of this section is to calculate the rate of change of $\mathcal{E}$, and to then write it in a form that allows us to determine the energy flux (the flow of energy across a given surface). We will find that the energy flux is given by the so-called *Poynting vector*.

If we write E^2 and B^2 as $\mathbf{E}\cdot\mathbf{E}$ and $\mathbf{B}\cdot\mathbf{B}$, then the rate of change of $\mathcal{E}$ becomes

$$\frac{\partial\mathcal{E}}{\partial t} = \epsilon_0\mathbf{E}\cdot\frac{\partial\mathbf{E}}{\partial t} + \frac{1}{\mu_0}\mathbf{B}\cdot\frac{\partial\mathbf{B}}{\partial t}\,. \tag{41}$$

(The product rule works here for the dot product of vectors for the same reason it works for a regular product. You can verify this by explicitly writing out the components.) The third and fourth Maxwell's equations turn this into

$$\begin{aligned}\frac{\partial\mathcal{E}}{\partial t} &= \epsilon_0\mathbf{E}\cdot\left(\frac{1}{\mu_0\epsilon_0}\nabla\times\mathbf{B}\right) + \frac{1}{\mu_0}\mathbf{B}\cdot\big(-\nabla\times\mathbf{E}\big)\\ &= \frac{1}{\mu_0}\Big(\mathbf{E}\cdot(\nabla\times\mathbf{B}) - \mathbf{B}\cdot(\nabla\times\mathbf{E})\Big).\end{aligned} \tag{42}$$

The righthand side of this expression conveniently has the same form as the righthand side of the vector identity (see Problem [to be added] for the derivation),

$$\nabla\cdot(\mathbf{C}\times\mathbf{D}) = \mathbf{D}\cdot(\nabla\times\mathbf{C}) - \mathbf{C}\cdot(\nabla\times\mathbf{D}). \tag{43}$$

So we now have

$$\frac{\partial\mathcal{E}}{\partial t} = \frac{1}{\mu_0}\nabla\cdot(\mathbf{B}\times\mathbf{E}). \tag{44}$$

Now consider a given volume V in space. Integrating Eq. (44) over this volume V yields

$$\int_V \frac{\partial\mathcal{E}}{\partial t} = \frac{1}{\mu_0}\int_V \nabla\cdot(\mathbf{B}\times\mathbf{E}) \quad\Longrightarrow\quad \frac{\partial W_V}{\partial t} = \frac{1}{\mu_0}\int_A (\mathbf{B}\times\mathbf{E})\cdot d\mathbf{A}, \tag{45}$$

where W_V is the energy contained in the volume V (we've run out of forms of the letter E), and where we have used the divergence theorem to rewrite the volume integral as a surface integral over the area enclosing the volume. $d\mathbf{A}$ is defined to be the vector perpendicular to the surface (with the positive direction defined to be outward), with a magnitude equal to the area of a little patch.

Let's now make a slight change in notation. $d\mathbf{A}$ is defined to be an outward-pointing vector, but let's define $d\mathbf{A}_{\text{in}}$ to be the inward-pointing vector, $d\mathbf{A}_{\text{in}} \equiv -d\mathbf{A}$. Eq. (45) can then be written as (switching the order of $\mathbf{E}$ and $\mathbf{B}$)

$$\frac{\partial W_V}{\partial t} = \frac{1}{\mu_0}\int_A (\mathbf{E}\times\mathbf{B})\cdot d\mathbf{A}_{\text{in}}. \tag{46}$$

We can therefore interpret the vector,

$$\boxed{\mathbf{S} \equiv \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B}} \qquad (\textit{energy flux}:\ \text{energy}/(\text{area}\cdot\text{time})) \tag{47}$$

as giving the flux of energy *into* a region. This vector $\mathbf{S}$ is known as the *Poynting vector*. And since $\mathbf{E}\times\mathbf{B} \propto \mathbf{k}$, the Poynting vector points in the same direction as the velocity of the wave. Integrating $\mathbf{S}$ over any surface (or rather, just the component perpendicular to the surface, due to the dot product with $d\mathbf{A}_{\text{in}}$) gives the energy flow across the surface. This result holds for any kind of wave – traveling, standing, or whatever. Comparing the units on both sides of Eq. (46), we see that the Poynting vector has units of energy per area per time. So if we multiply it (or its perpendicular component) by an area, we get the energy per time crossing the area.

The Poynting vector falls into a wonderful class of phonetically accurate theorems/results. Others are the Low energy theorem (named after S.Y. Low) dealing with low-energy photons, and the Schwarzschild radius of a black hole (kind of like a shield).

8.4.2 Traveling waves

Let's look at the energy density $\mathcal{E}$ and the Poynting vector $\mathbf{S}$ for a traveling wave. A traveling wave has $B = E/c$, so the energy density is (using $c^2 = 1/\mu_0\epsilon_0$ in the last step)

$$\mathcal{E} = \frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}B^2 = \frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}\frac{E^2}{c^2} \quad\Longrightarrow\quad \boxed{\mathcal{E} = \epsilon_0 E^2} \tag{48}$$

We have suppressed the (x, y, z, t) arguments of $\mathcal{E}$ and E. Note that this result holds only for traveling waves. A standing wave, for example, doesn't have $B = E/c$ anywhere, so $\mathcal{E}$ doesn't take this form. We'll discuss standing waves below.

The Poynting vector for a traveling wave is

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B} = \frac{1}{\mu_0}E\left(\frac{E}{c}\right)\hat{\mathbf{k}}, \tag{49}$$

where we have used the facts that $\mathbf{E}\perp\mathbf{B}$ and that their cross product points in the direction of $\mathbf{k}$. Using $1/\mu_0 = c^2\epsilon_0$, arrive at

$$\boxed{\mathbf{S} = c\epsilon_0 E^2\hat{\mathbf{k}}} = c\mathcal{E}\hat{\mathbf{k}}. \tag{50}$$

This last equality makes sense, because the energy density $\mathcal{E}$ moves along with the wave, which moves at speed c. So the energy per unit area per unit time that crosses a surface is $c\mathcal{E}$ (which you can verify has the correct units). Eq. (50) is true at all points (x, y, z, t) individually, and not just in an average sense. We'll derive formulas for the averages below.

In the case of a sinusoidal traveling wave of the form,

$$\mathbf{E} = \hat{\mathbf{x}}E_0 \cos(kz - \omega t) \qquad \text{and} \qquad \mathbf{B} = \hat{\mathbf{y}}(E_0/c)\cos(kz - \omega t), \tag{51}$$

the above expressions for $\mathcal{E}$ and $\mathbf{S}$ yield

$$\mathcal{E} = \epsilon_0 E_0^2 \cos^2(kz - \omega t) \qquad \text{and} \qquad \mathbf{S} = c\epsilon_0 E_0^2 \cos^2(kz - \omega t)\hat{\mathbf{k}}. \tag{52}$$

Since the average value of $\cos^2(kz - \omega t)$ over one period (in either space or time) is 1/2, we see that the average values of $\mathcal{E}$ and $|\mathbf{S}|$ are

$$\mathcal{E}_{\text{avg}} = \frac{1}{2}\epsilon_0 E_0^2 \qquad \text{and} \qquad |\mathbf{S}|_{\text{avg}} = \frac{1}{2}c\epsilon_0 E_0^2. \tag{53}$$

$|\mathbf{S}|_{\text{avg}}$ is known as the *intensity* of the wave. It is the average amount of energy per unit area per unit time that passes through (or hits) a surface. For example, at the location of the earth, the radiation from the sun has an intensity of 1360 Watts/m^2. The energy comes from traveling waves with many different frequencies, and the total intensity is just the sum of the intensities of the individual waves (see Problem [to be added]).

8.4.3 Standing waves

Consider the standing wave in Eqs. (35) and (36). With $2E_0$ defined to be A, this wave becomes

$$\mathbf{E} = \hat{\mathbf{x}}A\cos kz \cos \omega t, \qquad \text{and} \qquad \mathbf{B} = \hat{\mathbf{y}}(A/c)\sin kz \sin \omega t. \tag{54}$$

The energy density in Eq. (48) for a traveling wave isn't valid here, because it assumed $B = E/c$. Using the original expression in Eq. (40), the energy density for the above standing wave is (using $1/c^2 = \mu_0\epsilon_0$)

$$\mathcal{E} = \frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}B^2 = \frac{1}{2}\epsilon_0 A^2\big(\cos^2 kz \cos^2 \omega t + \sin^2 kz \sin^2 \omega t\big). \tag{55}$$

If we take the average over a full cycle in time (a full wavelength in space would work just as well), then the $\cos^2 \omega t$ and $\sin^2 \omega t$ factors turn into 1/2's, so the time average of $\mathcal{E}$ is

$$\mathcal{E}_{\text{avg}} = \frac{1}{2}\epsilon_0 A^2\left(\frac{1}{2}\cos^2 kz + \frac{1}{2}\sin^2 kz\right) = \frac{\epsilon_0 A^2}{4}\,, \tag{56}$$

which is independent of z. It makes sense that it doesn't depend on z, because a traveling wave has a uniform average energy density, and a standing wave is just the sum of two traveling waves moving in opposite directions.

The Poynting vector for our standing wave is given by Eq. (47) as (again, the traveling-wave result in Eq. (50) isn't valid here):

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B} = \frac{A^2}{\mu_0 c}\hat{\mathbf{k}}\cos kz \sin kz \cos \omega t \sin \omega t. \tag{57}$$

At any given value of z, the time average of this is zero (because $\cos \omega t \sin \omega t = (1/2)\sin 2\omega t$), so there is no net energy flow in a standing wave. This makes sense, because a standing wave is made up of two traveling waves moving in opposite directions which therefore have opposite energy flows (on average). Similarly, for a given value of t, the spatial average is zero. Energy sloshes back and forth between points, but there is no net flow.

Due to the fact that a standing wave is made up of two traveling waves moving in opposite directions, you might think that the Poynting vector (that is, the energy flow) should be *identically* equal to zero, for all z and t. But it isn't, because each of the two Poynting vectors depends on z and t, so only at certain discrete times and places do they anti-align and exactly cancel each other. But on average they cancel.

8.5 Momentum

Electromagnetic waves carry momentum. However, all the other waves we've studied (longitudinal spring/mass and sound waves, transverse string waves, etc.) *don't* carry momentum. (However, see Footnote 1 above.) Therefore, it is certainly not obvious that electromagnetic waves carry momentum, because it is quite possible for waves to carry energy without also carrying momentum.

A quick argument that demonstrates why an electromagnetic wave (that is, light) carries momentum is the following argument from relativity. The relativistic relation between a particle's energy, momentum, and mass is $E^2 = p^2c^2 + m^2c^4$ (we'll just accept this here). For a massless particle ($m = 0$), this yields $E^2 = p^2c^2 \Longrightarrow E = pc$. Since photons (which is what light is made of) are massless, they have a momentum given by $p = E/c$. We already know that electromagnetic waves carry energy, so this relation tells us that they must also carry momentum. In other words, a given part of an electromagnetic wave with energy E also has momentum $p = E/c$.

However, although this argument is perfectly valid, it isn't very satisfying, because (a) it invokes a result from relativity, and (b) it invokes the fact that electromagnetic waves (light) can be considered to be made up of particle-like objects called photons, which is by no means obvious. But why should the *particle* nature of light be necessary to derive the fact that an electromagnetic *wave* carries momentum? It would be nice to derive the $p = E/c$ result by working only in terms of waves and using only the results that we have developed so far in this book. Or said in a different way, it would be nice to understand how would someone living in, say, 1900 (that is, pre-relativity) would demonstrate that an electromagnetic waves carries momentum. We can do this in the following way.

Consider a particle with charge q that is free to move around in some material, and let it be under the influence of a traveling electromagnetic wave. The particle will experience forces due to the $\mathbf{E}$ and $\mathbf{B}$ fields that make up the wave. There will also be damping forces from the material. And the particle will also lose energy due to the fact that it is accelerating and hence radiating (see Section 8.7). But the exact nature of the effects of the damping and radiation won't be important for this discussion.[3]

Assume that the wave is traveling in the z direction, and let the $\mathbf{E}$ field point along the x direction. The $\mathbf{B}$ field then points along the y direction, because $\mathbf{E} \times \mathbf{B} \propto \mathbf{k}$. The complete motion of the particle will in general be quite complicated, but for the present purposes it suffices to consider the x component of the particle's velocity, that is, the component that is parallel to $\mathbf{E}$.[4] Due to the oscillating electric field, the particle will (mainly) oscillate back and forth in the x direction. However, we don't know the phase. In general, part of the velocity will be in phase with $\mathbf{E}$, and part will be $\pm 90°$ out of phase. The latter will turn out not to matter for our purposes,[5] so we'll concentrate on the part of the velocity that is

[3] If the particle is floating in outer space, then there is no damping, so only the radiation will extract energy from the particle.

[4] If we assume that the velocity v of the particle satisfies $v \ll c$ (which is generally a good approximation), then the magnetic force, $q\mathbf{v} \times \mathbf{B}$ is small compared with the electric force, $q\mathbf{E}$. This is true because $B = E/c$, so the magnetic force is suppressed by a factor of v/c (or more, depending on the angle between $\mathbf{v}$ and $\mathbf{B}$) compared with the electric force. The force on the particle is therefore due mainly to the electric field.

[5] We'll be concerned with the work done by the electric field, and this part of the velocity will lead to

in phase with **E**. Let's call it $\mathbf{v}_\mathrm{E}$. We then have the pictures shown in Fig. 8.

You can quickly verify with the righthand rule that the magnetic force $q\mathbf{v}_\mathrm{E} \times \mathbf{B}$ points forward along **k** in *both* cases. $\mathbf{v}_\mathrm{E}$ and **B** switch sign in phase with each other, so the two signs cancel, and there is a net force forward. The particle therefore picks up some forward momentum, and this momentum must have come from the wave. In a small time dt, the magnitude of the momentum that the wave gives to the particle is

$$|d\mathbf{p}| = |\mathbf{F}_\mathrm{B}\, dt| = |q\mathbf{v}_\mathrm{E} \times \mathbf{B}|\, dt = qv_\mathrm{E}B\, dt = \frac{qv_\mathrm{E}E\, dt}{c}\,. \tag{58}$$

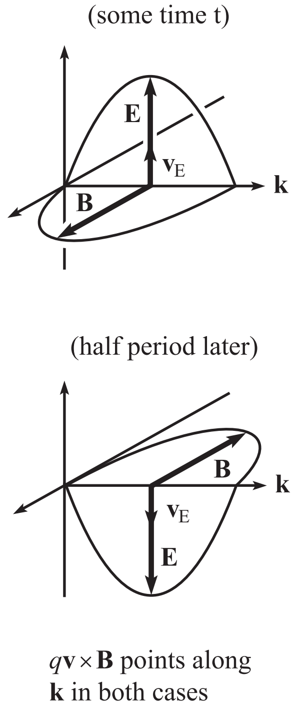


Figure 8

What is the energy that the wave gives to the particle? That is, what is the work that the wave does on the particle? (In the steady state, this work is balanced, on average, by the energy that the particle loses to damping and radiation.) Only the electric field does work on the particle. And since the electric force is qE, the amount of work done on the particle in time dt is

$$dW = \mathbf{F}_\mathrm{E} \cdot d\mathbf{x} = (qE)(v_\mathrm{E}\, dt) = qv_\mathrm{E}E\, dt. \tag{59}$$

(The part of the velocity that is $\pm 90^\circ$ out of phase with **E** will lead to zero net work, on average; see Problem [to be added].) Comparing this result with Eq. (58), we see that

$$|d\mathbf{p}| = \frac{dW}{c}\,. \tag{60}$$

In other words, the amount of momentum the particle gains from the wave equals $1/c$ times the amount of energy it gains from the wave. This holds for any extended time interval Δt, because any interval can be built up from infinitesimal times dt.

Since Eq. (60) holds whenever any electromagnetic wave encounters a particle, we conclude that the wave actually carries this amount of momentum. Even if we didn't have a particle in the setup, we could imagine putting one there, in which case it would acquire the momentum given by Eq. (60). This momentum must therefore be an intrinsic property of the wave.

Another way of writing Eq. (60) is

$$\frac{1}{A}\left|\frac{d\mathbf{p}}{dt}\right| = \frac{1}{c} \cdot \frac{1}{A}\frac{dW}{dt}\,, \tag{61}$$

where A is the cross-sectional area of the wave under consideration. The lefthand side is the force per area (in other words, the pressure) that the wave applies to a material. And from Eqs. (46) and (47), the righthand side is $|\mathbf{S}|/c$, where **S** is the Poynting vector. The pressure from an electromagnetic wave (usually called the *radiation pressure*) is therefore

$$\text{Radiation pressure} = \frac{|\mathbf{S}|}{c} = \frac{|\mathbf{E} \times \mathbf{B}|}{\mu_0 c} = \frac{E^2}{\mu_0 c^2}\,. \tag{62}$$

You can show (see Problem [to be added]) that the total force from the radiation pressure from sunlight hitting the earth is roughly $6 \cdot 10^8\,\mathrm{kg\,m/s^2}$ (treating the earth like a flat coin and ignoring reflection, but these won't affect the order of magnitude). This force is negligible compared with the attractive gravitational force, which is about $3.6 \cdot 10^{22}\,\mathrm{kg\,m/s^2}$. But for a small enough sphere, these two forces are comparable (see Problem [to be added]).

Electromagnetic waves also carry angular momentum *if* they are polarized (see Problem [to be added]).

zero net work being done.

8.6 Polarization

8.6.1 Linear polarization

Consider the traveling wave in Eq. (34) (we'll ignore the overall phase ϕ here):

$$\mathbf{E} = \hat{\mathbf{x}} E_0 \cos(kz - \omega t), \qquad \text{and} \qquad \mathbf{B} = \hat{\mathbf{y}} \frac{E_0}{c} \cos(kz - \omega t). \tag{63}$$

This wave has $\mathbf{E}$ always pointing in the x direction and $\mathbf{B}$ always pointing in the y direction. A wave like this, where the fields always point along given directions, is called a *linearly polarized* wave. The direction of the linear polarization is defined to be the axis along which the $\mathbf{E}$ field points.

But what if we want to construct a wave where the fields don't always point along given directions? For example, what if we want the $\mathbf{E}$ vector to rotate around in a circle instead of oscillating back and forth along a line?

Let's try making such a wave by adding on an $\mathbf{E}$ field (with the same magnitude) that points in the y direction. The associated $\mathbf{B}$ field then points in the negative x direction if we want the orientation to be the same so that the wave still travels in the same direction (that is, so that the $\hat{\mathbf{k}}$ vector still points in the $+\hat{\mathbf{z}}$ direction). The total wave is now

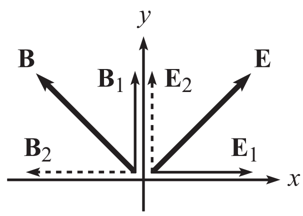


Figure 9

$$\mathbf{E} = (\hat{\mathbf{x}} + \hat{\mathbf{y}}) E_0 \cos(kz - \omega t), \qquad \text{and} \qquad \mathbf{B} = (\hat{\mathbf{y}} - \hat{\mathbf{x}}) \frac{E_0}{c} \cos(kz - \omega t). \tag{64}$$

If the two waves we added are labeled as "1" and "2" respectively, then the sum given in Eq. (64) is shown in Fig. 9. The wave travels in the positive z direction, which is out of the page. The $\mathbf{E}$ field in this wave always points along the (positive or negative) diagonal $\hat{\mathbf{x}} + \hat{\mathbf{y}}$ direction, and the $\mathbf{B}$ field always points along the $\hat{\mathbf{y}} - \hat{\mathbf{x}}$ direction. So we still have a linearly polarized wave. All we've done is rotate the fields by 45° and multiply the amplitudes by $\sqrt{2}$. Therefore, if our goal is to produce a wave that isn't linearly polarized (that is, to produce a wave where the directions of $\mathbf{E}$ and $\mathbf{B}$ change), we're going to have to come up with a more clever method than adding on fields that point in different directions.

Before proceeding further, we should note that no matter what traveling wave we have, the $\mathbf{B}$ field is completely determined by the $\mathbf{E}$ field and the $\mathbf{k}$ vector via Eq. (30). For a given $\mathbf{k}$ vector (we'll generally pick $\mathbf{k}$ to point along $\hat{\mathbf{z}}$), this means that the $\mathbf{E}$ field determines the $\mathbf{B}$ field (and vice versa). So we won't bother writing down the $\mathbf{B}$ field anymore. We'll just work with the $\mathbf{E}$ field.

We should also note that the x and y components of $\mathbf{B}$ are determined *separately* by the y and x components of $\mathbf{E}$, respectively. This follows from Eq. (30) and the properties of the cross product:

$$\begin{aligned}
\mathbf{k} \times \mathbf{E} = \omega \mathbf{B} \quad &\Longrightarrow \quad (k\hat{\mathbf{z}}) \times (E_x \hat{\mathbf{x}} + E_y \hat{\mathbf{y}}) = B_x \hat{\mathbf{x}} + B_y \hat{\mathbf{y}} \\
&\Longrightarrow \quad kE_x \hat{\mathbf{y}} + kE_y(-\hat{\mathbf{x}}) = B_x \hat{\mathbf{x}} + B_y \hat{\mathbf{y}} \\
&\Longrightarrow \quad B_x = -kE_y, \quad \text{and} \quad B_y = kE_x.
\end{aligned} \tag{65}$$

We see that the B_x and E_y pair of components is "decoupled" from the B_y and E_x pair. The two pairs have nothing to do with each other. Each pair can be doing whatever it feels like, independent of the other pair. So we basically have two independent electromagnetic waves. This is the key to understanding polarization.

8.6.2 Circular and elliptical polarization

Our above attempt at forming a non-linearly polarized wave failed because the two $\mathbf{E}$ fields that we added together had the *same phase*. This resulted in the two pairs of components

(E_x and B_y, and E_y and B_x) having the same phase, which in turn resulted in a simple (tilted) line for each of the total **E** and **B** fields.

Let us therefore try some different relative phases between the components. As mentioned above, from here on we'll write down only the **E** field. The **B** field can always be obtained from Eq. (30). Let's add a phase of, say, $\pi/2$ to the y component of **E**. As above, we'll have the magnitudes of the components be equal, so we obtain

$$\begin{aligned}\mathbf{E} &= \hat{\mathbf{x}}E_0\cos(kz-\omega t)+\hat{\mathbf{y}}E_0\cos(kz-\omega t+\pi/2)\\ &= \hat{\mathbf{x}}E_0\cos(kz-\omega t)-\hat{\mathbf{y}}E_0\sin(kz-\omega t).\end{aligned} \tag{66}$$

What does **E** look like as a function of time, for a given value of z? We might as well pick $z=0$ for simplicity, in which case we have (using the facts that cosine and sine are even and odd functions, respectively)

$$\mathbf{E}=\hat{\mathbf{x}}E_0\cos(\omega t)+\hat{\mathbf{y}}E_0\sin(\omega t). \tag{67}$$

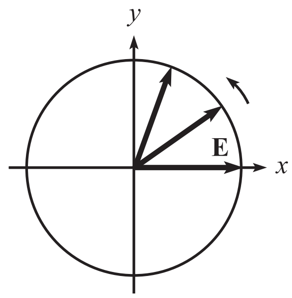


Figure 10

This is the expression for a vector with magnitude E_0 that swings around in a counterclockwise circle in the x-y plane, as shown in Fig. 10. (And at all times, **B** is perpendicular to **E**.) This is our desired *circular polarization*. The phase difference of $\pi/2$ between the x and y components of **E** causes **E** to move in a circle, as opposed to simply moving back and forth along a line. This is consistent with the fact that the phase difference implies that E_x and E_y can't both be zero at the same time, which is a necessary property of linear polarization, because the vector passes through the origin after each half cycle.

If we had chosen a phase of $-\pi/2$ instead of $\pi/2$, we would still have obtained circular polarization, but with the circle now being traced out in a clockwise sense (assuming that $\hat{\mathbf{z}}$ still points out of the page).

A phase of zero gives linear polarization, and a phase of $\pm\pi/2$ gives circular polarization. What about something in between? If we choose a phase of, say, $\pi/3$, then we obtain

$$\mathbf{E}=\hat{\mathbf{x}}E_0\cos(kz-\omega t)+\hat{\mathbf{y}}E_0\cos(kz-\omega t+\pi/3) \tag{68}$$

As a function of time at $z=0$, this takes the form,

$$\mathbf{E}=\hat{\mathbf{x}}E_0\cos(\omega t)+\hat{\mathbf{y}}E_0\cos(\omega t-\pi/3). \tag{69}$$

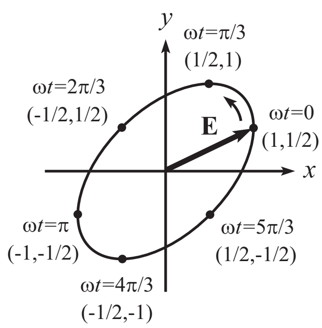


Figure 11

So the y component lags the x component by $\pi/3$. E_y therefore achieves its maximum value at a time given by $\omega t=\pi/3$ *after* E_x achieves its maximum value. A plot of **E** is shown in Fig. 11. A few of the points are labeled with their ωt values. This case is reasonably called *elliptical polarization*. The shape in Fig. 11 is indeed an ellipse, as you can show in Problem [to be added]. And as you can also show in this problem, the ellipse is always titled at $\pm 45°$. The ellipse is very thin if the phase is near 0 or π, and it equals a diagonal line in either of these limits. Linear and circular polarization are special cases of elliptical polarization. If you want to produce a tilt angle other than $\pm 45°$, you need to allow for E_x and E_y to have different amplitudes (see Problem [to be added]).

REMARK: We found above in Eq. (26) that the general solution for **E** is

$$\mathbf{E}=\mathbf{E}_0e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{70}$$

In looking at this, it appears that the various components of **E** have the same phase. So were we actually justified in throwing in the above phases of $\pi/2$ and $\pi/3$, or anything else? Yes, because as we mentioned right after Eq. (26), the $\mathbf{E}_0$ vector (and likewise the $\mathbf{B}_0$ vector) doesn't have to be real. Each component can be complex and have an arbitrary phase (although the three phases in $\mathbf{B}_0$ are determined by the three phases phases in $\mathbf{E}_0$ by Maxwell's equations). For example, we can have $E_{0,x} = A$ and $E_{0,y} = Ae^{i\phi}$. When we take the real part of the solution in Eq. (70), we then obtain

$$E_x = A\cos(\mathbf{k}\cdot\mathbf{r} - \omega t) \qquad \text{and} \qquad E_y = A\cos(\mathbf{k}\cdot\mathbf{r} - \omega t + \phi). \tag{71}$$

So this is the source of the relative phase, which in turn is the source of the various types of polarizations. ♣

Standing waves can also have different types of polarizations. Such a wave can be viewed as the sum of polarized waves traveling in opposite directions. But there are different cases to consider, depending on the orientation of the polarizations; see Problem [to be added].

8.6.3 Wave plates

Certain anisotropic materials (that is, materials that aren't symmetric around a given axis) have the property that electromagnetic waves that are linearly polarized along one axis travel at a different speed from waves that are linearly polarized along another (perpendicular) axis. This effect is known as *birefringence*, or *double refraction*, because it has two different speeds and hence two different indices of refraction, n_x and n_y. The difference in speeds and n values arises from the difference in permittivity values, ϵ, in the two directions. This difference in speed implies, as we will see, that the electric field components in the two directions will gradually get out of phase as the wave travels through the material. If the thickness of the material is chosen properly, we can end up with a phase difference of, say, $\pi/2$ (or anything else) which implies circular polarization. Let's see how this phase difference arises.

Let the two transverse directions be x and y, and let the E_x wave travel faster than the E_y wave. As mentioned right after Eq. (65), we can consider these components to be two separate waves. Let's assume that linearly polarized light traveling in the z direction impinges on the material and that it has nonzero components in both the x and y directions. Since this single wave is driving both the E_x and E_y waves in the material, these two components will be in phase with each other at the front end of the material. The material is best described as a plate (hence the title of this subsection), because the dimension along the direction of the wave's motion is generally small compared with the other two dimensions.

What happens as the E_x and E_y waves propagate through the material? Since the same external waves is driving both the E_x and E_y waves in the material, the frequencies of these waves must be equal. However, since the speeds are different, the $\omega = vk$ relation tells us that the k values must be different. Equivalently, the relation $\lambda\nu = v$ tells us that (since ν is the same) the wavelength is proportional to the velocity. So a smaller speed means a shorter wavelength. A possible scenario is shown in Fig. 12. We have assumed that the y speed is slightly smaller than the x speed, which means that λ_y is slightly shorter than λ_x. Equivalently, k_y is slightly larger than k_x. Therefore, slightly more E_y waves fit in the material than E_x waves, as shown. What does this imply about the phase difference between the E_x and E_y waves when they exit the material?

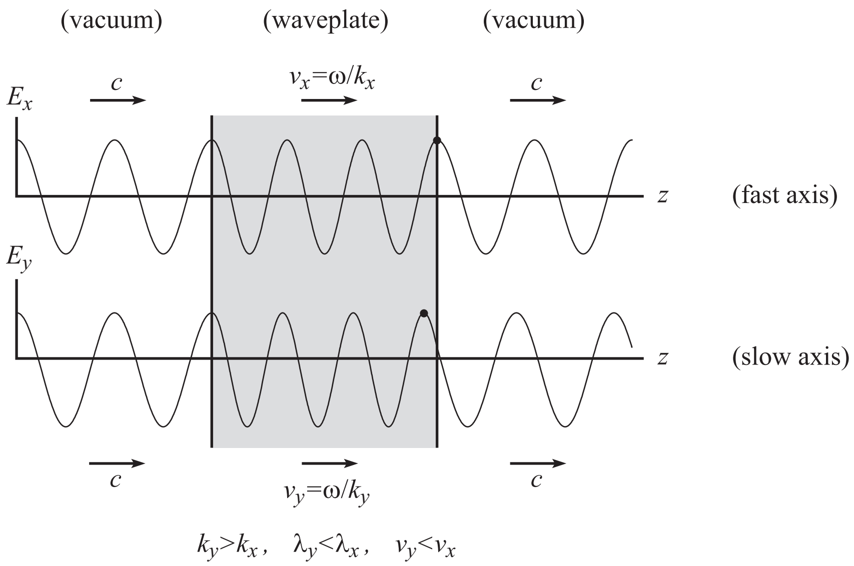


Figure 12

At the instant shown, the E_x field at the far end of the material has reached its maximum value, indicated by the dot shown. (It isn't necessary for an integral number of wavelengths to fit into the material, but it makes things a little easier to visualize.) But E_y hasn't reached its maximum quite yet. The E_y wave needs to travel a little more to the right before the crest marked by the dot reaches the far end of the material. So the phase of the E_y wave at the far end is slightly *behind* the phase of the E_x wave.[6]

By how much is the E_y phase behind the E_x phase at the far end? We need to find the phase of the E_y wave that corresponds to the extra little distance between the dots shown in Fig. 12. Let the length of the material (the thickness of the wave plate) be L. Then the number of E_x and E_y wavelengths that fit into the material are L/λ_x and L/λ_y, respectively, with the latter of these numbers being slightly larger. The number of *extra* wavelengths of E_y compared with E_x is therefore $L/\lambda_y - L/\lambda_x$. Each wavelength is worth 2π radians, so the phase of the E_y wave at the far end of the material is *behind* the phase of the E_x wave by an amount (we'll give four equivalent expressions here)

$$\Delta\phi = 2\pi L\left(\frac{1}{\lambda_y} - \frac{1}{\lambda_x}\right) = L(k_y - k_x) = \omega L\left(\frac{1}{v_y} - \frac{1}{v_x}\right) = \frac{\omega L}{c}(n_y - n_x), \tag{72}$$

where we have used $v_i = c/n_i$. In retrospect, we could have simply written down the second of these expressions from the definition of the wavenumber k, but we have to be careful to get the sign right. The fact that a *larger* number of E_y waves fit into the material means that the phase of the E_y wave is *behind* the phase of the E_x wave. Of course, if E_y is behind by a large enough phase, then it is actually better described as being ahead. For example, being behind by $7\pi/4$ is equivalent to being ahead by $\pi/4$. We'll see shortly how we can use wave plates to do various things with polarization, including creating circularly polarized light.

[6]You might think that the E_y phase should be ahead, because it has more wiggles in it. But this is exactly backwards. Of course, if you're counting from the left end of the plate, E_y does sweep through more phase than E_x. But that's not what we're concerned with. We're concerned with the phase of the wave as it passes the far end of the plate. And since the crest marked with the dot in the E_y wave hasn't reached the end yet, the E_y phase is behind the E_x phase. So in the end, it is correct to use the simplistic reasoning of, "a given crest on the slower wave takes longer to reach the end, so the phase of the slower wave is behind."

8.6.4 Making polarized light

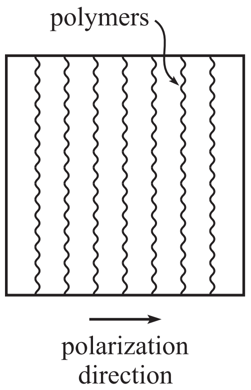


Figure 13

How can we produce polarized light? Let's look at linear polarization and then circular polarization. One way to produce linearly polarized light is to wiggle a charged particle in a certain way, so that the radiation is linearly polarized. We'll talk about how this works in Section 8.7.

Another way is to create a sheet of polymers (long molecules with repeating parts) that are stretched out parallel to each other, as shown in Fig. 13. If you have a light source that emits a random assortment of polarizations, then the sheet can filter out the light that is linearly polarized along a certain direction and leave you with only the light that is linearly polarized along the orthogonal direction. This works in the following way.

The electrons in the polymers are free to vibrate in the direction *along* the polymer, but not perpendicular to it. From the mechanism we discussed in Section 8.5, this leads to the absorption of the energy of the electric field that points along the polymer. So this component of the electric field shrinks to zero. Only the component *perpendicular* to the polymer survives. So we end up with linearly polarized light in the direction perpendicular to the polymers. You therefore can't think of the polymers as a sort of fence which lets through the component of the field that is parallel to the boards in the fence. It's the opposite of this.

Now let's see how to make circularly polarized light. As in the case of linear polarization, we can wiggle a charged particle in a certain way. But another way is to make use of the wave-plate results in the previous subsection. Let the fast and slow axes of the wave plate be the x and y axes, respectively. If we send a wave into the material that is polarized in either the x or y directions, then nothing exciting happens. The wave simply stays polarized along that direction. The phase difference we found in Eq. (72) is irrelevant if only one of the components exists.

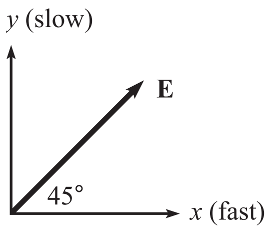


Figure 14

If we want to have anything useful come out of the phase difference caused by the plate, we need to have both components be nonzero. So let's assume that we have linearly polarized light that enters the material with a polarization direction at a 45° angle between the axes, as shown in Fig. 14. Given the frequency ω of the light, and given the two indices of refraction, n_x and n_y, let's assume that we've chosen the thickness L of the plate to yield a phase difference of $\Delta\phi = \pi/2$ in Eq. (72). Such a plate is called a *quarter-wave plate*, because $\pi/2$ is a quarter of a cycle. When the wave emerges from the plate, the E_y component is then $\pi/2$ behind the E_y component. So if we choose $t = 0$ to be the time when the wave enters the plate, we have (the phase advance of ϕ is unimportant here)

$$\begin{aligned}
\text{Enter plate}: \quad \mathbf{E} &\propto \hat{\mathbf{x}}\cos\omega t + \hat{\mathbf{y}}\cos\omega t, \\
\text{Leave plate}: \quad \mathbf{E} &\propto \hat{\mathbf{x}}\cos(\omega t + \phi) + \hat{\mathbf{y}}\cos(\omega t + \phi - \pi/2) \\
&= \hat{\mathbf{x}}\cos(\omega t + \phi) + \hat{\mathbf{y}}\sin(\omega t + \phi). \qquad (73)
\end{aligned}$$

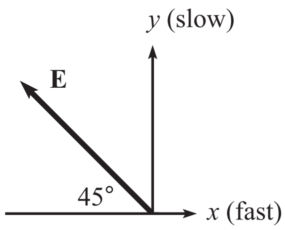


Figure 15

This wave has the same form as the wave in Eq. (67), so it is circularly polarized light with a counterclockwise orientation, as shown in above in Fig. 10.

What if we instead have an incoming wave polarized in the direction shown in Fig. 15? The same phase difference of $\pi/2$ arises, but the coefficient of the E_x part of the wave now has a negative sign in it. So we have

$$\begin{aligned}
\text{Enter plate}: \quad \mathbf{E} &\propto -\hat{\mathbf{x}}\cos\omega t + \hat{\mathbf{y}}\cos\omega t, \\
\text{Leave plate}: \quad \mathbf{E} &\propto -\hat{\mathbf{x}}\cos(\omega t + \phi) + \hat{\mathbf{y}}\cos(\omega t + \phi - \pi/2) \\
&= -\hat{\mathbf{x}}\cos(\omega t + \phi) + \hat{\mathbf{y}}\sin(\omega t + \phi). \qquad (74)
\end{aligned}$$

This is circularly polarized light, but now with a clockwise orientation.

The thickness of a quarter-wave plate (or a half-wave plate, or anything else) depends on the wavelength of the light, or equivalently on the various other quantities in Eq. (72). Intuitively, a longer wavelength means a longer distance to get ahead by a given fraction of that wavelength. So there is no "universal" quarter-wave plate that works for all wavelengths.

8.6.5 Malus' law

If a given electromagnetic wave encounters a linear polarizer, how much of the light makes it through? Consider light that is linearly polarized in the $\hat{\mathbf{x}}$ direction, and assume that we have a polarizer whose axis (call it the $\hat{\mathbf{x}}'$ axis) makes an angle of θ with the $\hat{\mathbf{x}}$ axis, as shown in Fig. 16. (This means that the polymers are oriented at an angle $\theta \pm \pi/2$ with the $\hat{\mathbf{x}}$ axis.) In terms of the primed coordinate system, the amplitude of the electric field is

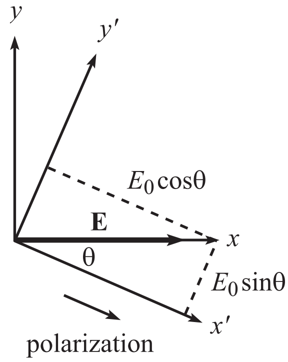


Figure 16

$$\mathbf{E} = E_0\hat{\mathbf{x}} = E_0(\hat{\mathbf{x}}' \cos\theta + \hat{\mathbf{y}}' \sin\theta). \tag{75}$$

The y' component gets absorbed by the polarizer, so we're left with only the x' component, which is $E_0 \cos\theta$. Hence, the polarizer decreases the amplitude by a factor of $\cos\theta$. The intensity (that is, the energy) is proportional to the square of the amplitude, which means that it is decreased by a factor of $\cos^2\theta$. Therefore,

$$|\mathbf{E}_{\text{out}}| = |\mathbf{E}_{\text{in}}| \cos\theta, \qquad \text{and} \qquad \boxed{I_{\text{out}} = I_{\text{in}} \cos^2\theta} \tag{76}$$

The second of these relations is known as *Malus' law*. Note that if $\theta = 90°$, then $I_{\text{out}} = 0$. So two successive polarizers that are oriented at 90° with respect to each other block all of the light that impinges on them, because whatever light makes it through the first polarizer gets absorbed by the second one.

What happens if we put a third polarizer between these two at an angle of 45° with respect to each? It seems that adding another polarizer can only make things "worse" as far as the transmission of light goes, so it seems like we should still get zero light popping out the other side. However, if a fraction f of the light makes it through the first polarizer (f depends on what kind of light you shine in), then $f \cos^2 45°$ makes it through the middle polarizer. And then a fraction $\cos^2 45°$ of this light makes it through the final polarizer. So the total amount that makes it through all three polarizers is $f \cos^4 45° = f/4$. This isn't zero! Adding the third polarizer makes things better, not worse.

This strange occurrence is due to the fact that polarizers don't act like filters of the sort where, say, a certain fraction of particles make it through a screen. In that kind of filter, a screen is always "bad" as far as letting particles through goes. The difference with actual polarizers is that the polarizer *changes* the polarization direction of whatever light makes it through. In contrast, if a particle makes it through a screen, then it's still the same particle. Another way of characterizing this difference is to note that a polarization is a vector, and vectors can be described in different ways, depending on what set of basis vectors is chosen. In short, in Fig. 17 the projection of **A** onto **C** is zero. But if we take the projection of **A** onto **B**, and then take the projection of the result onto **C**, the result isn't zero.

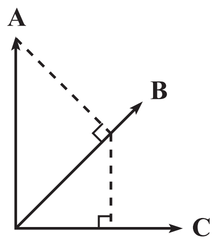


Figure 17

What happens if instead of inserting one intermediate polarizer at 45°, we insert two polarizers at angles 30° and 60°? Or three at 22.5°, 45°, and 67.5°, etc? Does more light or less light make it all the way through? The task of Problem [to be added] is to find out. You will find in this problem that something interesting happens in the case of a very large number of polarizers. The idea behind this behavior has countless applications in physics.

8.7 Radiation from a point charge

8.7.1 Derivation

In our above discussions of electromagnetic waves, we didn't worry about where the waves came from. We just studied their properties, given that they existed. Let's now look at how an electromagnetic wave can be created in the first place. We will find that an accelerating charge produces an electromagnetic wave.

Consider first a stationary charge. The field is simply radial,

$$\mathbf{E} = \hat{\mathbf{r}}\frac{q}{4\pi\epsilon_0 r^2}\,, \tag{77}$$

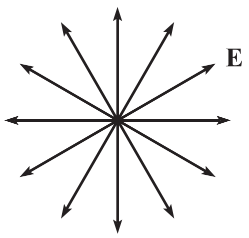


Figure 18

and there is no **B** field. This is shown in Fig. 18. If we instead have a charge moving with *constant* velocity **v**, then the field is also radial, but it is bunched up in the transverse direction, as shown in Fig. 19. This can be derived with the Lorentz transformations. However, the proof isn't important here, and neither is half of the result. All we care about is the radial nature of the field. We'll be dealing with speeds that are generally much less than c, in which case the bunching-up effect is negligible. We therefore again have

$$\mathbf{E} \approx \hat{\mathbf{r}}\frac{q}{4\pi\epsilon_0 r^2} \qquad \text{(for } v \ll c\text{)}. \tag{78}$$

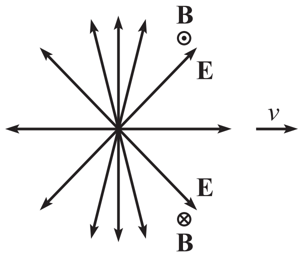


Figure 19

This is shown in Fig. 20. There is also a **B** field if the charge is moving. It points out of the page in the top half of Figs. 19 and 20, and into the page in the bottom half. This also follows from the Lorentz transformations (or simply by the righthand rule if you think of the moving charge as a current), but it isn't critical for the discussion. You can check with the righthand rule that $\mathbf{S} \propto \mathbf{E} \times \mathbf{B}$ points tangentially, which means that no power is radiated outward. And you can also check that **S** always has a forward component in the direction of the charge's velocity. This makes sense, because the field (and hence the energy) increases as the charge moves to the right, and **S** measures the flow of energy.

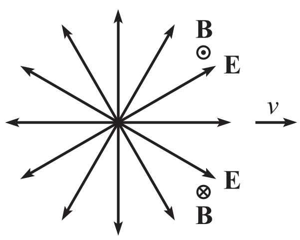


Figure 20

This radial nature of **E** for a moving charge seems reasonable (and even perhaps obvious), but it's actually quite bizarre. In Fig. 21, the field at point P points radially away from the present position of the charge. But how can P know that the charge is where it is *at this instant*? What if, for example, the charge stops shortly before the position shown? The field at P would still be directed radially away from where the charge *would have been* if it had kept moving with velocity **v**. At least for a little while. The critical fact is that the information that the charge has stopped can travel only at speed c, so it takes a nonzero amount of time to reach P. After this time, the field will point radially away from the stopped position, as expected.

A reasonable question to ask is then: What happens to the field during the transition period when it goes from being radial from one point (the projected position if the charge kept moving) to being radial from another point (the stopped position)? In other words, what is the field that comes about due to the acceleration? The answer to this question will tell us how an electromagnetic field is created and what it looks like.

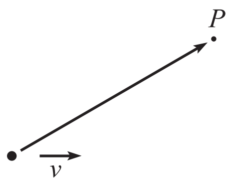


Figure 21

For concreteness, assume that the charge is initially traveling at speed v, and then let it decelerate with constant acceleration $-a$ for a time Δt (starting at $t = 0$) and come to rest. So $v = a\Delta t$, and the distance traveled during the stopping period is $(1/2)a\Delta t^2/2$. Let the origin of the coordinate system be located at the place where the deceleration starts.

Consider the situation at time T, where $T \gg \Delta t$. For example, let's say that the charge takes $\Delta t = 1\,\text{s}$ to stop, and we're looking at the setup $T = 1$ hour later. The distance $(1/2)a\Delta t^2/2$ is negligible compared with the other distances we'll be involved with, so we'll ignore it. At time T, positions with $r > cT$ have no clue that the charge has started to decelerate, so they experience a field directed radially away from the future projected position. Conversely, positions with $r < c(T - \Delta t)$ know that the charge has stopped, so they experience a field directed radially away from the origin (or actually a position $(1/2)a\Delta t^2/2$, but this is negligible). So we have the situation shown in Fig. 22.

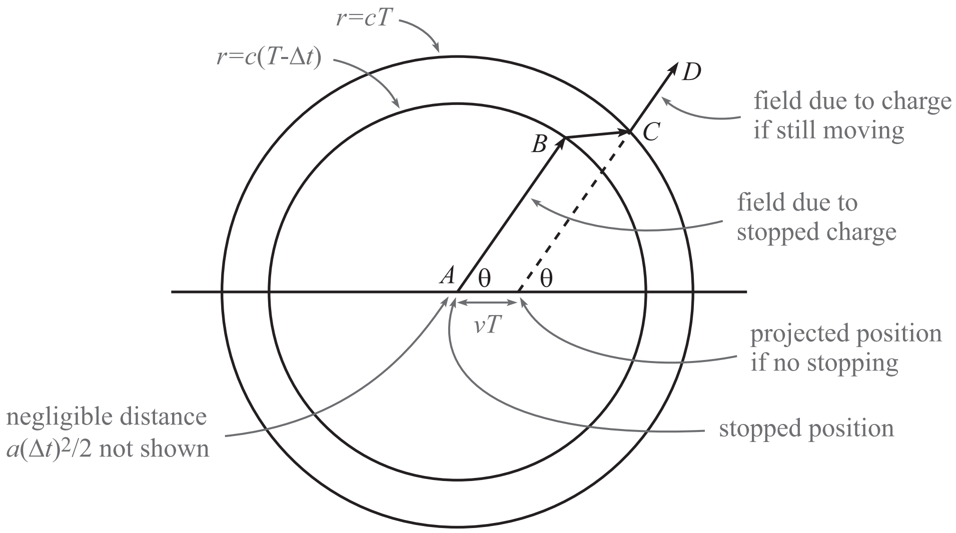


Figure 22

Since we're assuming $v \ll c$, the "exterior" field lines (the ones obtained by imagining that the charge is still moving) are essentially not compressed in the transverse direction. That is, they are spherically symmetric, as shown in Fig. 20. (More precisely, the compression effect is of order v^2/c^2 which is small compared with the effects of order v/c that we will find.) Consider the segments AB and CD in Fig. 22. These segments are chosen to make the same angle θ (which can be arbitrary) with the x axis, with AB passing through the stopped position, and the line of CD passing through the projected position. Due to the spherically symmetric nature of both the interior and exterior field lines, the surfaces of revolutions of AB and CD (which are parts of cones) enclose the same amount of flux. AB and CD must therefore be part of the same field line. This means that they are indeed connected by the "diagonal" field line BC shown.[7]

If we expand the relevant part of Fig. 22, we obtain a picture that takes the general form shown in Fig. 23. Let the radial and tangential components of the $\mathbf{E}$ field in the transition region be E_r and E_θ. From similar triangles in the figure, we have

$$\frac{E_\theta}{E_r} = \frac{vT \sin\theta}{c\Delta t}\,. \tag{79}$$

Note that the righthand side of this grows with T. Again, the units of E_θ and E_r aren't distance, so the size of the $\mathbf{E}$ vector in Fig. 23 is meaningless. But all the matters in the above similar-triangle argument is that the vector points in the direction shown.

[7] We're drawing both field lines and actual distances in this figure. This technically makes no sense, of course, because the fields don't have units of distance. But the point is to show the directions of the fields. We could always pick pick our unit size of the fields to be the particular value that makes the lengths on the paper be the ones shown.

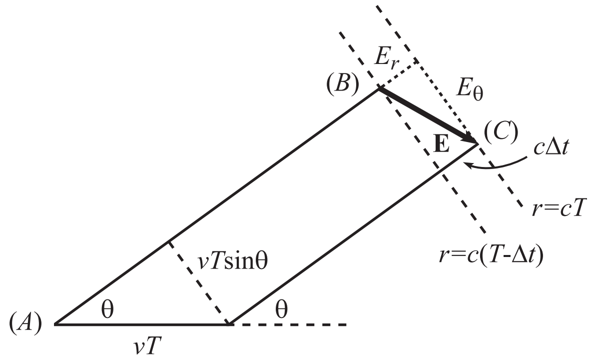


Figure 23

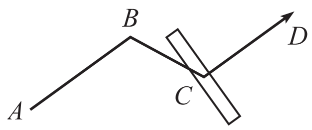


Figure 24

We now claim that E_r has the same value just outside and just inside the transition region. This follows from a Gauss's-law argument. Consider the pillbox shown in Fig. 24, which is located at the kink in the field at point C. The long sides of the box are oriented in the tangential direction, and also in the direction perpendicular to the page. The short sides are chosen to be infinitesimally small, so E_θ contributes essentially nothing to the flux. The flux is therefore due only to the radial component, so we conclude that $E_r^{\text{transition}} = E_r^{\text{outside}}$. (And likewise, a similar pillbox at point B tells us that $E_r^{\text{transition}} = E_r^{\text{inside}}$, but we won't need this. $E_r^{\text{transition}}$ varies slightly over the transition region, but the change is negligible if Δt is small.) We know that

$$E_r^{\text{outside}} = \frac{q}{4\pi\epsilon_0 r^2} = \frac{q}{4\pi\epsilon_0 (cT)^2}\,. \tag{80}$$

Eq. (79) then gives

$$\begin{aligned} E_\theta = \frac{vT\sin\theta}{c\Delta t}E_r &= \frac{vT\sin\theta}{c\Delta t}\cdot\frac{q}{4\pi\epsilon_0 (cT)^2} \\ &= \frac{q\sin\theta}{4\pi\epsilon_0 c^2 (cT)}\cdot\frac{v}{\Delta t} \\ &= \frac{qa\sin\theta}{4\pi\epsilon_0 rc^2} \qquad (\text{using } \ a = v/\Delta t \ \text{ and } \ r = cT) \end{aligned} \tag{81}$$

So the field inside the transition region is given by

$$\boxed{(E_r, E_\theta) = \frac{q}{4\pi\epsilon_0}\left(\frac{1}{r^2}\,,\,\frac{a\sin\theta}{rc^2}\right)} \tag{82}$$

Both of the components in the parentheses have units of $1/\text{m}^2$, as they should. We will explain below why this field leads to an electromagnetic wave, but first some remarks.

REMARKS:

1. The location of the E_θ field (that is, the transition region) propagates outward with speed c.
2. E_θ is proportional to a. Given this fact, you can show that the only way to obtain the right units for E_θ using a, r, c, and θ, is to have a function of the form, $af(\theta)/rc^2$, where $f(\theta)$ is an arbitrary function of θ. And $f(\theta)$ happens to be $\sin\theta$ (times $q/4\pi\epsilon_0$).
3. If $\theta = 0$ or $\theta = \pi$, then $E_\theta = 0$. In other words, there is no radiation in the forward or backward directions. E_θ is maximum at $\theta = \pm\pi/2$, that is, in the transverse direction.

4. The acceleration vector associated with Fig. 22 points to the left, since the charge was decelerating. And we found that E_θ has a rightward component. So in general, the E_θ vector is

$$\mathbf{E}_\theta(\mathbf{r}, t) = -\frac{q}{4\pi\epsilon_0} \cdot \frac{\mathbf{a}_\perp(t')}{rc^2}\,, \tag{83}$$

where $t' \equiv t - r/c$ is the time at which the kink at point C in Fig. 22 was emmitted, and where $\mathbf{a}_\perp$ is the component of $\mathbf{a}$ that is perpendicular to the radial direction. In other words, it is the component with magnitude $a\sin\theta$ that you "see" across your vision if you are located at position $\mathbf{r}$. This is consistent with the previous remark, because $\mathbf{a}_\perp = \mathbf{0}$ if $\theta = 0$ or $\theta = \pi$. Note the minus sign in Eq. (83).

5. Last, but certainly not least, we have the extremely important fact: For sufficiently large r, E_r is negligible compared with E_θ. This follows from the fact that in Eq. (82), E_θ has only one r in the denominator, whereas E_r has two. So for large r, we can ignore the "standard" radial part of the field. We essentially have only the new "strange" tangential field. By "large r," we mean $a/rc^2 \gg 1/r^2 \Longrightarrow r \gg c^2/a$. Or equivalently $\sqrt{ra} \gg c$. In other words, ignoring relativity and using the kinematic relation $v = \sqrt{2ad}$, the criterion for large r is that (in an order-of-magnitude sense) if you accelerate something with acceleration a for a distance r, its velocity will exceed c.

   The reason why E_θ becomes so much larger than E_r is because there is a T in the numerator of Eq. (79). This T follows from the fact that in Fig. 23, the E_θ component of $\mathbf{E}$ grows with time (because vT, which is the projected position of the charge if it kept moving, grows with time), whereas E_r is always proportional to the constant quantity, $c\Delta t$.

The above analysis dealt with constant acceleration. However, if the acceleration is changing, we can simply break up time into little intervals, with the above result holding for each interval (as long as T is large enough so that all of our approximations hold). So even if $\mathbf{a}$ is changing, $\mathbf{E}_\theta(\mathbf{r}, t)$ is proportional to whatever $-\mathbf{a}_\perp(t')$ equaled at time $t' \equiv t - r/c$. In particular, if the charge is wiggling sinusoidally, then $\mathbf{E}_\theta(\mathbf{r}, t)$ is a sinusoidal wave.

The last remark above tells us that if we're far away from an accelerating charge, then the only electric field we see is the tangential one; there is essentially no radial component. There is also a magnetic field, which from Maxwell's equations can be shown to also be tangential, perpendicular to the page in Fig. 22; see Problem [to be added]. So we have electric and magnetic fields that oscillate in the tangential directions while propagating with speed c in the radial direction. But this is exactly what happens with an electromagnetic wave. We therefore conclude that an electromagnetic wave can be generated by an accelerating charge.

Of course, we know from Section 8.3 that Maxwell's equations in vacuum imply that the direction of the $\mathbf{E}$ and $\mathbf{B}$ fields must be perpendicular to the propagation direction, so in retrospect we know that this also has to be the case for whatever fields popped out of the above analysis. The main new points of this analysis are that (1) an accelerating charge can generate the electromagnetic wave (before doing this calculation, for all we know a nonzero, say, third derivative of the position is needed to generate a wave), and (2) the radial field in Eq. (82) essentially disappears due to the $1/r^2$ vs. $1/r$ behavior, leaving us with only the tangential field.

Eq. (30) gives the magnetic field as $\mathbf{k} \times \mathbf{E} = \omega\mathbf{B} \Longrightarrow \hat{\mathbf{r}} \times \mathbf{E} = c\mathbf{B}$ (using $\mathbf{k} = k\hat{\mathbf{r}}$). So in the top half of Fig. 22, $\mathbf{B}$ points into the page with magnitude $B = E/c$. And in the bottom half it points out of the page. These facts are consistent with the cylindrical symmetry of the system around the horizontal axis. If the charge is accelerating instead of decelerating as we chose above, then the $\mathbf{E}$ and $\mathbf{B}$ fields are reversed.

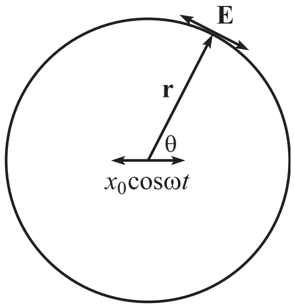


Figure 25

8.7.2 Poynting vector

What is the energy flow of a wave generated by a sinusoidally oscillating charge? Let the position of the charge be $x(t) = x_0 \cos\omega t$. The acceleration is then $a(t) = -\omega^2 x_0 \cos\omega t$. The resulting electric field at an arbitrary point is shown in Fig. 25. The energy flow at this point is given by the Poynting vector, which from Eqs. (49) and (50) is

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B} = c\epsilon_0 E^2 \hat{\mathbf{r}}. \tag{84}$$

Since E is given by the E_θ in Eq. (82), the average value of the magnitude of $\mathbf{S}$ is (using $a = -\omega^2 x_0 \cos\omega t$, along with the fact that the average value of $\cos^2\omega t$ is 1/2)

$$\begin{aligned} S_{\text{avg}} &= c\epsilon_0 \left(\frac{q}{4\pi\epsilon_0}\cdot\frac{a\sin\theta}{rc^2}\right)^2 \\ &= \frac{q^2}{16\pi^2\epsilon_0 c^3}\cdot\frac{1}{r^2}(\omega^2 x_0)^2 \sin^2\theta\cdot\frac{1}{2} \\ &= \boxed{\frac{\omega^4 x_0^2 q^2 \sin^2\theta}{32\pi^2\epsilon_0 c^3}\cdot\frac{1}{r^2}} \end{aligned} \tag{85}$$

If you want, you can write the $1/\epsilon_0 c^3$ part of this as μ_0/c. Note that S_{avg} falls off like $1/r^2$ and is proportional to ω^4. Note also that it is zero if $\theta = 0$ or $\theta = \pi$.

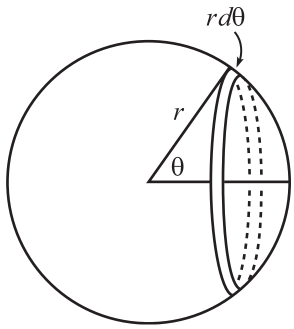


Figure 26

The Poynting vector has units of Energy/(time · area). Let's integrate it over a whole sphere of radius r to find the total Energy/time, that is, the total power. If we parameterize the integral by θ, then we can slice the sphere into rings as shown in Fig. 26. The circumference of a ring is $2\pi r\sin\theta$, and the width is $r\,d\theta$. So the total power is

$$\begin{aligned} P = \int_{\text{sphere}} S_{\text{avg}} &= \int_0^\pi S_{\text{avg}}(2\pi r\sin\theta)(r\,d\theta) \\ &= \frac{\omega^4 x_0^2 q^2}{32\pi^2\epsilon_0 c^3}\cdot\frac{2\pi r^2}{r^2}\int_0^\pi \sin^3\theta\,d\theta \\ &= \boxed{\frac{\omega^4 x_0^2 q^2}{12\pi\epsilon_0 c^3}} \end{aligned} \tag{86}$$

where we have used the fact that $\int_0^\pi \sin^3\theta\,d\theta = 4/3$. You can quickly verify this by writing the integrand as $(1-\cos^2\theta)\sin\theta$. There are two important features of this result for P. First, it is independent of r. This must be the case, because if more (or less) energy crosses a sphere at radius r_1 than at radius r_2, then energy must be piling up (or be taken from) the region in between. But this can't be the case, because there is no place for the energy to go. Second, P is proportional to $(\omega^2 x_0)^2$, which is the square of the amplitude of the acceleration. So up to constant numbers and physical constants, we have

$$P \propto a_0^2 q^2, \tag{87}$$

were a_0 is the amplitude of the acceleration.

8.7.3 Blue sky

The fact that the P in Eq. (86) is proportional to ω^4 has a very noticeable consequence in everyday life. It is the main reason why the sky is blue. (The exponent doesn't have to be 4. Any reasonably large number would do the trick, as we'll see.) In a nutshell, the

ω^4 factor implies that blue light (which is at the high-frequency end of visible spectrum) scatters more easily than red light (which is at the low-frequency end of visible spectrum). So if random white light (composed of many different frequencies) hits the air molecule shown in Fig. 27, the blue light is more likely to scatter and hit your eye, whereas the other colors with smaller frequencies are more likely to pass straight through. The sky in that direction therefore looks blue to you. More precisely, the intensity (power per area) of blue light that is scattered to your eye is larger than the intensity of red light by a factor of $P_{\text{blue}}/P_{\text{red}} = \omega^4_{\text{blue}}/\omega^4_{\text{red}}$. And since $\omega_{\text{blue}}/\omega_{\text{red}} \approx 1.5$, we have $P_{\text{blue}}/P_{\text{red}} \approx 5$. So 5 times as much blue light hits your eye.

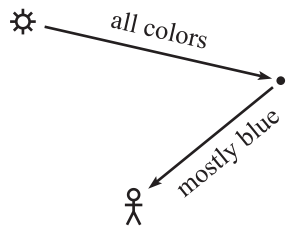


Figure 27

This also explains why sunsets are red. When the sun is near the horizon, the light must travel a large distance through the atmosphere (essentially tangential to the earth) to reach your eye, much larger than when the sun is high up in the sky. Since blue light scatters more easily, very little of it makes it straight to your eye. Most of it gets scattered in various directions (and recall that none of it gets scattered directly forward, due to the $\sin^2\theta$ factor in Eq. (85)). Red light, on the other hand, scatters less easily, so it is more likely to make it all the way through the atmosphere in a straight line from the sun to your eye. Pollution adds to this effect, because it adds particles to the air, which strip off even more of the blue light by scattering. So for all the bad effects of pollution, cities sometimes have the best sunsets. A similar situation arises with smoke. If you look at the sun through the smoke of a forest fire, it appears as a crisp red disk (but don't look at it for too long).

The actual scattering process is a quantum mechanical one involving photons, and it isn't obvious how this translates to our electromagnetic waves. But for the present purposes, it suffices to think about the scattering process as one where a wave with a given intensity encounters a region of molecules, and the molecules grab chunks of energy and throw them off in some other direction. (The electrons in the molecules are the things that are vibrating/accelerating and creating the radiation). The point is that with blue light, the chunks of energy are 5 times as large as they are for red light.

There are, however, a number of issues that we've glossed over. The problem is rather complicated when everything is included. In particular, one issue is that in addition to the ω^4 factor, the P in Eq. (86) is also proportional to x_0^2. What if the electron's x_0 value for red light is larger than the value for blue light? It turns out that it isn't; the x_0's are all essentially the same size. This can be shown by treating the electron in the atom as an essentially undamped driven oscillator. The natural frequency ω_0 depends on the nature of the atom, and it turns out that it is much larger than the frequency ω of light in the visible spectrum (we'll just accept this fact). The driven-oscillator amplitude is given in Eq. (1.88), and when $\gamma \approx 0$ and $\omega_0 \gg \omega$, it reduces to being proportional to $1/\omega_0^2$. That is, it is independent of ω, as we wanted to show.

Other issues that complicate things are: Is there multiple scattering? (The answer is generally no.) Why is the sky not violet, in view of the fact that $\omega_{\text{violet}} > \omega_{\text{blue}}$? How does the eye's sensitivity come into play? (It happens to be peaked at green.) So there are certainly more things to consider. But the ω^4 issue we covered above can quite reasonably be called the main issue.

8.7.4 Polarization in the sky

If you look at the daytime sky with a polarizer (polarized sunglasses do the trick) and rotate it in a certain way (the polarizer, not the sky, although that would suffice too), you will find that a certain region of the sky look darker. The reason for this is that the light that makes it to your eye after getting scattered from this region is polarized. To see why, consider an electron in the air that is located at a position such that the line from it to you is perpendicular to the line from the sun to it (which is the $\hat{\mathbf{k}}$ direction of the sun's

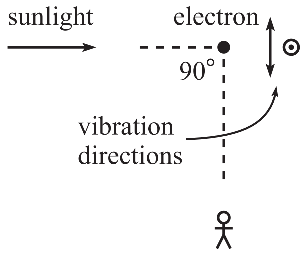


Figure 28

light); see Fig. 28. The radiation from the sun may cause this electron to vibrate. And then because it is vibrating/accelerating, it radiates light which may end up in your eye.

The electric field in the sun's light lies in the plane perpendicular to $\hat{\mathbf{k}}$. So the field is some combination of the two directions indicated in Fig. 28. The field can have a component along the line from the electron to you, and also a component perpendicular to the page, signified by the $\odot$ in the figure. The sun's light is randomly polarized, so it contains some of each of these. The electric field causes the electron to vibrate, and from the general force law $\mathbf{F} = q\mathbf{E}$, the electron vibrates in some combination of these two directions. However, due to the $\sin^2\theta$ factor in Eq. (85), you don't see any of the radiation that arises from the electron vibration along the line between it and you. Therefore, the only light that reaches your eye is the light that was created from the vibration pointing perpendicular to the page. Hence all of the light you see is polarized in this direction. So if your sunglasses are oriented perpendicular to this direction, then not much light makes it through, and the sky looks dark.

Note how the three possible directions of the resulting $\mathbf{E}$ field got cut down to one. First, the electric field that you see must be perpendicular to the $\hat{\mathbf{k}}$, due to the transverse nature of light (which is a consequence of Maxwell's equations), and due to the fact that the electron vibrates along the direction of $\mathbf{E}$. And second, the field must be perpendicular to the line from the electron to you, due to the "no forward scattering" fact that arises from the $\sin^2\theta$ factor in Eq. (85). Alternatively, we know that the $\mathbf{E}$ field that you see can't have a longitudinal component.

It's easy to see that conversely if the electron is *not* located at a position such that the line from it to you is perpendicular to the line from the sun to it, then you will receive some light that came from the vertical (on the page, in Fig. 28) oscillation of the electron. But the amount will be small if the angle is near 90°. So the overall result is that there is a reasonably thick band in the sky that looks fairly dark when viewed through a polarizer. If the sun is directly overhead, then the band is a circle at the horizon. If the sun is on the horizon, then the band is a semicircle passing directly overhead, starting and ending at the horizon.

8.8 Reflection and transmission

We'll now study the reflection and transmission that arise when light propagating in one medium encounters another medium. For example, we might have light traveling through air and then encountering a region of glass. We'll begin with the case of normal incidence and then eventually get to the more complicated case of non-normal incidence. For the case of normal incidence, we'll start off by considering only one boundary. So if light enters a region of glass, we'll assume that the glass extends infinitely far in the forward direction.

8.8.1 Normal incidence, single boundary

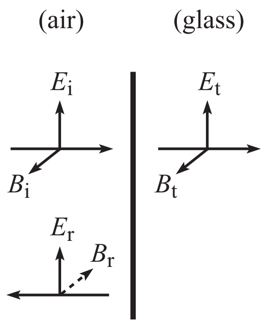


Figure 29

Consider light that travels to the right and encounters an air/glass boundary. As with other waves we've discussed, there will be a reflected wave and a transmitted wave. Because the wave equation in Eq. (15) is dispersionless, all waves travel with the same speed (in each region). So we don't need to assume anything about the shape of the wave, although we generally take it to be a simple sinusoidal one.

The incident, reflected, and transmitted waves are shown in Fig. 29 (the vertical displacement in the figure is meaningless). We have arbitrarily defined all three electric fields to be positive if they point upward on the page. A negative value of E_i, E_r, or E_t simply means that the vector points downward. We aren't assuming anything about the polarization of the wave. No matter what kind of wave we have, the incident electric field at a given instant in time points in some direction, and we are taking that direction to be upward (which we can arrange by a suitable rotation of the axes). The waves we have drawn are understood to be the waves infinitesimally close to the boundary on either side.

Given these positive conventions for the E's, and given the known directions of the three $\hat{\mathbf{k}}$ vectors, the positive directions of the magnetic fields are determined by $\hat{\mathbf{k}} \times \mathbf{E} = \omega \mathbf{B}$, and they are shown. Note that if E_r has the same orientation as E_i (the actual sign will be determined below), then B_r must have the opposite orientation as B_i, because the $\hat{\mathbf{k}}$ vector for the reflected waves is reversed. If we define positive B to be out of the page, then the total $\mathbf{E}$ and $\mathbf{B}$ fields in the left and right regions (let's call these regions 1 and 2, respectively) are

$$\begin{aligned} E_1 &= E_\text{i} + E_\text{r}, \qquad \text{and} \qquad E_2 = E_\text{t}, \\ B_1 &= B_\text{i} - B_\text{r}, \qquad \text{and} \qquad B_2 = B_\text{t}. \end{aligned} \tag{88}$$

What are the boundary conditions? There are four boundary conditions in all: two for the components of the electric and magnetic fields parallel to the surface, and two for the components perpendicular to the surface. However, we'll need only the parallel conditions for now, because all of the fields are parallel to the boundary. The perpendicular ones will come into play when we deal with non-normal incidence in Section 8.8.3 below. The two parallel conditions are:

Figure 30

- Let's first find the boundary condition on $E^{\parallel}$ (the superscript "$\parallel$" stands for parallel). The third of Maxwell's equations in Eq. (23) is $\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t$. Consider the very thin rectangular path shown in Fig. 30. If we integrate each side of Maxwell's equation over the surface S bounded by this path, we obtain

$$\int_S \nabla \times \mathbf{E} = -\int_S \frac{\partial \mathbf{B}}{\partial t}\,. \tag{89}$$

By Stokes' theorem, the lefthand side equals the integral $\int \mathbf{E} \cdot d\ell$ over the rectangular path. And the right side is $-\partial \Phi_B/\partial t$, where Φ_B is the magnetic flux through the surface S. But if we make the rectangle arbitrarily thin, then the flux is essentially zero. So Eq. (89) becomes $\int \mathbf{E} \cdot d\ell = 0$. The short sides of the rectangle contribute essentially nothing to this integral, and the contribution from the long sides is $E_2^{\parallel}\ell - E_1^{\parallel}\ell$, where ℓ is the length of these sides. Since this difference equals zero, we conclude that

$$\boxed{E_1^{\parallel} = E_2^{\parallel}} \tag{90}$$

The component of the electric field that is parallel to the boundary is therefore continuous across the boundary. This makes sense intuitively, because the effect of the dielectric material (the glass) is to at most have charge pile up on the boundary, and this charge doesn't affect the field parallel to the boundary. (Or if it does, in the case where the induced charge isn't uniform, it affects the two regions in the same way.)

- Now let's find the boundary condition on $B^{\parallel}$. Actually, what we'll find instead is the boundary condition on $H^{\parallel}$, where the $\mathbf{H}$ field is defined by $\mathbf{H} \equiv \mathbf{B}/\mu$. We're using $\mathbf{H}$ here instead of $\mathbf{B}$ because $\mathbf{H}$ is what appears in the fourth of Maxwell's equations in Eq. (23), $\nabla \times \mathbf{H} = \partial \mathbf{D}/\partial t + \mathbf{J}_\text{free}$. We need to use this form because we're working with a dielectric.

  There are no free currents anywhere in our setup, so $\mathbf{J}_\text{free} = 0$. We can therefore apply to $\nabla \times \mathbf{H} = \partial \mathbf{D}/\partial t$ the exact same reasoning with the thin rectangle that we used above for the electric field, except now the long sides of the rectangle are perpendicular to the page. We immediately obtain

$$\boxed{H_1^{\|} = H_2^{\|}} \tag{91}$$

The boundary condition for the $B^{\|}$ fields is then $B_1^{\|}/\mu_1 = B_2^{\|}/\mu_2$. However, since the μ value for most materials is approximately equal to the vacuum μ_0 value, the $B^{\|}$ fields also approximately satisfy $B_1^{\|} = B_2^{\|}$.

We can now combine the above two boundary conditions and solve for E_r and E_t in terms of E_i, and also H_r and H_t in terms of H_i. (The B's are then given by $B \equiv H/\mu$.) We can write the $H^{\|}$ boundary condition in terms of E fields by using

$$H \equiv \frac{B}{\mu} = \frac{E}{v\mu} \equiv \frac{E}{Z}, \qquad \text{where } \boxed{Z \equiv v\mu} \tag{92}$$

This expression for Z is by definition the impedance for an electromagnetic field. (We're using v for the speed of light in a general material, which equals $1/\sqrt{\mu\epsilon}$ from Eq. (19). We'll save c for the speed of light in vacuum.) Z can alternatively be written as

$$Z \equiv v\mu = \frac{\mu}{\sqrt{\mu\epsilon}} = \sqrt{\frac{\mu}{\epsilon}}\,. \tag{93}$$

The two boundary conditions can now be written as

$$\begin{aligned} E_1^{\|} = E_2^{\|} &\implies E_\text{i} + E_\text{r} = E_\text{t}, \\ H_1^{\|} = H_2^{\|} &\implies H_\text{i} - H_\text{r} = H_\text{t} \implies \frac{E_\text{i}}{Z_1} - \frac{E_\text{r}}{Z_1} = \frac{E_\text{t}}{Z_2}\,. \end{aligned} \tag{94}$$

These are two equations in the two unknowns E_r and E_t. Solving for them in terms of E_i gives

$$\boxed{E_\text{r} = \frac{Z_2 - Z_1}{Z_2 + Z_1}\, E_\text{i}} \qquad \text{and} \qquad \boxed{E_\text{t} = \frac{2Z_2}{Z_2 + Z_1}\, E_\text{i}} \tag{95}$$

These are similar (but not identical) to the reflection and transmission expressions for a transverse wave on a string; see Eq. (4.38).

We'll generally be concerned with just the E values, because the B values can always be found via Maxwell's equations. But if you want to find the H and B values, you can write the E's in the first boundary condition in terms of the H's. The boundary conditions become

$$\begin{aligned} E_1^{\|} = E_2^{\|} &\implies Z_1 H_\text{i} + Z_1 H_\text{r} = Z_2 H_\text{t}, \\ H_1^{\|} = H_2^{\|} &\implies H_\text{i} - H_\text{r} = H_\text{t}. \end{aligned} \tag{96}$$

Solving for H_r and H_t gives

$$H_\text{r} = \frac{Z_2 - Z_1}{Z_2 + Z_1}\, H_\text{i}, \qquad \text{and} \qquad H_\text{t} = \frac{2Z_1}{Z_2 + Z_1}\, H_\text{i}. \tag{97}$$

Remember that positive H_r is defined to point into the page, whereas positive H_i points out of the page, as indicated in Fig. 29. So the signed statement for the vectors is $\mathbf{H}_\text{r} = \mathbf{H}_\text{i} \cdot (Z_1 - Z_2)/(Z_1 + Z_2)$. If you want to find the B values, they are obtained via $B = \mu H$.

But again, we'll mainly be concerned with just the E values. Note that you can also quickly obtain these H's by using $E = Hz$ in the results for the E's in Eq. (95),

We can write the above expressions in terms of the index of refraction, which we defined in Eq. (20). We'll need to make an approximation, though. Since most dielectrics have $\mu \approx \mu_0$, Eq. (21) gives $n \propto \sqrt{\epsilon}$, and Eq. (93) gives $Z \propto 1/\sqrt{\epsilon}$. So we have $Z \propto 1/n$. The expressions for E_r and E_t in Eq. (95) then become

$$\boxed{E_\text{r} = \frac{n_1 - n_2}{n_1 + n_2}\, E_\text{i}} \qquad \text{and} \qquad \boxed{E_\text{t} = \frac{2n_1}{n_1 + n_2}\, E_\text{i}} \qquad (\text{if } \mu \approx \mu_0). \tag{98}$$

In going from, say, air to glass, we have $n_1 = 1$ and $n_2 \approx 1.5$. Since $n_1 < n_2$, this means that E_r has the opposite sign of E_i. A reflection like this, where the wave reflects off a region of higher index n, is called a "hard reflection." The opposite case with a lower n is called a "soft reflection."

What is the power in the reflected and transmitted waves? The magnitude of the Poynting vector for a traveling wave is given by Eq. (49) as $S = E^2/v\mu$, where we are using v and μ instead of c and μ_0 to indicate an arbitrary dielectric. But $v\mu$ is by definition the impedance Z, so the instantaneous power is (we'll use "P" instead of "S" here)

$$\boxed{P = \frac{E^2}{Z}} \tag{99}$$

Now, it must be the case that the incident power P_i equals the sum of the reflected and transmitted powers, P_r and P_t. (All of these P's are the instantaneous values at the boundary.) Let's check that this is indeed true. Using Eq. (95), the reflected and transmitted powers are

$$\begin{aligned} P_\text{r} &= \frac{E_\text{r}^2}{Z_1} = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2 \frac{E_\text{i}^2}{Z_1} = \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2}\, P_\text{i}, \\ P_\text{t} &= \frac{E_\text{t}^2}{Z_2} = \left(\frac{2Z_2}{Z_2 + Z_1}\right)^2 \frac{E_\text{i}^2}{Z_2} = \frac{4Z_1Z_2}{(Z_2 + Z_1)^2}\,\frac{E_\text{i}^2}{Z_1} = \frac{4Z_1Z_2}{(Z_2 + Z_1)^2}\, P_\text{i}. \end{aligned} \tag{100}$$

We then quickly see that $P_\text{r} + P_\text{t} = P_\text{i}$, as desired.

The following topics will eventually be added:

8.8.2 Normal incidence, double boundary

8.8.3 Non-normal incidence

Huygens' principle, Snell's law

Maxwell's equations

Chapter 9

Interference and diffraction

 (Version 1, June 25, 2010)

This file contains the "Interference and diffraction" chapter of a potential book on Waves, designed for college sophomores.

In this chapter we'll study what happens when waves from two or more sources exist at a given point in space. In the case of two waves, the total wave at the given point is the sum of the two waves. The waves can add constructively if they are in phase, or destructively if they are out of phase, or something inbetween for other phases. In the general case of many waves, we need to add them all up, which involves keeping track of what all the phases are. The results in this chapter basically boil down to (as we'll see) getting a handle on the phases and adding them up properly. We won't need to worry about various other wave topics, such as dispersion, polarization, and so on; it pretty much all comes down to phases. The results in this chapter apply to any kind of wave, but for convenience we'll generally work in terms of electromagnetic waves.

The outline of this chapter is as follows. In Section 9.1 we do the warm-up case of two waves interfering. The setup consists of a plane wave passing through two very narrow (much narrower than the wavelength of the wave) slits in a wall, and these two slits may be considered to be the two sources. We will calculate the interference pattern on a screen that is located far away. We'll be concerned with this "far-field" limit for most of this chapter, with the exception of Section 9.5. In Section 9.2 we solve the general case of interference from N narrow slits. In addition to showing how the phases can be added algebraically, we show how they can be added in an extremely informative geometric manner. In Section 9.3 we switch gears from the case of many narrow slits to the case of one wide slit. The word "diffraction" is used to describe the interference pattern that results from a slit with non-negligible width. We will see, however, that this still technically falls into the category of N narrow slits, because one wide slit can be considered to be a collection of a large (infinite) number of narrow slits. In section 9.4 we combine the results of the two previous sections and calculate the interference pattern from N wide slits. Finally, in Section 9.5 we drop the assumption that the screen is far away from the slit(s) and discuss "near-field" interference and diffraction. This case is a bit more complicated, but fortunately there is still a nice geometric way of seeing how things behave. This involves a very interesting mathematical curve known as the *Cornu spriral*.

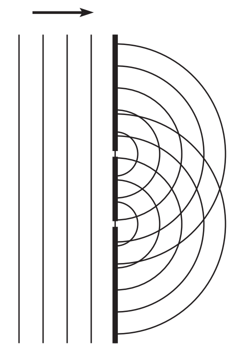


Figure 1

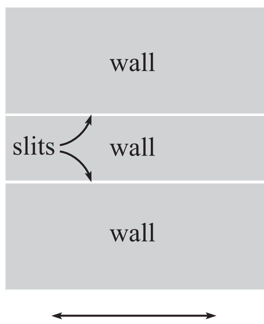


Figure 2

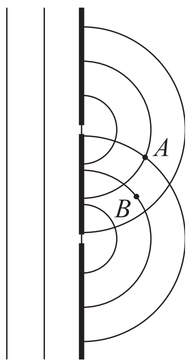


Figure 3

9.1 Two-slit interference

Consider a plane wave moving toward a wall, and assume that the wavefronts are parallel to the wall, as shown in Fig. 1. If you want, you can consider this plane wave to be generated by a point source that is located a very large distance to the left of the wall. Let there be two very small slits in the wall that let the wave through. (We'll see in Section 9.3 that by "very small," we mean that the height is much smaller than the wavelength.) We're assuming that the slits are essentially infinite in length in the direction perpendicular to the page. So they are very wide but very squat rectangles. Fig. 2 shows a head-on view from the far-away point source.

By Huygens' principle we can consider each slit to be the source of a *cylindrically* propagating wave. It is a cylindrical (and not spherical) wave because the wave has no dependence in the direction perpendicular to the page, due to the fact that it is generated by a line source (the slit). If we had a point source instead of a line source, then we would end up with a standard spherically propagating wave. The reason why we're using a line source is so that we can ignore the coordinate perpendicular to the page. However, having said this, the fact that we have a cylindrical wave instead of a spherical wave will be largely irrelevant in this chapter. The main difference is that the amplitude of a cylindrical wave falls off like $1/\sqrt{r}$ (see Section [to be added] in Chapter 7) instead of the usual $1/r$ for a spherical wave. But for reasons that we will see, we can usually ignore this dependence. In the end, since we're ignoring the coordinate perpendicular to the page, we can consider the setup to be a planer one (in the plane of the page) and effectively think of the line source as a point source (namely, the point on the line that lies in the page) that happens to produce a wave whose amplitude falls off like $1/\sqrt{r}$ (although this fact won't be very important).

The important thing to note about our setup is that the two sources are *in phase* due to the assumption that the wavefronts are parallel to the wall.[1] Note that instead of this setup with the incoming plane wave and the slits in a wall, we could of course simply have two actual sources that are in phase. But it is sometimes difficult to generate two waves that are exactly in phase. Our setup with the slits makes this automatically be the case.

As the two waves propagate outward from the slits, they will interfere. There will be constructive interference at places where the two waves are in phase (where the pathlengths from the two slits differ by an integral multiple of the wavelength). And there will be destructive interference at places where the two waves are 180° out of phase (where the pathlengths from the two slits differ by an odd multiple of half of the wavelength). For example, there is constructive interference at point A in Fig. 3 and destructive interference at point B.

What is the interference pattern on a screen that is located very far to the right of the wall? Assume that the screen is parallel to the wall. The setup is shown in Fig. 4. The distance between the slits is d, the distance to the screen is D, the lengths of the two paths to a given point P are r_1 and r_2, and θ is the angle that the line to P makes with the normal to the wall and screen. The distance x from P to the midpoint of the screen is then $x = D\tan\theta$.

[1]Problem 9.1 shows how things are modified if the wavefronts aren't parallel to the wall. This is done in the context of the N-slit setup in Section 9.2. The modification turns out to be a trivial one.

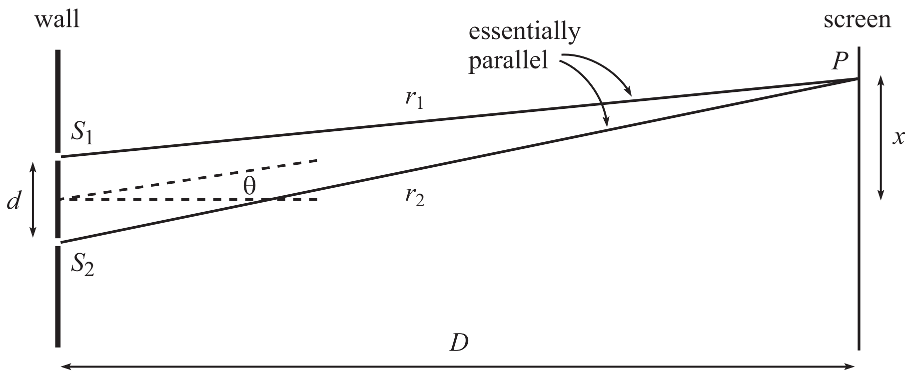


Figure 4

In finding the interference pattern on the screen, we will work in the so-called *far-field* limit where the screen is very far away. (We'll discuss the *near-field* case in Section 9.5.)[2] The quantitative definition of the far-field limit is $D \gg d$. This assumption that D is much larger than d leads to two important facts.

- If $D \gg d$, then we can say that the two pathlengths r_1 and r_2 in Fig. 4 are essentially equal in a *multiplicative* sense. That is, the ratio r_1/r_2 is essentially equal to 1. This follows from the fact that the additive difference $|r_1 - r_2|$ is negligible compared with r_1 and r_2 (because $|r_1 - r_2|$ can't be any larger than d, which we are assuming is negligible compared with D, which itself is less than r_1 and r_2). This $r_1/r_2 \approx 1$ fact then tells us that the amplitudes of the two waves at point P from the two slits are essentially equal (because the amplitudes are proportional to $1/\sqrt{r}$, although the exact power dependence here isn't important).

- If $D \gg d$, then we can say that the r_1 and r_2 paths in Fig. 4 are essentially parallel, and so they make essentially the same angle (namely θ) with the normal. The parallel nature of the paths then allows us to easily calculate the *additive* difference between the pathlengths. A closeup of Fig. 4 near the slits is shown in Fig. 5. The difference in the pathlengths is obtained by dropping the perpendicular line as shown, so we see that the difference $r_2 - r_1$ equals $d \sin\theta$. The phase difference between the two waves is then

$$k(r_2 - r_1) = kd\sin\theta = \frac{2\pi}{\lambda} d\sin\theta = 2\pi \cdot \frac{d\sin\theta}{\lambda}. \tag{1}$$

  In short, $d\sin\theta/\lambda$ is the fraction of a cycle that the longer path is ahead of the shorter path.

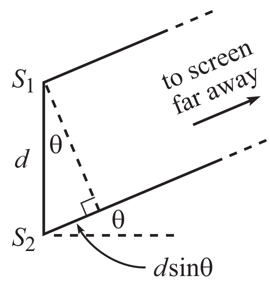


Figure 5

Remark: We found above that r_1 is essentially equal to r_2 in a *multiplicative* sense, but not in an *additive* sense. Let's be a little more explicit about this. Let ϵ be defined as the difference, $\epsilon \equiv r_2 - r_1$. Then $r_2 = r_1 + \epsilon$, and so $r_2/r_1 = 1 + \epsilon/r_1$. Since $r_1 > D$, the second term here is less than ϵ/D. As we mentioned above, this quantity is negligible because ϵ can't be larger than d, and because we're assuming $D \gg d$. We therefore conclude that $r_2/r_1 \approx 1$. In other words, $r_1 \approx r_2$ in a multiplicative sense. This then implies that the amplitudes of the two waves are essentially equal.

However, the phase difference equals $k(r_2 - r_1) = 2\pi(r_2 - r_1)/\lambda = 2\pi\epsilon/\lambda$. So if ϵ is of the same order as the wavelength, then the phase difference isn't negligible. So r_2 is *not* equal to r_1

[2] The fancier terminology for these two cases comes from the people who did pioneering work in them: the *Fraunhofer* limit for far-field, and the *Fresnel* limit for near-field. The correct pronunciation of "Fresnel" appears to be fray-NELL, although many people say freh-NELL.

in an *additive* sense. To sum up, the multiplicative comparison of r_2 and r_1 (which is relevant for the amplitudes) involves the comparison of ϵ and D, and we know that ϵ/D is negligible in the far-field limit. But the additive comparison of r_2 and r_1 (which is relevant for the phases) involves the comparison of ϵ and λ, and ϵ may very well be of the same order as λ. ♣

Having found the phase difference in Eq. (1), we can now find the total value of the wave at point P. Let A_P be the common amplitude of each of the two waves at P. Then up to an overall phase that depends on when we pick the $t = 0$ time, the total (complex) wave at P equals

$$\begin{aligned} E_{\text{tot}}(P) &= A_P e^{i(kr_1-\omega t)} + A_P e^{i(kr_2-\omega t)} \\ &= A_P(e^{ikr_1} + e^{ikr_2})e^{-i\omega t}. \end{aligned} \tag{2}$$

Our goal is to find the amplitude of the total wave, because that (or rather the square of it) yields the intensity of the total wave at point P. We can find the amplitude by factoring out the average of the two phases in the wave, as follows.

$$\begin{aligned} E_{\text{tot}}(P) &= A_P\Big(e^{ik(r_1-r_2)/2} + e^{-ik(r_1-r_2)/2}\Big)e^{ik(r_1+r_2)/2}e^{-i\omega t} \\ &= 2A_P\cos\left(\frac{k(r_1-r_2)}{2}\right)e^{i\left(k(r_1+r_2)/2-\omega t\right)} \\ &= 2A_P\cos\left(\frac{kd\sin\theta}{2}\right)e^{i\left(k(r_1+r_2)/2-\omega t\right)}, \end{aligned} \tag{3}$$

where we have used $k(r_2 - r_1) = kd\sin\theta$ from Eq. (1). The amplitude is the coefficient of the exponential term, so we see that the total amplitude at P is

$$A_{\text{tot}}(P) = 2A_P\cos\left(\frac{kd\sin\theta}{2}\right) \quad\longrightarrow\quad A_{\text{tot}}(\theta) = 2A(\theta)\cos\left(\frac{kd\sin\theta}{2}\right), \tag{4}$$

where we have rewritten A_P as $A(\theta)$, and $A_{\text{tot}}(P)$ as $A_{\text{tot}}(\theta)$, to emphasize the dependence on θ. Note that the amplitude at $\theta = 0$ is $2A(0)\cos(0) = 2A(0)$. Therefore,

$$\boxed{\frac{A_{\text{tot}}(\theta)}{A_{\text{tot}}(0)} = \frac{A(\theta)}{A(0)}\cos\left(\frac{kd\sin\theta}{2}\right)} \tag{5}$$

The intensity is proportional to the square of the amplitude, which gives

$$\boxed{\frac{I_{\text{tot}}(\theta)}{I_{\text{tot}}(0)} = \frac{A(\theta)^2}{A(0)^2}\cos^2\left(\frac{kd\sin\theta}{2}\right)} \tag{6}$$

Since the amplitude of a cylindrically propagating wave is proportional to $1/\sqrt{r}$, we have

$$\frac{A(\theta)}{A(0)} = \frac{1/\sqrt{r(\theta)}}{1/\sqrt{r(0)}} = \sqrt{\frac{r(0)}{r(\theta)}} = \sqrt{\frac{D}{D/\cos\theta}} = \sqrt{\cos\theta}. \tag{7}$$

Therefore,

$$\begin{aligned} \frac{I_{\text{tot}}(\theta)}{I_{\text{tot}}(0)} &= \cos\theta\cos^2\left(\frac{kd\sin\theta}{2}\right) \\ &= \cos\theta\cos^2\left(\frac{\pi d\sin\theta}{\lambda}\right). \end{aligned} \tag{8}$$

This result holds for all values of θ, even ones that approach 90°. The only approximation we've made so far is the far-field one, which allows us to say that (1) the amplitudes of the waves from the two slits are essentially equal, and (2) the two paths are essentially parallel. The far-field approximation has nothing to do with the angle θ.

If we want to write I_{tot} in terms of the distance x from the midpoint of the screen, instead of θ, then we can use $\cos\theta = D/\sqrt{x^2+D^2}$ and $\sin\theta = x/\sqrt{x^2+D^2}$. This gives

$$\begin{aligned}\frac{I_{\text{tot}}(x)}{I_{\text{tot}}(0)} &= \frac{D}{\sqrt{x^2+D^2}}\cos^2\left(\frac{xkd}{2\sqrt{x^2+D^2}}\right) \\ &= \frac{D}{\sqrt{x^2+D^2}}\cos^2\left(\frac{x\pi d}{\lambda\sqrt{x^2+D^2}}\right). \qquad (9)\end{aligned}$$

Plots of $I_{\text{tot}}(x)/I_{\text{tot}}(0)$ are shown in Fig. 6, for d values of $(.01)\lambda$, $(0.5)\lambda$, 5λ, and 50λ.

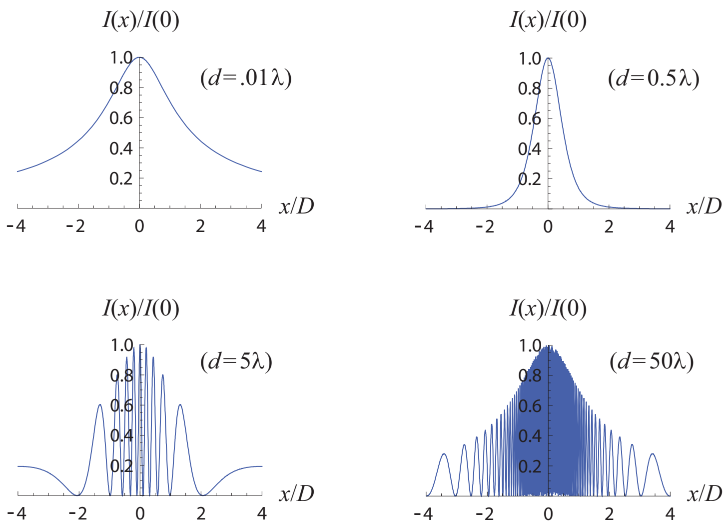


Figure 6

As you can see from the first plot, if d is much smaller than λ, the interference pattern isn't too exciting, because the two paths are essentially in phase with each other. The most they can be out of phase is when $\theta \to 90°$ (equivalently, $x \to \infty$), in which case the pathlength difference is simply $b = (.01)\lambda$, which is only 1% of a full phase. Since we have $d \ll \lambda$ in the first plot, the cosine-squared term in Eq. (9) is essentially equal to 1, so the curve reduces to a plot of the function $D/\sqrt{x^2+D^2}$. It decays to zero for the simple intuitive reason that the farther we get away from the slit, the smaller the amplitude is (more precisely, $A(\theta) = A(0)\sqrt{\cos\theta}$). In this $d \ll \lambda$ case, we effectively have a single light source from a single slit; interference from the two slits is irrelevant because the waves can never be much out of phase. The function in the first plot is simply the intensity we would see from a single slit.

The $d = (0.5)\lambda$ plot gives the cutoff case when there is barely destructive interference at $x = \infty$. (Of course, the amplitude of both waves is zero there, so the total intensity is zero anyway.) The $d = 5\lambda$ and $d = 50\lambda$ plots exhibit noticeable interference. The local maxima occur where the two pathlengths differ by an integral multiple of the wavelength. The local minima occur where the two pathlengths differ by an odd multiple of half of the wavelength. The $D/\sqrt{x^2+D^2}$ function in Eq. (9) is the envelope of the cosine-squared function. In the first plot, the $D/\sqrt{x^2+D^2}$ function is all there is, because the cosine-squared function

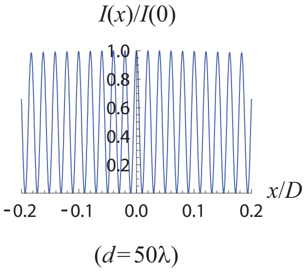


Figure 7

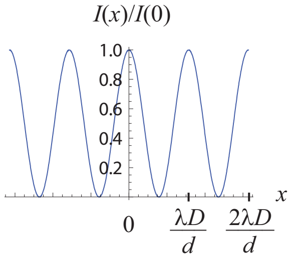


Figure 8

essentially never deviates from 1. But in the $d = 5\lambda$ and $d = 50\lambda$ cases, it actually goes through some cycles.

Fig. 7 shows a close-up version of the $d = 50\lambda$ case. For small x (equivalently, for small θ), the ratio $A(\theta)/A(0) = \sqrt{\cos\theta}$ is essentially equal to 1, so the envelope is essentially constant. We therefore simply have a cosine-squared function with a nearly-constant amplitude. In practice, we're usually concerned only with small x and θ values, in which case Eqs. (8) and (9) become

$$\boxed{\frac{I_{\text{tot}}(\theta)}{I_{\text{tot}}(0)} \approx \cos^2\left(\frac{\theta\pi d}{\lambda}\right)} \qquad (\text{for} \quad \theta \ll 1)$$
$$\boxed{\frac{I_{\text{tot}}(x)}{I_{\text{tot}}(0)} \approx \cos^2\left(\frac{x\pi d}{\lambda D}\right)} \qquad (\text{for} \quad x \ll D) \qquad (10)$$

For the remainder of this chapter, we will generally work in this small-angle approximation. So we won't need the exact (at least exact in the far-field approximation) results in Eqs. (8) and (9).

The plot of $I_{\text{tot}}(x)/I_{\text{tot}}(0)$ from Eq. (10) is shown in Fig. 8. The maxima occur at integer multiples of $\lambda D/d$. It makes sense that the spacing grows with λ, because the larger λ is, the more tilted the paths in Fig. 5 have to be to make the difference in their lengths (which is $d\sin\theta$) be a given multiple of λ. The approximations we've made in Fig. 8 are that we've ignored the facts that as we move away from the center of the screen, (a) the amplitude $A(\theta)$ of the two waves decreases, and (b) the peaks become spaced farther apart. You can compare Fig. 8 with the third and fourth plots in Fig. 6.

Remember that the small-angle approximation we've made here is different from the "far-field" approximation. The far-field approximation is the statement that the distances from the two slits to a *given* point P on the screen are essentially equal, multiplicatively. This holds if $d \ll D$. (We'll eventually drop this assumption in Section 9.5 when we discuss the near-field approximation.) The small-angle approximation that leads to Eq. (10) is the statement that the distances from the two slits to *different* points on the screen are all essentially equal. This holds if $x \ll D$, or equivalently $\theta \ll 1$. Note that the small-angle approximation has no chance of being valid unless the far-field approximation already holds.

Remark: The small-angle approximation in Eq. (10) shoves the $A(\theta)$ dependence in Eq. (6) under the rug. Another way to get rid of this dependance is to use a cylindrical screen instead of a flat screen, with the axis of the cylinder coinciding with the slits. So in Fig. 4 the screen would be represented by a semicircle in the plane of the page, with the slits located at the center. In the far-field limit, all of the paths in different θ directions now have the same length, multiplicatively. (The difference in pathlengths to a given point on the screen is still $d\sin\theta$.) So $A(\theta) = A(0)$ for all θ, and the A's cancel in Eq. (6). Note, however, that the spacing between the local maxima on the cylindrical screen still isn't uniform, because they occur where $\sin\theta = \lambda/d$. And $\sin\theta$ isn't a linear function of θ. At any rate, the reason why we generally work in terms of a flat screen isn't that there is anything fundamentally better about it compared with a cylindrical screen. It's just that in practice it's easier to find a flat screen. ♣

9.2 N-slit interference

9.2.1 Standard derivation

Let's now looks at the case where we have a general number, N, of equally-spaced slits, instead of 2. The setup is shown in Fig. 9 for the case of $N = 5$.

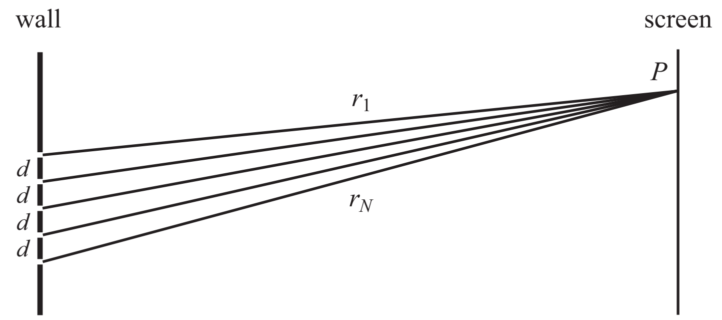


Figure 9

Similar to the $N = 2$ case above, we will make the far-field assumption that the distance to the screen is much larger than the total span of the slits, which is $(N-1)d$. We can then say, as we did in the $N = 2$ case, that all the paths to a given point P on the screen have essentially the same length in a multiplicative (but not additive) sense, which implies that the amplitudes of the waves are all essentially equal. And we can also say that all the paths are essentially parallel. A closeup version near the slits is shown in Fig. 10. Additively, each pathlength is $d \sin\theta$ longer than the one right above it. So the lengths take the form of $r_n = r_1 + (n-1)d\sin\theta$.

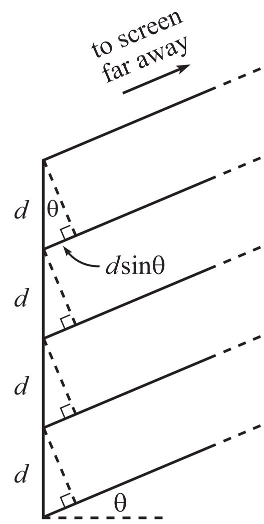


Figure 10

To find the total wave at a given point at an angle θ on the screen, we need to add up the N individual waves (call them E_n). The procedure is the same as in the $N = 2$ case, except that now we simply have more terms in the sum. In the $N = 2$ case we factored out the average of the phases (see Eq. (3)), but it will now be more convenient to factor out the phase of the top wave in Fig. 9 (the r_1 path). The total wave at an angle θ on the screen is then (with $A(\theta)$ being the common amplitude of all the waves)

$$\begin{aligned} E_{\text{tot}}(\theta) = \sum_{n=1}^{N} E_n &= \sum_{n=1}^{N} A(\theta) e^{i(kr_n - \omega t)} \\ &= A(\theta) e^{i(kr_1 - \omega t)} \sum_{n=1}^{N} e^{ik(n-1)d\sin\theta}. \end{aligned} \tag{11}$$

With $z \equiv e^{ikd\sin\theta}$, the sum here is $1 + z + z^2 + \cdots + z^{N-1}$. The sum of this geometric series is

$$\begin{aligned} \frac{z^N - 1}{z - 1} &= \frac{e^{ikNd\sin\theta} - 1}{e^{ikd\sin\theta} - 1} \\ &= \frac{e^{ik(N/2)d\sin\theta}}{e^{ik(1/2)d\sin\theta}} \cdot \frac{e^{ik(N/2)d\sin\theta} - e^{-ik(N/2)d\sin\theta}}{e^{ik(1/2)d\sin\theta} - e^{-ik(1/2)d\sin\theta}} \\ &= e^{ik((N-1)/2)d\sin\theta} \cdot \frac{\sin\left(\frac{1}{2}Nkd\sin\theta\right)}{\sin\left(\frac{1}{2}kd\sin\theta\right)}. \end{aligned} \tag{12}$$

Substituting this into Eq. (11) yields a total wave of

$$E_{\text{tot}}(\theta) = A(\theta)\frac{\sin\left(\frac{1}{2}Nkd\sin\theta\right)}{\sin\left(\frac{1}{2}kd\sin\theta\right)}\left(e^{i(kr_1-\omega t)}e^{ik((N-1)/2)d\sin\theta}\right). \tag{13}$$

The amplitude is the coefficient of the exponential factors, so we have

$$\boxed{A_{\text{tot}}(\theta) = A(\theta)\frac{\sin\left(\frac{1}{2}Nkd\sin\theta\right)}{\sin\left(\frac{1}{2}kd\sin\theta\right)} \equiv A(\theta)\frac{\sin(N\alpha/2)}{\sin(\alpha/2)}} \tag{14}$$

where

$$\alpha \equiv kd\sin\theta = \frac{2\pi d\sin\theta}{\lambda}\,. \tag{15}$$

Since adjacent pathlengths differ by $d\sin\theta$, the physical interpretation of α is that it is the phase difference between adjacent paths.

What is the value of $A_{\text{tot}}(\theta)$ at the midpoint of the screen where $\theta = 0$ (which implies $\alpha = 0$)? At $\alpha = 0$, Eq. (14) yields $A_{\text{tot}}(\theta) = 0/0$, which doesn't tell us much. But we can obtain the actual value by taking the limit of small α. Using $\sin\epsilon \approx \epsilon$, we have

$$A_{\text{tot}}(0) = \lim_{\theta\to 0} A_{\text{tot}}(\theta) = \lim_{\alpha\to 0} A(\theta)\frac{\sin(N\alpha/2)}{\sin(\alpha/2)} = A(0)\frac{N\alpha/2}{\alpha/2} = A(0)\cdot N. \tag{16}$$

It is customary to deal not with the amplitude itself, but rather with the amplitude relative to the amplitude at $\theta = 0$. Combining Eqs. (14) and (16) gives

$$\frac{A_{\text{tot}}(\theta)}{A_{\text{tot}}(0)} = \frac{A(\theta)}{A(0)}\cdot\frac{\sin(N\alpha/2)}{N\sin(\alpha/2)}\,. \tag{17}$$

Since we generally deal with small angles, we'll ignore the variation in the $A(\theta)$ coefficient. In other words, we'll set $A(\theta) \approx A(0)$. This gives

$$\boxed{\frac{A_{\text{tot}}(\alpha)}{A_{\text{tot}}(0)} \approx \frac{\sin(N\alpha/2)}{N\sin(\alpha/2)}} \qquad \text{(for small } \theta) \tag{18}$$

The intensity at θ relative to the intensity at $\theta = 0$ is then

$$\boxed{\frac{I_{\text{tot}}(\alpha)}{I_{\text{tot}}(0)} \approx \left(\frac{\sin(N\alpha/2)}{N\sin(\alpha/2)}\right)^2} \qquad \text{(for small } \theta) \tag{19}$$

Even for large angles, the effect of $A(\theta)$ is to simply act as an envelope function of the oscillating sine functions. We can always bring $A(\theta)$ back in if we want to, but the more interesting behavior of $A_{\text{tot}}(\theta)$ is the oscillatory part. We're generally concerned with the *locations* of the maxima and minima of the oscillations and not with the actual value of the amplitude. The $A(\theta)$ factor doesn't affect these locations.[3] We'll draw a plot of what $I_{\text{tot}}(\alpha)/I_{\text{tot}}(0)$ looks like, but first a remark.

REMARK: Technically, we're being inconsistent here in our small-angle approximation, because although we set $A(\theta) = A(0)$ (which from Eq. (7) is equivalent to setting $\cos\theta = 1$), we didn't set $\sin\theta = \theta$ in the expression for α. To be consistent, we should approximate $\alpha = kd\sin\theta$ by $\alpha = kd\cdot\theta$. The reason why we haven't made this approximation is that we want to keep the *locations* of the bumps in the interference pattern correct, even for large θ. And besides, the function $A(\theta)$ depends on the nature of the screen. A flat screen has $A(\theta)/A(0) = \sqrt{\cos\theta}$, which decreases with θ, while a cylindrical screen has $A(\theta)/A(0) = 1$, which is constant. Other shapes yield other functions of θ. But they're all generally slowly-varying functions of θ, compared with the oscillations of the $\sin(N\alpha/2)$ function (unless you use a crazily-shaped wiggly screen, which you have no good reason to do). The main point is that the function $A(\theta)$ isn't an inherent property of the interference pattern; it's a property of the screen. On the other hand, the angular locations of the maxima and minima of the oscillations *are* an inherent property of the pattern. So it makes sense to keep these locations exact and not lose this information when making the small-angle approximation. If you want, you can write the intensity in Eq. (19) as

$$\frac{I_{\text{tot}}(\theta)}{I_{\text{tot}}(0)} = F(\theta)\left(\frac{\sin(N\alpha/2)}{N\sin(\alpha/2)}\right)^2 \qquad (\text{where } \alpha \equiv kd\sin\theta), \tag{20}$$

[3] Strictly speaking, $A(\theta)$ does affect the locations of the maxima in a very slight manner (because when taking the overall derivative, the derivative of $A(\theta)$ comes into play). But $A(\theta)$ doesn't affect the locations of the minima, because those are where $I_{\text{tot}}(\alpha)$ is zero.

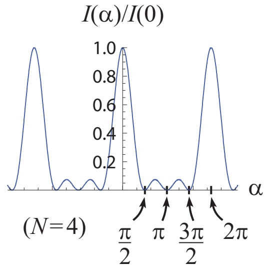


Figure 11

where $F(\theta)$ is a slowly-varying function of θ that depends on the shape of the screen. We will generally ignore this dependence and set $F(\theta) = 1$. ♣

What does the $I_{\text{tot}}(\alpha)/I_{\text{tot}}(0)$ ratio in Eq. (19) look like as a function of α? The plot for $N = 4$ is shown in Fig. 11. If we're actually talking about small angles, then we have $\alpha = kd \sin\theta \approx kd \cdot \theta$. But the distance from the center of the screen is $x = D \tan\theta \approx D \cdot \theta$. So for small angles, we have $\alpha \propto x$. You can therefore think of Fig. 11 as showing the actual intensity on the screen as a function of x (up to a scaling constant).

Note that although we generally assume θ to be small, α is *not* necessarily small, because $\alpha \equiv kd \sin\theta$ involves a factor of k which may be large. Said in another way, α is the phase difference between adjacent slits, so if k is large (more precisely, if $\lambda \ll d$), then even a small angle θ can lead to a pathlength difference (which is $d \sin\theta$) equal to λ. This corresponds to a phase difference of $\alpha = kd \sin\theta = (2\pi/\lambda) d \sin\theta = 2\pi$. Consistent with this, the values on the horizontal axis in Fig. 11 are on the order of π (that is, they are not small), and the first side peak is located at 2π.

A number of things are evident from both Fig. 11 and Eq. (19):

1. The value of $I_{\text{tot}}(\alpha)/I_{\text{tot}}(0)$ at $\theta = 0$ is 1, by construction.

2. $I_{\text{tot}}(\alpha)/I_{\text{tot}}(0)$ has a period of 2π in α. The sine function in the denominator picks up a minus sign when α increases by 2π, and likewise in the numerator if N is odd. But an overall minus sign is irrelevant because the intensity involves the squares of the sines.

3. $I_{\text{tot}}(\alpha)$ has zeroes whenever $N\alpha/2$ is a multiple of π, that is, whenever $N\alpha/2 = m\pi \implies \alpha = 2m\pi/N$, which means that α is an even multiple of π/N. The one exception to this is when $\alpha/2$ is also a multiple of π, that is, when $\alpha/2 = m'\pi \implies \alpha = 2m'\pi$, because then the denominator in Eq. (19) is also zero. (In this case, Eq. (16) tells us that the value at $\theta = 0$ is 1. And likewise at any integer multiple of 2π. These are the locations of the main peaks.) In the $N = 4$ case in Fig. 11, you can see that the zeros do indeed occur at

$$\frac{0\cancel{\pi}}{\cancel{4}}, \quad \frac{2\pi}{4}, \quad \frac{4\pi}{4}, \quad \frac{6\pi}{4}, \quad \frac{8\cancel{\pi}}{\cancel{4}}, \quad \frac{10\pi}{4}, \quad \frac{12\pi}{4}, \quad \frac{14\pi}{4}, \quad \frac{16\cancel{\pi}}{\cancel{4}}, \quad \ldots. \tag{21}$$

   And likewise for negative values. In general, the number of zeros between the main peaks is $N - 1$.

4. If you take the derivative of $I_{\text{tot}}(\alpha)$, you will find that the local maxima (of the small bumps) occur when $\tan(N\alpha/2) = N\tan(\alpha/2)$. This has to be solved numerically. However, for large N, the solutions for α are generally very close to the odd multiples of π/N (except for values of the form of $2\pi \pm \pi/N$); see Problem [to be added]. In other words, the local maxima are approximately right between the local minima (the zeros) which themselves occur exactly at the even multiples of π/N, except at the integral multiples of 2π where the main peaks are. In Fig. 11 you can see that the small bumps do indeed occur at approximately

$$\frac{1\cancel{\pi}}{\cancel{4}}, \quad \frac{3\pi}{4}, \quad \frac{5\pi}{4}, \quad \frac{7\cancel{\pi}}{\cancel{4}}, \quad \frac{9\cancel{\pi}}{\cancel{4}}, \quad \frac{11\pi}{4}, \quad \frac{13\pi}{4}, \quad \frac{15\cancel{\pi}}{\cancel{4}}, \quad \frac{17\cancel{\pi}}{\cancel{4}}, \quad \frac{19\pi}{4}, \quad \ldots. \tag{22}$$

   And likewise for negative values. In general, the number of little bumps between the main peaks is $N - 2$.

5. The little bumps in Fig. 11 have the same height, simply because there are only two of them. For larger values of N, the bump sizes are symmetric around $\alpha = \pi$ (or in

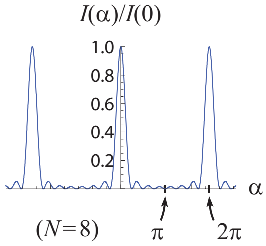


Figure 12

general any multiple of π). They are the shortest there (because the denominator in Eq. (19) is largest at $\alpha = \pi$), and they grow in size as they get closer to the main peaks. Fig. 12 shows the interference pattern for $N = 8$.

Note that if $d < \lambda$, then $\alpha \equiv kd\sin\theta = (2\pi/\lambda)d\sin\theta < 2\pi\sin\theta \leq 2\pi$. So α can't achieve the value of 2π, which means that none of the tall side peaks in Fig. 11 exist. We have only one tall peak at $\alpha = 0$, and then a number of small peaks. This makes sense physically, because the main peaks occur when the waves from all the slits are in phase. And if $d < \lambda$ there is no way for the pathlengths to differ by λ, because the difference can be at most d (which occurs at $\theta = 90°$). In general, the upper limit on α is kd, because $\sin\theta$ can't exceed 1. So no matter what the relation between d and λ is, a plot such as the one in Fig. 12 exists out to the $\alpha = kd$ point (which corresponds to $\theta = 90°$), and then it stops.

In the case of $N = 2$, we should check that the expression for $I_{\text{tot}}(\alpha)/I_{\text{tot}}(0)$ in Eq. (19) reduces properly to the expression in Eq. (6) (with $A(\theta)$ set equal to $A(0)$). Indeed, if $N = 2$, then the quotient in Eq. (19) becomes $\sin(2\cdot\alpha/2)/2\sin(\alpha/2)$. Using the double-angle formula in the numerator turns this into $\cos(\alpha/2) = \cos\big((1/2)kd\sin\theta\big)$, which agrees with Eq. (6).

9.2.2 Geometric construction

Let's now derive the amplitude in Eq. (14) in a different way. It turns out that there is an extremely informative geometric way of seeing how this amplitude arises. The main task in finding the amplitude is calculating the sum in Eq. (11). With $\alpha \equiv kd\sin\theta$, this sum is

$$\sum_{n=1}^{N} e^{ik(n-1)d\sin\theta} = 1 + e^{i\alpha} + e^{i2\alpha} + \cdots + e^{i(N-1)\alpha}. \tag{23}$$

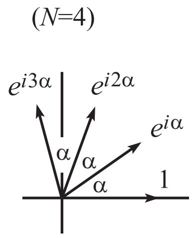


Figure 13

Each term in this sum is a complex number with magnitude 1. If we plot these numbers as vectors in the complex plane, they make angles of 0, α, 2α, etc. with respect to the x axis. For example, in the case of $N = 4$ we might have the unit vectors shown in Fig. 13. (Remember that α depends on θ, which depends on where the point P is on the screen. So for any point P, we have a set of N vectors in the plane. The angle α between them increases as P moves farther off to the side.) The easiest way to add up these vectors is to put them tail-to-head, as shown in Fig. 14. Each of the unit vectors is tilted at an angle α with respect to the one before it. The desired sum is the thick vector shown. As with any complex number, we can write this sum as a magnitude times a phase, that is, as $Re^{i\phi}$.

The total amplitude $A_{\text{tot}}(\theta)$ equals R times the $A(\theta)$ in Eq. (11), because the phase $e^{i(kr_1-\omega t)}$ in Eq. (11) and the phase $e^{i\phi}$ in the sum don't affect the amplitude. So our goal is to find R, which can be done in the following way. (If you want to find the value of ϕ, see Problem [to be added].)

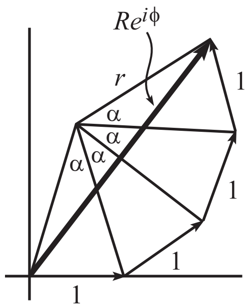


Figure 14

The thick vector in Fig. 14 is the base of an isosceles triangle with vertex angle 4α, which in general is $N\alpha$. So we have

$$R = 2 \cdot r\sin(N\alpha/2), \tag{24}$$

where r is the length shown in the figure. But from looking at any one of the four thinner isosceles triangles with vertex angle α and base 1, we have

$$1 = 2 \cdot r\sin(\alpha/2). \tag{25}$$

Taking the quotient of the two preceding equations eliminates the length r, and we arrive at

$$R = \frac{\sin(N\alpha/2)}{\sin(\alpha/2)}\,. \tag{26}$$

This reproduces Eq. (14), because the total amplitude $A_{\text{tot}}(\theta)$ equals $R \cdot A(\theta)$.

As time increases, the whole picture in Fig. 14 rotates clockwise in the plane, due to the $-\omega t$ in the phase. There is also a phase shift due to the $k_1 r$ and ϕ terms in the phase, but this simply affects the starting angle. Since all the little vectors keep their same relative orientation, the figure keeps its same shape. That is, it rotates as a rigid "object." The sum (the thick vector) therefore always has the same length. This (constant) length is therefore the amplitude, while the (changing) horizontal component is the real part that as usual gives the actual physical wave.

The above geometric construction makes it easy to see why the main peaks and all the various local maxima and minima appear in Fig. 12. The main peaks occur when α is a multiple of 2π, because then all the little vectors point in the same direction (rightward at a given instant, if the first little vector points to the right at that instant). The physical reason for this is that $\alpha = m \cdot 2\pi$ implies that

$$kd\sin\theta = 2m\pi \implies \frac{2\pi d\sin\theta}{\lambda} = 2m\pi \implies d\sin\theta = m\lambda. \tag{27}$$

This says that the difference in pathlengths from adjacent slits is a multiple of the wavelength, which in turn says that the waves from all of the slits constructively interfere. Hence the maximal amplitude.

A local minimum (a zero) occurs if the value of α is such that the chain of little vectors in Fig. 14 forms a closed regular polygon (possibly wrapped around multiple times). In this case the sum (the thick vector in Fig. 14) has no length, so the amplitude is zero. The "polygons" for the seven zeros in the $N = 8$ case in Fig. 12 are shown in Fig. 15. We've taken the first of the vectors to always point horizontally to the right, although this isn't necessary. We've drawn the figures slightly off from the case where the sum of the vectors is zero, to make it easier to see what's going on. The last three figures are mirror images of the first three.

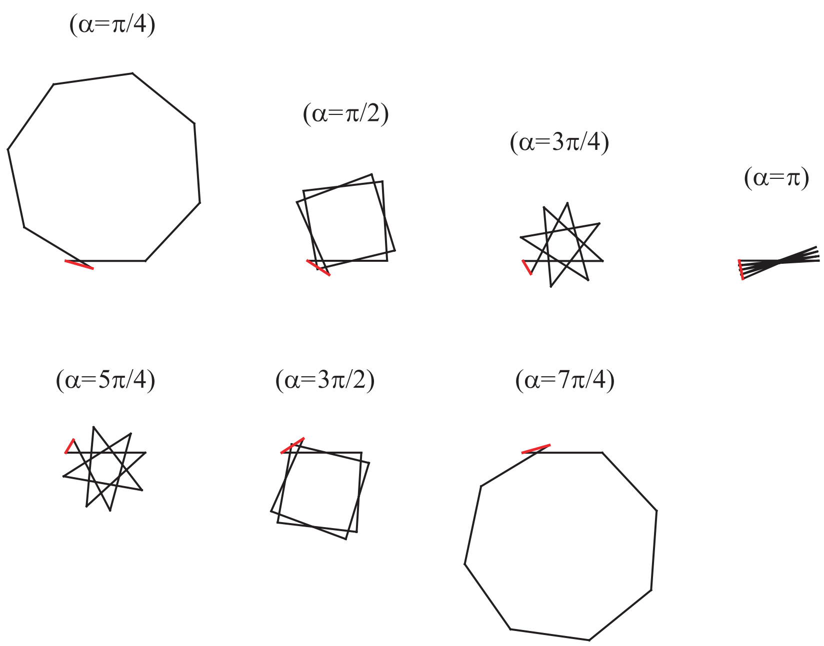


Figure 15

The local maxima occur between the local minima. In the case of large N, it's easy to determine the approximate locations of these maxima. For large N, the vectors form an essentially smooth curve, and the maxima occur roughly when the amplitude is a diameter of a circle. The first few of these occurrences are shown in Fig. 16 for the case of $N = 50$. We've made the curve spiral slightly inward so that you can see how many times it wraps around. But in reality (in the far-field limit), the curve just keeps tracing over itself.

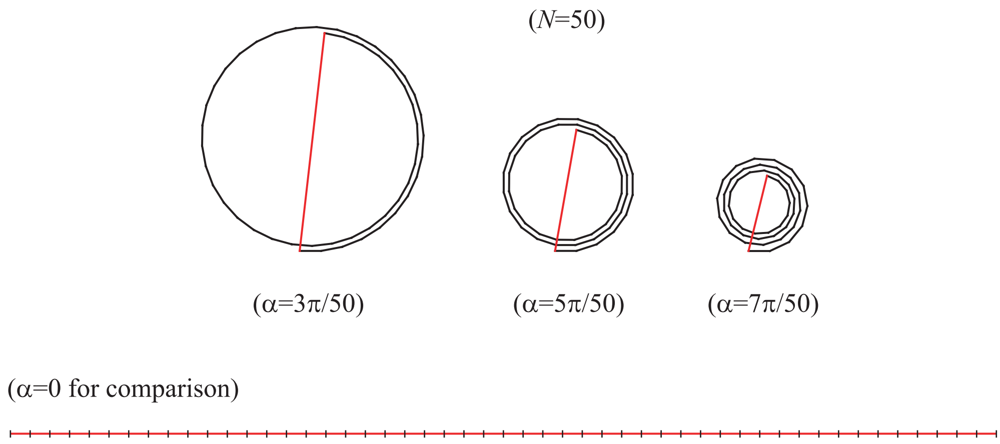


Figure 16

The maxima don't occur exactly at the diameters, because the circle shrinks as the little vectors wrap around further as α increases, so there are competing effects. But it is essentially the diameter if the wrapping number is large (because in this case the circle hardly changes size as the amplitude line swings past the diameter, so the shrinking effect is basically nonexistent). So we want the vectors to wrap (roughly) 3/2, 5/2, 7/2, etc. times around the circle. Since each full circle is worth 2π radians, this implies that the total angle, which is $N\alpha$, equals 3π, 5π, 7π, etc. In other words, α is an odd multiple of π/N, excluding π/N itself (and also excluding the other multiples adjacent to multiples of 2π). This agrees with the result in the paragraph preceding Eq. (22). The amplitude of the main peaks that occur when α equals zero or a multiple of 2π is also shown in Fig. 16 for comparison. In this case the circular curve is unwrapped and forms a straight line. The little tick marks indicate the $N = 50$ little vectors.

9.2.3 Diffraction gratings

A *diffraction grating* is a series of a large number, N, of slits with a very small spacing d between them. If a source emits light that consists of various different wavelengths, a diffraction grating provides an extremely simple method for determining what these wavelengths are.

As we saw above, the N-slit interference pattern consists of the main peaks, plus many smaller peaks in between. However, we will be concerned here only with the main peaks, because these completely dominate the smaller peaks, assuming N is large. Let's justify this statement rigorously and discuss a few other things, and then we'll do an example.

From the discussion that led to Eq. (22), the smaller peaks occur when α takes on values that are approximately the odd multiples of π/N (except for values of the form $2\pi \pm \pi/N$), that is, when α equals $3\pi/N$, $5\pi/N$, etc. The corresponding values of $I_{\text{tot}}(\alpha)/I_{\text{tot}}(0)$ are obtained from Eq. (19). The numerator equals $(\pm 1)^2$, and since N is large we can use a small-angle approximation in the denominator, which turns the denominator into $N\alpha/2$. The resulting values of $I_{\text{tot}}(\alpha)/I_{\text{tot}}(0)$ are then $(2/3\pi)^2$, $(2/5\pi)^2$, etc. Note that these are independent of N.

The first of the side peaks isn't negligible compared with the main peak (it's about $(2/3\pi)^2 = 4.5\%$ as tall). But by the 10th peak, the height is negligible (it's about 0.1% as tall). However, even though the first few side peaks aren't negligible, they are squashed very close to the main peaks if N is large. This follows from the fact that the spacing between the main peaks is $\Delta\alpha = 2\pi$, whereas the side peaks are on the order of π/N away from the main peaks. The figure for $N = 20$ is shown in Fig. 17. We can therefore make the approximation that the interference pattern is non-negligible only at (or extremely close to) the main peaks where α is a multiple of 2π.

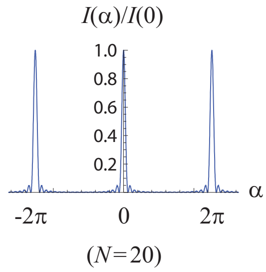


Figure 17

When dealing with diffraction gratings, we're generally concerned only with the location of the bright spots in the interference pattern, and not with the actual intensity. So any extra intensity from the little side peaks is largely irrelevant. And since they're squashed so close to the main peaks, it's impossible to tell that they're distinct bumps anyway. The location of the main peaks tells us what the various wavelengths are, by using $kd\sin\theta = 2\pi \Longrightarrow (2\pi/\lambda)d\sin\theta = 2\pi \Longrightarrow \lambda = d\sin\theta$. The intensity tells us how much of each wavelength the light is made of, but for most purposes we're not so concerned about this.

Remarks:

1. A diffraction grating should more appropriately be called an "interference grating," because it is simply an example of N-slit interference. It is *not* an example of diffraction, which we will define and discuss in Section 9.3.1. We'll see there that a feature of a diffraction pattern is that there are no tall side peaks, whereas these tall side peaks are the whole point of an "interference grating." However, we'll still use the term "diffraction grating" here, since this is the generally accepted terminology.

2. If we view the interference pattern on a screen, we know that it will look basically like Fig. 17 (we'll assume for now that only one wavelength is involved). However, if you put your eye right behind the grating, very close to it, what do you see? If you look straight at the light source, then you of course see the source. But if you look off at an angle (but still through the grating; so your eye has to be close to it), then you will also see a bright spot there. And you will also see bright spots at other angles. The number of spots depends on the relation between the wavelength and the spacing. We'll discuss a concrete case in the example below.

   The angles at which you see the spots are the same as the angles of the main peaks in Fig. 17, for the following reason. Fig. 18 shows the typical locations of the first few main peaks in the interference pattern from a standard set of slits contained in a small span in a wall. Imagine putting additional sets of slits in the wall at locations such that a given spot on the screen (your eye) is located at the angles of successive off-center peaks. This scenario is shown in Fig. 19. Each set of slits also produces many other peaks, of course, but you don't see them because your eye is at only one location.

   A diffraction grating is a continuous set of slits, but most of the slits are irrelevant. The only slits that matter are the ones that are located at positions such that the angle to your eye is one of the main-peak angles. In other words, we can replace the entire wall in Fig. 19 with a continuous set of slits, and you will still see the same thing. Only the small regions of slits shown in the figure will produce bright spots. In short, a diffraction grating acts like a collection of interference setups at specific locations in the grating.

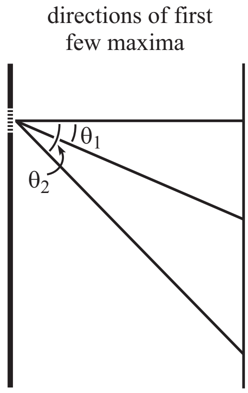


Figure 18

3. You might be concerned that if your eye is close enough to the grating, then the far-field approximation (and hence all of the result so far in this chapter) will be invalid. After all, the distance D from your eye to the grating isn't large compared with the total span of the slits in the grating. However, the far-field approximation does indeed still hold, because from the previous remark we're not concerned with the total span of the slits in the grating, but rather with the span of a small region near each of the main-peak angles. Assuming that the spacing between the lines in the grating is very small (it's generally on the order of 10^{-6} m), the span

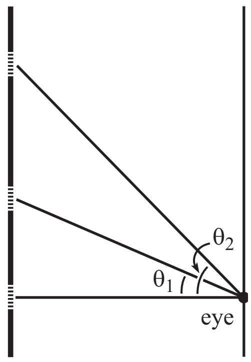


Figure 19

of a few hundred lines will still be very small compared with the distance from your eye to the grating (assuming that your eyelashes don't touch it). So the far-field approximation still holds. That is, the distances from these few hundred slits to your eye are all essentially equal (multiplicatively).

4. In reality, most diffraction gratings are made by etching regularly-spaced lines into the material. The exact details of the slits/etchings are't critical. Any periodic structure with period d will do the trick. The actual intensities depend on the details, but the locations of the main peaks don't. This follows from the usual argument that if $d \sin\theta$ is a multiple of λ, then there is constructive interference from the slits (whatever they may look like).
5. Problem 9.1 shows that it doesn't matter whether or not the incident light is normal to the wall (which is the diffraction grating here), as long as the deviation angle is small. If we measure all angles relative to the incident angle, then all of our previous result still hold. This is fortunate, of course, because if you hold a diffraction grating in front of your eye, it is highly unlikely that you will be able orient it at exactly a 90° angle to the line between it and the light source. ♣

Example (Blue and red light): A diffraction grating has 5000 lines per cm. Consider a white-light source (that is, it includes all wavelengths), and assume that it is essentially a point source far away. Taking the wavelengths of blue and red light to be roughly $4.5 \cdot 10^{-5}$ cm and $7 \cdot 10^{-5}$ cm, find the angles at which you have to look to the side to see the off-center blue and red maxima. What is the total number of maxima for each color that you can theoretically see on each side of the light source?

Solution: We basically have to do the same problem here twice, once for blue light and once for red light. As usual, the main peaks occur where the difference in pathlengths from adjacent slits is an integral multiple of the wavelength. So we want $d \sin\theta = m\lambda$. (Equivalently, we want $\alpha = m \cdot 2\pi$, which reduces to $d \sin\theta = m\lambda$.) We therefore want $\sin\theta = m\lambda/d$, where $d = (1\,\text{cm})/5000 = 2 \cdot 10^{-4}\,\text{cm}$.

For blue light, this gives $\sin\theta = m(4.5 \cdot 10^{-5}\,\text{cm})/(2 \cdot 10^{-4}\,\text{cm}) = m(0.225)$. So we have the following four possible pairs of m and θ values:

$$(m, \theta): \quad (1,\ 13.0^\circ) \quad (2,\ 26.7^\circ) \quad (3,\ 42.5^\circ) \quad (4,\ 64.2^\circ) \tag{28}$$

There are only four possible angles (plus their negatives), because $m = 5$ gives a value of $\sin\theta$ that is larger than 1.

For red light, we have $\sin\theta = m(7 \cdot 10^{-5}\,\text{cm})/(2 \cdot 10^{-4}\,\text{cm}) = m(0.35)$. So we have the following two possible pairs of m and θ values:

$$(m, \theta): \quad (1,\ 20.5^\circ) \quad (2,\ 44.4^\circ) \tag{29}$$

There are only two possible angles (plus their negatives), because $m = 3$ gives a value of $\sin\theta$ that is larger than 1. The red angles are larger than the corresponding blue angles because the red wavelength is longer, so it takes a larger angle to make adjacent pathlengths differ by a wavelength (or two wavelengths, etc.).

The rest of the spectrum falls between blue and red, so we obtain rainbow bands of colors. Note, however, that the first band (from 13.0° to 20.5°, although the endpoints are fuzzy) is the only "clean" band that doesn't overlap with another one. The second band ends at 44.4°, which is after the third band starts at 42.5°. And the third band doesn't even finish by the time the angle hits 90°. Your viewing angle has to be less than 90°, of course, because you have to be looking at least a little bit toward the grating. The angles of the various bands are shown in Fig. 20. The mirror images of these angles on the left side work too.

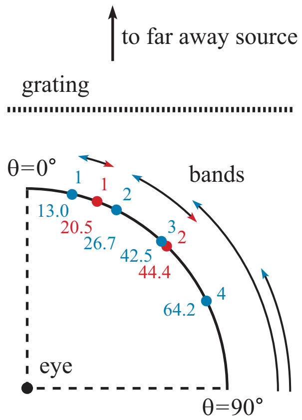


Figure 20

9.3 Diffraction from a wide slit

9.3.1 Derivation

We'll now discuss what happens when a plane wave impinges on just one wide slit with width a, instead of a number of infinitesimally thin ones. See Fig. 21. We'll find that if the width a isn't negligible compared with the wavelength λ, then something interesting happens. The interference pattern will depend on a in a particular way, whereas it didn't depend on the infinitesimal width in the previous sections. We'll keep working in the far-field limit, which here means that $D \gg a$.

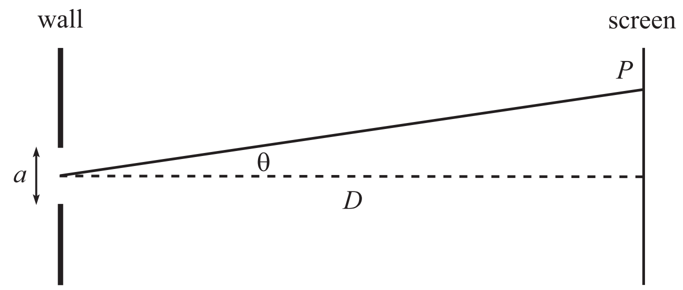


Figure 21

By Huygen's principle, we can consider the wide slit to consist of an infinite number of line sources (or point sources, if we ignore the direction perpendicular to the page) next to each other, each creating a cylindrical wave. In other words, the diffraction pattern from one continuous wide slit is equivalent to the $N \to \infty$ limit of the N-slit result in Eq. (19). So we've already done most of the work we need to do. We'll present three ways we can go about taking the continuum limit. But first some terminology.

The word *diffraction* refers to a situation with a continuous aperture. The word *interference* refers to a situation involving two or more apertures whose waves interfere. On one hand, since diffraction is simply the $N \to \infty$ limit of interference, there is technically no need to introduce a new term for it. But on the other hand, a specific kind of pattern arises, so it makes sense to give it its own name. Of course, we can combine interference and diffraction by constructing a setup with waves coming from a number of *wide* apertures. We'll deal with this in Section 9.4. A name that causes confusion between the words "interference" and "diffraction" is the *diffraction grating* that we discussed above. As we mentioned in the first remark in Section 9.2.3, this should technically be called an interference grating.

$N \to \infty$ **limit**

For our first derivation of the diffraction pattern, we'll take the $N \to \infty$ limit of Eq. (19). The α in Eq. (19) equals $kd\sin\theta$. But if we imagine the slit of width a to consist of N infinitesimal slits separated by a distance $d = a/N$, then we have $\alpha = k(a/N)\sin\theta$.[4] (The N here should perhaps be $N-1$, depending on where you put the slits, but this is irrelevant

[4]Of course, if we actually have infinitesimal slits separated by little pieces of wall, then the intensity will go down. But this doesn't matter since our goal is only to find the relative intensity $I_{\text{tot}}(\theta)/I_{\text{tot}}(0)$. As we'll see below, if the distance $d = a/N$ is much smaller than the wavelength (which it is, in the $N \to \infty$ limit) then we actually don't even need to have little pieces of wall separating the slits. The slits can bump right up against each other.

in the $N \to \infty$ limit.) Plugging this value of α into Eq. (19) gives

$$\frac{I_{\text{tot}}(\theta)}{I_{\text{tot}}(0)} \approx \left(\frac{\sin\left(\frac{ka\sin\theta}{2}\right)}{N\sin\left(\frac{ka\sin\theta}{2N}\right)} \right)^2 . \tag{30}$$

In the $N \to \infty$ limit, we can use $\sin\epsilon \approx \epsilon$ in the denominator to obtain

$$\frac{I_{\text{tot}}(\theta)}{I_{\text{tot}}(0)} \approx \left(\frac{\sin\left(\frac{1}{2}ka\sin\theta\right)}{\frac{1}{2}ka\sin\theta} \right)^2 \quad \Longrightarrow \quad \boxed{\frac{I_{\text{tot}}(\theta)}{I_{\text{tot}}(0)} \approx \left(\frac{\sin(\beta/2)}{\beta/2} \right)^2} \tag{31}$$

where

$$\beta \equiv ka\sin\theta = \frac{2\pi a\sin\theta}{\lambda} . \tag{32}$$

Another convention is to define $\beta \equiv (1/2)ka\sin\theta$, in which case the result in Eq. (31) takes the simpler form of $\big((\sin\beta)/\beta\big)^2$. The reason why we chose $\beta \equiv ka\sin\theta$ here is because it parallels the definition of α in Eq. (15). The results in Eqs. (19) and (31) are then similar, in that they both involve factors of 2. The physical meaning of α in Eq. (15) is that it is the phase difference between adjacent paths. The physical meaning of β is that is it the phase difference between the paths associated with the endpoints of the wide slit of width a.

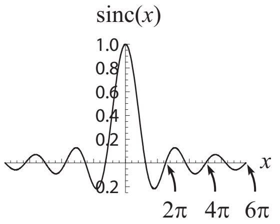


Figure 22

The function $(\sin x)/x$ is knows as the "sinc" function, $\text{sinc}(x) \equiv (\sin x)/x$. A plot is shown in Fig. 22. It is a sine function with a $1/x$ envelope. The result in Eq. (31) can therefore be written as $I_{\text{tot}}(\theta)/I_{\text{tot}}(0) = \text{sinc}^2(\beta/2)$. A plot of this is shown in Fig. 23. The factor of 2 in the argument makes the plot expanded by a factor of 2 in the horizontal direction compared with the plot in Fig. 22. Since $\sin\theta$ can't exceed 1, β can't exceed ka. So the plot in Fig. 23 exists out to the $\beta = ka$ point (which corresponds to $\theta = 90°$), and then it stops.

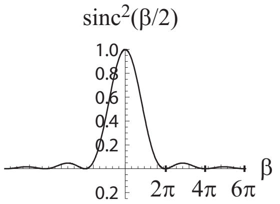


Figure 23

Note that the diffraction pattern has only one tall bump, whereas the interference patterns we've seen generally have more than one tall bump (assuming that $d > \lambda$). This is consistent with the discussion in the second-to-last paragraph in Section 9.2.1. We saw there that if $d < \lambda$, then there is only one tall bump. And indeed, in the present case we have $d = a/N$, which becomes infinitesimal as $N \to \infty$. So d is certainly smaller than λ.

The zeros of $I_{\text{tot}}(\theta)$ occur when β is a multiple of 2π (except $\beta = 0$). And since $\beta \equiv 2\pi a\sin\theta/\lambda$, this is equivalent to $a\sin\theta$ being a multiple of λ. We'll give a physical reason for this relation below in Section 9.3.2, but first let's give two other derivations of Eq. (31).

Geometric derivation

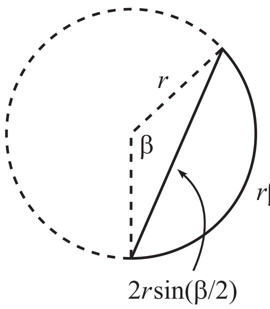


Figure 24

We can give another derivation of the diffraction pattern by using the geometric construction in Section 9.2.2. In the $N \to \infty$ limit, the little vectors in Fig. 14 become infinitesimal, so the crooked curve becomes a smooth curve with no kinks. If $\beta = 0$ (which corresponds to $\theta = 0$ and hence $\alpha = 0$ in Fig. 14), then all of the infinitesimal vectors are in phase, so we get a straight line pointing to the right. If β is nonzero, then the vectors curl around, and we get something like the picture shown in Fig. 24. The bottom infinitesimal vector corresponds to one end of the wide slit, and the top infinitesimal vector corresponds to the other end. The pathlength difference between the ends is $a\sin\theta$, so the phase difference is $ka\sin\theta$, which is by definition β. This phase difference is the angle between the top and bottom vectors in Fig. 24. But this angle equals the central angle subtended by the arc. The central angle is therefore β, as shown.

Now, the amplitude $A_{\text{tot}}(0)$ is the length of the straight line in the $\beta = 0$ case. But this is also the length of the arc in Fig. 24, which we know is $r\beta$, where r is the radius of the circle. And $A_{\text{tot}}(\theta)$ is the sum of all the infinitesimal vectors, which is the straight line shown. From the isosceles triangle in the figure, this sum has length $2r\sin(\beta/2)$. Therefore, since the intensity is proportional to the square of the amplitude, we have

$$\frac{I_{\text{tot}}(\theta)}{I_{\text{tot}}(0)} = \left(\frac{2r\sin(\beta/2)}{r\beta}\right)^2 = \left(\frac{\sin(\beta/2)}{\beta/2}\right)^2, \tag{33}$$

in agreement with Eq. (31).

Continuous integral

We can also find the diffraction pattern by doing a continuous integral over all the phases from the possible paths from different parts of the wide slit. Let the slit run from $y = -a/2$ to $y = a/2$. And let $B(\theta)\,dy$ be the amplitude that would be present at a location θ on the screen if only an infinitesimal slit of width dy was open. So $B(\theta)$ is the amplitude (on the screen) per unit length (in the slit). $B(\theta)\,dy$ is the analog of the $A(\theta)$ in Eq. (11). If we measure the pathlengths relative to the midpoint of the slit, then the path that starts at position y is shorter by $y\sin\theta$ (so it is longer if $y < 0$). It therefore has a relative phase of $e^{-iky\sin\theta}$. Integrating over all the paths that emerge from the different values of y (through imaginary slits of width dy) gives the total wave at position θ on the screen as (up to an overall phase from the $y = 0$ point, and ignoring the ωt part of the phase)

$$E_{\text{tot}}(\theta) = \int_{-a/2}^{a/2} \big(B(\theta)\,dy\big)e^{-iky\sin\theta}. \tag{34}$$

This is the continuous version of the discrete sum in Eq. (11). $B(\theta)$ falls off like $1/\sqrt{r}$, where $r = D/\cos\theta$. However, as in Section 9.2.1, we'll assume that θ is small, which mean that we can let $\cos\theta \approx 1$. (And even if θ isn't small, we're not so concerned about the exact intensities and the overall envelope of the diffraction pattern.) So we'll set $B(\theta)$ equal to the constant value of $B(0)$. We therefore have

$$\begin{aligned} E_{\text{tot}}(\theta) \approx B(0)\int_{-a/2}^{a/2} e^{-iky\sin\theta}\,dy &= \frac{B(0)}{-ik\sin\theta}\Big(e^{-ik(a/2)\sin\theta} - e^{ik(a/2)\sin\theta}\Big) \\ &= B(0)\cdot\frac{-2i\sin\left(\frac{ka\sin\theta}{2}\right)}{-ik\sin\theta} \\ &= B(0)a\cdot\frac{\sin\left(\frac{1}{2}ka\sin\theta\right)}{\frac{1}{2}ka\sin\theta}. \end{aligned} \tag{35}$$

There aren't any phases here, so this itself is the amplitude $A_{\text{tot}}(\theta)$. Taking the usual limit at $\theta = 0$, we obtain $A_{\text{tot}}(0) = B(0)a$. Therefore, $A_{\text{tot}}(\theta)/A_{\text{tot}}(0) = \sin(\beta/2)/(\beta/2)$, where $\beta \equiv ka\sin\theta$. Since the intensity is proportional to the square of the amplitude, we again arrive at Eq. (31).

9.3.2 Width of the diffraction pattern

From Fig. 23, we see that most of the intensity of the diffraction pattern is contained within the main bump where $|\beta| < 2\pi$. Numerically, you can shown that the fraction of the total area that lies under the main bump is about 90%. So it makes sense to say that the angular half-width of the pattern is given by

$$\beta = 2\pi \quad\Longrightarrow\quad \frac{2\pi a\sin\theta}{\lambda} = 2\pi \quad\Longrightarrow\quad \boxed{\sin\theta = \frac{\lambda}{a}} \tag{36}$$

For small θ, this becomes

$$\theta \approx \frac{\lambda}{a} \qquad \text{(for small } \theta\text{)} \tag{37}$$

Note that this is inversely proportional to a. The narrower the slit, the wider the diffraction pattern. There are two ways of understanding the $\sin\theta = \lambda/a$ result, or equivalently why the intensity is zero when this relation holds.

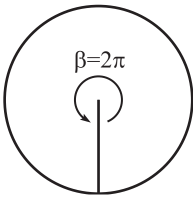


Figure 25

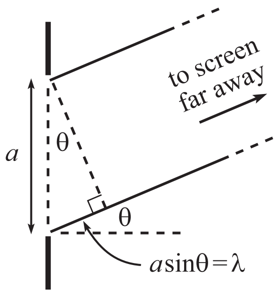


Figure 26

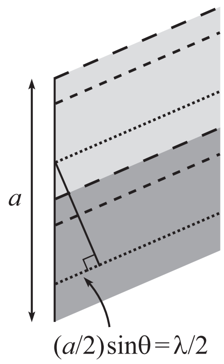


Figure 27

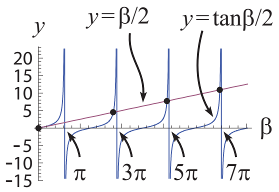


Figure 28

- As indicated in Eq. (36), $\sin\theta = \lambda/a$ is equivalent to $\beta = 2\pi$. So in the geometric construction that led to Eq. (33), this means that the arc in Fig. 24 is actually a full circle, as shown in Fig. 25. The sum of all the infinitesimal vectors is therefore zero, so the amplitude and intensity are zero.

- If $\sin\theta = \lambda/a$, then the pathlength from one end of the slit is $a\sin\theta = \lambda$ longer than the pathlength from the other end, as shown in Fig. 26. So the waves from the two ends are in phase. You might think that this implies that there should be constructive interference (which would be the case if we simply had two infinitesimal slits separated by a). But in fact it's exactly the opposite in the present case of a continuous wide slit. We have complete destructive interference when the whole slit is taken into account, for the following reason.

  Imagine dividing the slit into two halves, as shown in Fig. 27. For every path in the upper half (the lightly shaded region), there is a path in the lower half (the darkly shaded region) that is $\lambda/2$ longer. So the two waves are exactly out of phase. Three pairs of dotted lines are shown. The waves therefore cancel in pairs throughout the slit, and we end up with zero amplitude. This is equivalent to saying that the phases cancel at diametrically opposite points in Fig. 25.

You can quickly show by taking a derivative that the local maxima of $I(\beta)/I(0) = \big((\sin(\beta/2))/(\beta/2)\big)^2$ occur when $\tan\beta/2 = \beta/2$. This must be solved numerically, but you can get a sense of what the roots are by plotting the functions $y = \beta/2$ and $y = \tan\beta/2$, and then looking at where the curves intersect. In Fig. 28 the intersections are indicated by the dots, and you can see that the associated β values (in addition to the $\beta = 0$ root) are close to 3π, 5π, 7π, etc. (the larger β is, the better these approximations are). These maxima therefore occur roughly halfway between the zeros, which themselves occur when β equals (exactly) 2π, 4π, 6π, etc. The task of Problem [to be added] is to show how the 3π, 5π, 7π, etc. values follow from each of the above two bullet-point reasonings.

The (half) angular spread of the beam, $\theta \approx \lambda/a$, is large if a is small. Physically, the reason for this is that if a is small, then the beam needs to tilt more in order to generate the path differences (and hence phase differences) that lead to the total cancelation at the first zero (at $\beta = 2\pi$). Said in a different way, if a is small, then the beam can tilt quite a bit and still have all the different paths be essentially in phase.

If $a < \lambda$, then even if the beam is titled at $\theta = \pi/2$, there still can't be total cancelation. So if $a < \lambda$, then the diffraction pattern has no zeros. It simply consists of one bump that is maximum at $\theta = 0$ and decreases as $\theta \to \pi/2$. It actually does approach zero in this limit (assuming we have a flat screen and not a cylindrical one) because of the $B(\theta)$ factor in Eq. (34). $\theta \to \pi/2$ corresponds to points on the screen that are very far from the slit, so the amplitude of all the waves is essentially zero.

In the limit $a \ll \lambda$, all of the waves from the different points in the slit are essentially in phase for any angle θ. Interference effects therefore don't come into play, so the slit behaves essentially like a point (or rather a line) source. The only θ dependence in the diffraction pattern comes from the $B(\theta)$ factor. If we have a cylindrical screen, then we don't even have this factor, so the diffraction pattern is constant. If we have a flat screen and deal only with small angles (for which $B(\theta) \approx B(0)$), then the diffraction pattern is constant there, too. Since we generally deal with small angles, it is customary to say that $a \ll \lambda$ leads to a constant diffraction pattern. We now see what we meant by "narrow slits" or "infinitesimal slits" in Sections 9.1 and 9.2. We meant that $a \ll \lambda$. This allowed us to ignore any nontrivial diffraction effects from the individual slits.

If we have the other extreme where $a \gg \lambda$, then even the slightest tilt of the beam will lead to a pathlength difference of λ between the paths associated with the two ends of the slit. This corresponds to the first zero at $\beta = 2\pi$. So the diffraction pattern is very narrow in an angular sense. In the far-field limit, the distances on the screen arising from the angular spread (which take the rough form of $D\theta$) completely dominate the initial spread of the beam due to the thickness a of the slit. So as long as D is very large, increasing the value of a will *decrease* the size of the bright spot in the screen. If the screen were right next to the slit, then increasing a would of course increase the size of the spot. But we're working in the far-field limit here, where the angular spread is all that matters.

Let's now do two examples that illustrate various aspects of diffraction. For both of these examples, we'll need to use the diffraction pattern from a wide slit, but with it *not* normalized to the value at $\theta = 0$. Conveniently, this is the result we found in Eq. (35), which we'll write in the form,

$$A_{\text{tot}}(\theta) = B(0) \cdot \frac{\sin\left(\frac{1}{2}ka\sin\theta\right)}{\frac{1}{2}k\sin\theta} \quad \Longrightarrow \quad A_{\text{tot}}(\theta) \propto \frac{\sin\left(\frac{1}{2}ka\sin\theta\right)}{\frac{1}{2}k\sin\theta}. \tag{38}$$

The $B(0)$ term (which is simply a measure of how bright the light is as it impinges on the slit) will cancel out in these two examples, so all that matters is the second proportionality relation in Eq. (38). The intensity is then

$$I_{\text{tot}}(\theta) \propto \left(\frac{\sin\left(\frac{1}{2}ka\sin\theta\right)}{\frac{1}{2}k\sin\theta}\right)^2. \tag{39}$$

Example (Four times the light?): If we let $\theta = 0$ in Eq. (39), and if we make the usual $\sin\epsilon \approx \epsilon$ approximation, we see that $I_{\text{tot}}(0) \propto a^2$. This means that if we double a, then $I_{\text{tot}}(0)$ increases by a factor of 4. Intensity equals energy per unit time per unit area, so 4 times as much energy is now hitting a given tiny region around $\theta = 0$. Does this makes sense? Does it mean that if we double the width of the slit, then 4 times as much light makes it through?

Solution: The answers to the above two questions are yes and no, respectively. The answer to the second one had better be no, because otherwise energy would be created out of nowhere. If we double the width of the slit, then our intuition is entirely correct: twice as much light makes it through, not 4 times as much. The reason why the 4-fold increase in $I_{\text{tot}}(0)$ doesn't imply that 4 times as much light makes it through is the following.

The critical point is that although the intensity goes up by a factor of 4 at $\theta = 0$, the diffraction pattern gets *thinner*. So the range of θ vales that have a significant intensity decreases. It turns out that the combination of theses effects leads to just a factor of 2 in the end. This is quite believable, and we can prove it quantitatively as follows. We'll assume that the bulk of the diffraction pattern is contained in the region where θ is small. For the general case without this assumption, see Problem [to be added].

If θ is small, then we can use $\sin\theta \approx \theta$ in Eq. (39) to write[5]

$$I_{\text{tot}}(\theta) \propto \left(\frac{\sin\left(\frac{1}{2}ka\theta\right)}{\frac{1}{2}k\theta}\right)^2. \tag{40}$$

[5]Note that even though we're assuming that θ is small, we *cannot* assume that $ka\theta/2$ is small and thereby make another $\sin\epsilon \approx \epsilon$ approximation in the numerator. This is because k may be large, or more precisely λ may be much smaller than a.

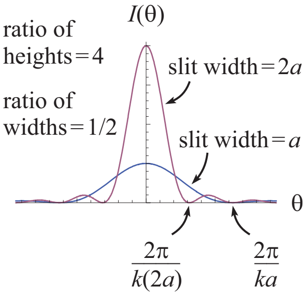


Figure 29

The larger a is, the quicker $\sin(ka\theta/2)$ runs through its cycles. In particular, the first zero (which gives the "width" of the diffraction pattern) occurs at $\theta = 2\pi/ka$. This is proportional to $1/a$, so increasing a by a general factor f *shrinks* the pattern by a factor f in the horizontal direction. And since we saw above that $I_{\text{tot}}(0) \propto a^2$, increasing a by a factor f *expands* the pattern by a factor of f^2 in the vertical direction. The combination of these two effects makes the total area under the curve (which is the total intensity) increase by a factor $f^2/f = f$. This is consistent with the fact that f times as much light makes it through the widened slit of width fa, as desired. This reasoning is summarized in Fig. 29 for the case where $f = 2$ (with arbitrary units on the vertical axis).

REMARKS: The main point here is that intensity equals energy per unit time *per unit area*. So we can't conclude anything by using only the fact that I_{tot} increases by a factor of f^2 at the specific point $\theta = 0$. We need to integrate over all θ values on the screen (and then technically multiply by some length in the direction perpendicular to the page to obtain an actual area, but this isn't important for the present discussion). From Fig. 29, the curve as a whole is most certainly *not* simply scaled up by a factor f^2.

There are two issues we glossed over in the above solution. First, in finding the area under the intensity curve, the integral should be done over the position x along the screen, and not over θ. But since x is given by $D\tan\theta \approx D\theta$ for small θ, the integral over x is the same (up to the constant factor D) as the integral over θ. Second, we actually showed only that $I_{\text{tot}}(\theta)$ increases by a factor of f^2 right at the origin. What happens at other corresponding points isn't as obvious. If you want to be more rigorous about the integral $\int I_{\text{tot}}(\theta)\,d\theta$, you can let $a \to fa$ in Eq. (40), and then make the change of variables $\theta' \equiv f\theta$. The integral will pick up a factor of $f^2/f = f$. But having said this, you can do things completely rigorously, with no approximations, in Problem [to be added]. ♣

Example (Increasing or decreasing intensity?): Given a slit with width a, consider the intensity at a particular point on the screen that is a reasonable distance off to the side. (By this we mean that the distance is large compared with the width λ/a of the central bump.) If we make a larger, will the intensity increase or decrease at the point? By intensity here, we mean the average intensity in a small region, so that we take the average over a few bumps in the diffraction pattern.

On one hand, increasing a will allow more light through the slit, so the intensity should increase. But on the other hand, increasing a will make the diffraction pattern narrower, so the intensity should decrease. Which effect wins?

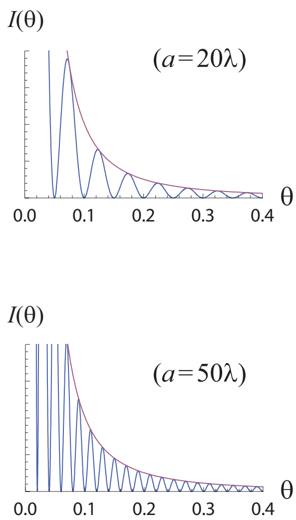


Figure 30

Solution: It turns out that these two effects exactly cancel, for the following reason. If we take the average over a few oscillations of the $I_{\text{tot}}(\theta)$ function in Eq. (40), the $\sin^2(ka\theta/2)$ term averages to 1/2 (we can ignore the variation of the denominator over a few oscillations of the sine term). So the average value of $I_{\text{tot}}(\theta)$ in a small region near a given value of θ is $I_{\text{tot,avg}}(\theta) \propto 2/(k^2\theta^2)$. This is independent of a. So the intensity at the given point doesn't change as we widen the slit. In short, the envelope of the wiggles in the diffraction pattern behaves like a $1/\theta^2$ function, and this is independent of a. Fig. 30 shows the diffraction patterns for $a = 20\lambda$ and $a = 50\lambda$. The envelope is the same for each.

9.3.3 Relation to the Fourier transform

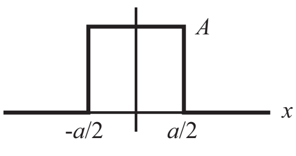


Figure 31

If the $\sin\left(\frac{1}{2}ka\sin\theta\right)/\frac{1}{2}ka\sin\theta$ function in Eq. (31) looks familiar to you, it's because this function is basically (up to an overall constant) the Fourier transform of the square-wave function shown in Fig. 31. We discussed this function in Chapter 3, but let's derive the transform again here since it's quick. If we let the argument of the Fourier transform be $k\sin\theta$ instead of the usual k (we're free to pick it to be whatever we want; if you wish, you can define $k' \equiv k\sin\theta$ and work in terms of k'), then Eq. (3.43) gives

$$\begin{aligned} C(k\sin\theta) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-i(k\sin\theta)x}\,dx \\ &= \frac{1}{2\pi}\int_{-a/2}^{a/2} Ae^{-ikx\sin\theta}\,dx \\ &= \frac{A}{2\pi}\frac{e^{-i(ka\sin\theta)/2} - e^{i(ka\sin\theta)/2}}{-ik\sin\theta} \\ &= \frac{A}{2\pi}\frac{-2i\sin\left(\frac{1}{2}ka\sin\theta\right)}{-ik\sin\theta} \\ &= \frac{aA}{2\pi}\frac{\sin\left(\frac{1}{2}ka\sin\theta\right)}{\frac{1}{2}ka\sin\theta}\,. \end{aligned} \tag{41}$$

So in view of Eq. (31), the intensity on the screen is (up to an overall constant) the square of the Fourier transform of the slit. This might seem like a random coincidence, but there's actually a good reason or it: In Eq. (35) we saw that the amplitude of the diffraction pattern was obtained by integrating up a bunch of $e^{-iky\sin\theta}$ phases. But this is exactly the same thing we do when we compute a Fourier transform. So that's the reason, and that's pretty much all there is to it.

More generally, instead of a slit we can have a wall with transmittivity $T(y)$. $T(y)$ gives the fraction (compared with no wall) of the amplitude coming through the wall at position y. For example, a normal slit has $T(y) = 1$ inside the slit and $T(y) = 0$ outside the slit. But you can imagine having a partially opaque wall where $T(y)$ takes on values between 0 and 1 in various regions. In terms of $T(y)$, the total wave at an angle θ on the screen is given by Eq. (35), but with the extra factor of $T(y)$ in the integrand:

$$E_{\text{tot}}(\theta) = B(0)\int_{-\infty}^{\infty} T(y)e^{-iky\sin\theta}\,dy. \tag{42}$$

Note that the integral now runs from $-\infty$ to ∞, although there may very well be only a finite region where $T(y)$ is nonzero. Up to an overall constant, the result of this integral is simply $\tilde{T}(k\sin\theta)$, where $\tilde{T}$ denotes the Fourier transform of $T(y)$. So the diffraction pattern is the (absolute value of the square of the) Fourier transform of the transmittivity function. (We're assuming that the region of nonzero $T(y)$ is small compared with the distance to the screen, so that we can use the standard far-field approximation that all the paths from the different points in the "slit" to a given point on the screen have equal lengths (multiplicatively).)

Recall the uncertainty principle from Problem [to be added] in Chapter 3, which stated that the thinner a function $f(x)$ is, the broader the Fourier transform $\tilde{f}(k)$ is, and vice versa. The present result (that the diffraction pattern is the square of the Fourier transform of the slit) is consistent with this. A narrow slit gives a wide diffraction pattern, and a wide slit gives a narrow (in an angular sense) pattern.

There are two ways of defining the Fourier transform. The definition we used above is the statement in the second equation in Eq. (3.43): The Fourier transform is the result of multiplying each $f(x)$ value by a phase e^{-ikx} and then integrating. This makes it clear why the diffraction pattern is the Fourier transform of the transmittivity function, because the diffraction pattern is the result of attaching an extra phase of $e^{-iky\sin\theta}$ to the Huygens wavelets coming from each point in the slit.

The other definition of the Fourier transform comes from the first equation in Eq. (3.43): The Fourier transform gives a measure of how much the function $f(x)$ is made up of the function e^{ikx}. (This holds in a simpler discrete manner in the case of a Fourier series for a periodic function.) Does this interpretation of the Fourier transform have an analog in the diffraction setup? That is, does the diffraction pattern somehow give a measure of how much the transmittivity function is made up of the function $e^{iky\sin\theta}$? Indeed it does, for the following reason.

We'll be qualitative here, but this should suffice to illustrate the general idea. Let's assume that we observe a large amplitude in the diffraction pattern at an angle θ. This means that the wavelets from the various points in the slit generally constructively interfere at the angle θ. From Fig. 32, we see that the transmittivity function must have a large component with spatial period $\lambda/\sin\theta$. This means that the spatial frequency of the transmittivity function is $\kappa = 2\pi/(\lambda/\sin\theta) = (2\pi/\lambda)\sin\theta = k\sin\theta$, where k is the spatial frequency of the light wave. In other words, a large amplitude at angle θ means that $T(y)$ has a large component with spatial frequency $k\sin\theta$. The larger the amplitude at angle θ, the larger the component of $T(y)$ with spatial frequency $k\sin\theta$. But this is exactly the property that the Fourier transform of $T(y)$ has: The larger the value of $\tilde{T}(k\sin\theta)$, the larger the component of $T(y)$ with spatial frequency $k\sin\theta$. So this makes it believable that the amplitude of the diffraction pattern equals the Fourier transform of $T(y)$, with $k\sin\theta$ in place of the usual k. The actual proof of this fact is basically the statement in Eq. (42).

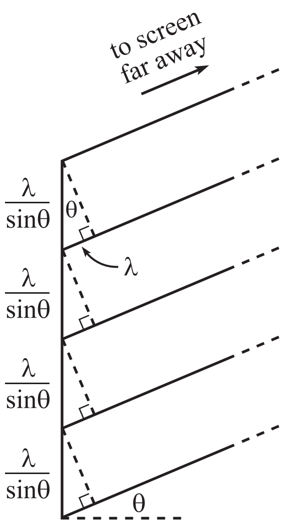


Figure 32

9.4 Combined Interference and diffraction

So far we've dealt with either N infinitesimally thin slits, or one wide slit. We'll now combine these two setups and consider N wide slits. Let the slits have width a, and let the spatial period be d (this is the distance between, say, two adjacent bottom ends). Fig. 33 shows the case with $N = 3$ and $d = 3a$. We'll continue to work in the far-field limit.

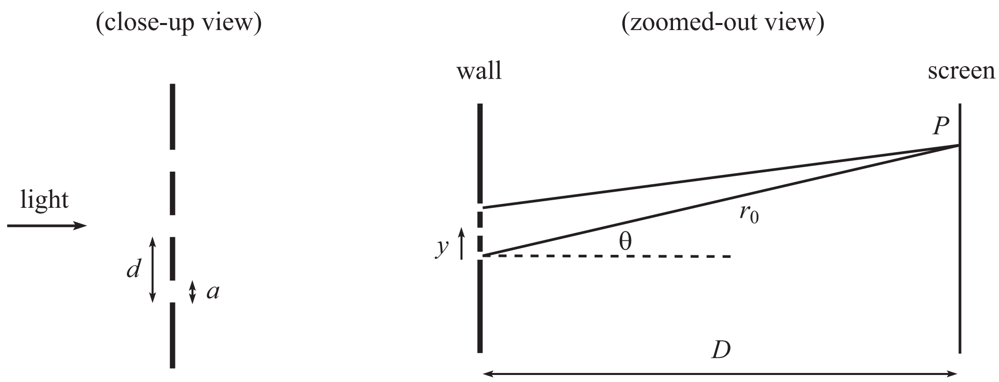


Figure 33

What is the amplitude of the wave at an angle θ on a distant screen? To answer this, we can use the reasoning in the "Continuous integral" derivation of the one-wide-slit result in Section 9.3.1. Let r_0 be the pathlength from the bottom of the bottom slit, as shown in Fig. 33. Define y to be the distance from the bottom of the bottom slit up to a given location in a slit. Then the relevant y values are from 0 to a for the bottom slit, then d to $d + a$ for the next slit, and so on.

The integral that gives the total wave from the bottom slit is simply the integral in Eq. (34), but with the integration now running from 0 to a. (We technically need to multiply by the phase $e^{i(kr_0-\omega t)}$, but this phase is tacked on uniformly to all the slits, so it doesn't affect the overall amplitude.) The integral that gives the total wave from the second slit is again the same, except with the integration running from d to $d + a$. And so on, up to limits of $(N-1)d$ and $(N-1)d+a$ for the top slit. So the total wave at angle θ from all the slits is (as usual, we'll approximate the $B(\theta)$ in Eq. (34) by $B(0)$)

$$E_{\text{tot}}(\theta) = B(0)\left(\int_0^a e^{-iky\sin\theta}\,dy + \int_d^{d+a} e^{-iky\sin\theta}\,dy + \cdots + \int_{(N-1)d}^{(N-1)d+a} e^{-iky\sin\theta}\,dy\right). \tag{43}$$

The second integral here is simply $e^{-ikd\sin\theta}$ times the first integral, because the y values are just shifted by a distance d. Likewise, the third integral is $e^{-2ikd\sin\theta}$ times the first. Letting $z \equiv e^{-ikd\sin\theta}$, we therefore have

$$E_{\text{tot}}(\theta) = B(0)\left(\int_0^a e^{-iky\sin\theta}\,dy\right)\left(1 + z + z^2 + \cdots + z^{N-1}\right). \tag{44}$$

Shifting the limits of this integral by $-a/2$ (which only introduces a phase, which doesn't affect the amplitude) puts it in the form of Eq. (34). So we can simply copy the result in Eq. (35). (Or you can just compute the integral with the 0 and a limits.) And the geometric series is the same one we calculated in Eq. (12), so we can copy that result too. (Our z here is the complex conjugate of the z in Eq. (12), but that will only bring in an overall minus sign in the final result, which doesn't affect the amplitude.) So the total amplitude at angle θ is

$$A_{\text{tot}}(\theta) = B(0)a \cdot \frac{\sin\left(\frac{1}{2}ka\sin\theta\right)}{\frac{1}{2}ka\sin\theta} \cdot \frac{\sin\left(\frac{1}{2}Nkd\sin\theta\right)}{\sin\left(\frac{1}{2}kd\sin\theta\right)}. \tag{45}$$

Taking the usual limit as $\theta \to 0$, the value of the amplitude at $\theta = 0$ is $B(0)aN$. The intensity relative to $\theta = 0$ is therefore

$$\boxed{\frac{I_{\text{tot}}(\theta)}{I_{\text{tot}}(0)} = \left(\frac{\sin\left(\frac{1}{2}ka\sin\theta\right)}{\frac{1}{2}ka\sin\theta} \cdot \frac{\sin\left(\frac{1}{2}Nkd\sin\theta\right)}{N\sin\left(\frac{1}{2}kd\sin\theta\right)}\right)^2} \tag{46}$$

This result really couldn't have come out any nicer. It is simply the product of the results for the two separate cases we've discussed. The first quotient is the one-wide-slit diffraction result, and the second quotient is the N-thin-slit interference result. Note that since $Nd > a$ (because $d > a$), the second quotient oscillates faster than the first. You can therefore think of this result as the N-thin-slit interference result modulated by (that is, with an envelope of) the one-wide-slit diffraction result.

In retrospect, it makes sense that we obtained the product of the two earlier results. At a given value of θ, we can think of the setup as just N-thin-slit interference, but where the amplitude from each slit is decreased by the one-wide-slit diffraction result. This is clear if we rearrange Eq. (45) and write it as (we'll switch the $B(0)$ back to $B(\theta)$)

$$A_{\text{tot}}(\theta) = \left(B(\theta)a\frac{\sin\left(\frac{1}{2}ka\sin\theta\right)}{\frac{1}{2}ka\sin\theta}\right) \cdot \frac{\sin\left(\frac{1}{2}Nkd\sin\theta\right)}{\sin\left(\frac{1}{2}kd\sin\theta\right)}. \tag{47}$$

This is the N-thin-slit result, with $B(\theta)a$ (which equals the $A(\theta)$ in Eq. (14)) replaced by $B(\theta)a \cdot \sin\left(\frac{1}{2}ka\sin\theta\right)\big/\frac{1}{2}ka\sin\theta$. Basically, at a given θ, you can't tell the difference between a wide slit, and an infinitesimal slit with an appropriate amount of light coming through.

Fig. 34 shows the interference/diffraction pattern for $N = 4$ and for various slit widths a, given the spatial period d. The coordinate on the horizontal axis is $\alpha \equiv kd\sin\theta$. The $a \approx 0$ plot is exactly the same as the thin-slit plot in Fig. 11, as it should be. As the width a increases, the envelope becomes narrower. Recall from Eq. (36) that the width of the one-wide-slit diffraction pattern is inversely proportional to a.

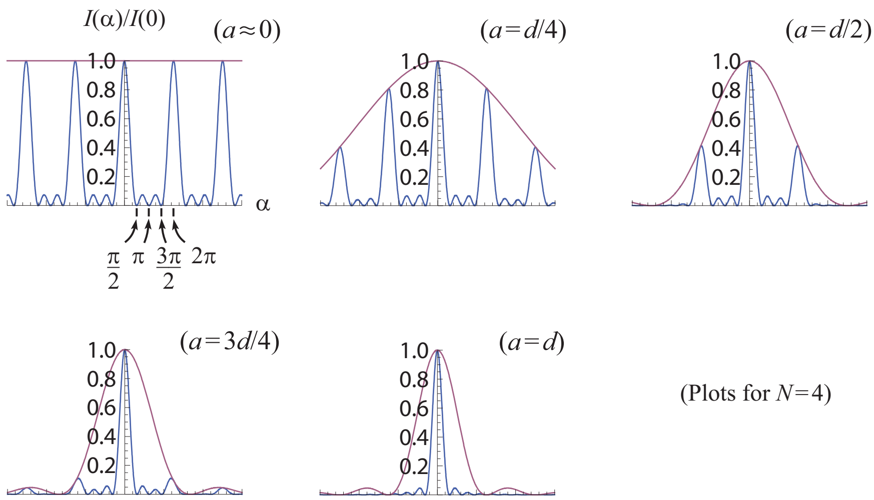


Figure 34

When a finally reaches the value of d in the last plot, the four slits blend together, and we simply have one slit with width $4a$. (It doesn't make any sense to talk about a values that are larger than d.) And the $a = d$ plot is indeed the plot for a single wide slit with width $4a$. The only difference between it and the envelope (which comes from a slit width a) is that it is squashed by a factor of 4 in the horizontal direction. It turns out that in the $a = d$ case, the zeros of the envelope fall exactly where the main peaks would be if the envelope weren't there (see the $a \approx 0$ case). This follows from the fact that the zeros of the diffraction envelope occur when $\beta \equiv ka \sin\theta$ equals 2π, while the main peaks of the N-slit interference pattern occur when $\alpha \equiv kd \sin\theta$ equals 2π. So if $a = d$, these occur at the same locations.

9.5 Near-field diffraction

9.5.1 Derivation

Everything we've done so far in this chapter has been concerned with the far-field approximation (the so-called Fraunhofer approximation). We have assumed that the distance to the screen is large compared with the span of the slit(s). As discussed in Section 9.1, this assumption leads to two facts:

- The pathlengths from the various points in the slit(s) to a given point on the screen are all essentially equal in a *multiplicative* sense. This implies that the amplitudes of all the various wavelets are equal. In other words, we can ignore the $1/\sqrt{r}$ dependence in the individual amplitudes of the cylindrically-propagating Huygens wavelets.

- The paths from the various points in the slit(s) to a given point on the screen are all essentially parallel. This implies that the *additive* difference between adjacent pathlengths equals $d \sin\theta$ (or $dy \sin\theta$ in the continuous case). The pathlengths therefore take the nice general form of $r_0 + nd \sin\theta$ (or $r_0 + y \sin\theta$), and the phases are easy to get a handle on.

In the first of these points, note that we're talking about the various distances from a *particular point on the screen* to all the *different points in the slit(s)*. We are *not* talking about the various distances from a particular point in the slit(s) to all the different points on the screen. These distances certainly aren't equal; the fact that they aren't equal is what brought in the factor of $A(\theta)$ in, say, Eq. (4) or Eq. (14). But this lack of equality is fine; it simply leads to an overall envelope of the interference curve. The relevant fact in the far-field approximation is that the various distances from a particular point on the *screen* to all the different points in the *slit(s)* are essentially equal. This lets us associate all the different wavelets (at a given point on the screen) with a single value of $A(\theta)$, whatever that value may be.

We'll now switch gears and discuss the near-field approximation (the so-called Fresnel approximation). That is, we will *not* assume that the distance to the screen is large compared with the span of the slit(s). The above two points are now invalid. To be explicit, in the near-field case:

- We cannot say that the pathlengths from the various points in the slit(s) to a given point on the screen are all equal in a *multiplicative* sense. We will need to take into account the $1/\sqrt{r}$ dependence in the amplitudes.

- We cannot say that the pathlengths take the nice form of $r_0 + nd\sin\theta$ (or $r_0 + y\sin\theta$). We will have to calculate the lengths explicitly as a function of the position in the slit(s).

The bad news is that all of the previous results in this chapter are now invalid. But the good news is that they're close to being correct. The strategy for the near-field case is basically the same as for the far-field case, as long as we incorporate the changes in the above two points.

The procedure is best described by an example. We'll look at a continuous case involving diffraction from a wide slit, but we could of course have a near-field setup involving interference from N narrow slits, or a combination of interference and diffraction from N wide slits.

Our wide slit will actually be an infinite slit. Our goal will be to find the intensity at the point P directly across from the top of a "half-wall" (see Fig. 35). Since our slit is infinitely large, we're automatically in the near-field case, because it is impossible for the wall-screen distance D to be much greater than the slit width a, since $a = \infty$. The various pathlengths (which are infinite in number) to the given point P from all of the possible points in the slit (three of these paths are indicated by dotted lines in Fig. 35) certainly cannot be approximated as having the same length. These paths have lengths $r(y) = \sqrt{D^2 + y^2}$, where y is measured from the top of the wall. If we instead had an infinite number of thin slits extending upward with separation d, the pathlengths would be $r_n = \sqrt{D^2 + (nd)^2}$.

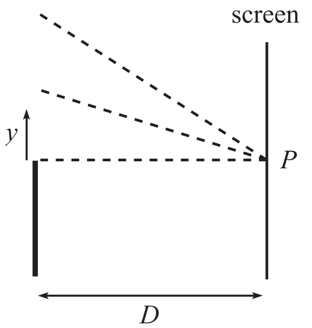


Figure 35

Since the amplitudes of the various cylindrically-propagating wavelets are proportional to $1/\sqrt{r}$, we need to tack on a factor of $1/\sqrt{r(y)}$ in front of each wavelet. More precisely, let $B_0\,dy$ be the amplitude of the wave that would hit point P due to an infinitesimal span dy in the slit at $y = 0$, if the distance D were equal to 1 (in whatever units we're using).[6] Then $B_0\,dy/\sqrt{r(y)}$ is the amplitude of the wave that hits point P due to a span dy in the slit at height y. The length $r(y)$ depends on where the screen is located (which gives D), and also on the height y.

As far as the phases go, the phase of the wavelet coming from a height y in the slit is $e^{ikr(y)}$, neglecting the $e^{-i\omega t}$ phase and an overall phase associated with the $y = 0$ path.

Using these facts about the amplitude and phase of the wavelets, we can integrate over the entire (infinite) slit to find the total wave at the point P directly across from the top

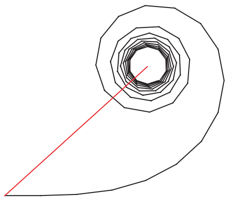
[6] B_0 is slightly different from the $B(0)$ in Eq. (35), because we didn't take into account the distance to the screen there. We assumed the position was fixed. But we want to be able to move the screen in the present setup and get a handle on how this affects things.

of the wall. The integral is similar to Eq. (34). But with the modified amplitude, the more complicated phase, and the new limits of integration, we now have

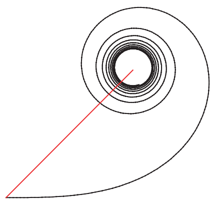
$$E_{\text{tot}}(P) = \int_0^\infty \frac{B_0\,dy}{\sqrt{r(y)}}\, e^{ikr(y)} = \int_0^\infty \frac{B_0\,dy}{(D^2+y^2)^{1/4}}\, e^{ik\sqrt{D^2+y^2}}\,. \tag{48}$$

This integral must be computed numerically, but we can get a sense of what's going on if we draw a picture similar to the far-field case in Fig. 14. In that figure we had little vectors of *equal* length wrapping around in a circle, with successive vectors always making the *same* angle with respect to each other. In the present near-field case, these two italicized words are modified for the following reasons.

Let's imagine discretizing the slit into equal dy intervals. Then as y increases, the lengths of the little vectors *decrease* due to the $(D^2+y^2)^{1/4}$ factor in the denominator in Eq. (48). Also, the phase doesn't increase at a constant rate. For small y, the phase hardly changes at all, because the derivative of the $\sqrt{D^2+y^2}$ term in the exponent is zero at $y=0$. But for large y, the rate of change of the phase approaches a constant, because the derivative of $\sqrt{D^2+y^2}$ equals 1 for $y \gg D$. So as y increases, the angle between successive vectors *increases* and asymptotically approaches a particular value. Both of these effects (the shortening lengths and the increasing rate of change of the phase) have the effect of decreasing the radius of curvature of the circle that is being wrapped around. In other words, the "circles" get tighter and tighter, and instead of a circle we end up with a spiral, as shown in Fig. 36 (we've arbitrarily chosen $\lambda = D$ here).



$\lambda = D,\ \Delta y = (0.1)D$



$\lambda = D,\ \Delta y \approx 0$

Figure 36

In the first spiral in Fig. 36, we have discretized the integral in Eq. (48) by doing a discrete sum over intervals with length $\Delta y = (0.1)D$ in the slit. You can see that the little vectors get smaller as they wrap around.[7] And you can also see that the angle between them starts off near zero and then increases. The second spiral shows the continuous limit where $\Delta y \approx 0$. So this corresponds to the actual integral in Eq. (48). In reality, this plot was generated by doing a discrete sum with $\Delta y = (0.01)D$. But the little vectors are too small to see, so the spiral is essentially continuous. So neither of these spirals actually corresponds to the continuous integral in Eq. (48). But the second one is a very good approximation. If you look closely, you can see that the slope of the straight line in the first spiral is slightly different from the slope in the second.

We haven't drawn the axes in these plots, because the absolute size of the resulting amplitude isn't so important. We're generally concerned with how large the amplitude is relative to a particular case. The most reasonable case to compare all others to is the one where there is no wall at all (so the slit extends from $y=-\infty$ to $y=\infty$). We'll talk about this below. But if you're curious about the rough size of the spiral, the horizontal and vertical spans (for the case in Fig. 36 where $\lambda = D$) are around $(0.5)B_0$.

This spiral is known as the *Cornu spiral*,[8] or the *Euler spiral*. In the present case where the upper limit on y is infinity, the spiral keeps wrapping around indefinitely (even though we stopped drawing it after a certain point in Fig. 36). The radius gets smaller and smaller, and the spiral approaches a definite point. This point is the sum of the infinite number of tiny vectors. The desired amplitude of the wave at P is the distance from the origin to this point, as indicated by the straight line in the figure. As usual, the whole figure rotates around in the plane with frequency ω as time progresses. The horizontal component of the straight line is the actual value of the wave.

[7] We've stopped drawing the vectors after a certain point, but they do spiral inward all the way to the center of the white circle you see in the figure. If we kept drawing them, they would end up forming a black blob where the white circle presently is.

[8] Technically, this name is reserved for the simpler approximate spiral we'll discuss in Section 9.5.3. But we'll still use the name here.

The shape of the spiral depends on the relative size of λ and D. If we define the dimensionless quantity z by $y \equiv zD$, then Eq. (48) can be written as (using $k = 2\pi/\lambda$ and $dy = D\,dz$)

$$E_{\text{tot}}(P) = \int_0^\infty \frac{B_0\sqrt{D}\,dz}{(1+z^2)^{1/4}}\, e^{2i\pi(D/\lambda)\sqrt{1+z^2}}. \tag{49}$$

For a given value of D/λ, the factor of $\sqrt{D}$ in the numerator simply scales the whole spiral, so it doesn't affect the overall shape. However, the factor of D/λ in the exponent does affect the shape, but it turns out that the dependence is fairly weak. If we instead had spherically propagating waves with $(D^2+y^2)^{1/2}$ instead of $(D^2+y^2)^{1/4}$ in the denominator of Eq. (48), then there would be a noticeable dependence on D/λ, especially for large λ.

9.5.2 Changing the slit

What happens if instead of extending to infinity, the slit runs from $y = 0$ up to a finite value y_{max}? The only change in Eq. (48) is that the upper limit is now y_{max}. The integrand is exactly the same. So Eqs. (48) and (49) become

$$E_{\text{tot}}(P) = \int_0^{y_{\text{max}}} \frac{B_0\,dy}{(D^2+y^2)^{1/4}}\, e^{ik\sqrt{D^2+y^2}} = \int_0^{z_{\text{max}}} \frac{B_0\sqrt{D}\,dz}{(1+z^2)^{1/4}}\, e^{2i\pi(D/\lambda)\sqrt{1+z^2}}. \tag{50}$$

In terms of Fig. 36 (we'll again assume $\lambda = D$), we now only march along the spiral until we get to the little vector associated with y_{max}, the location of which can be found numerically. (We know $r(y_{\text{max}})$, so we know the relative phase, so we know the angle (slope) of the spiral at the stopping point.) The amplitude is then the length of the line from the origin to the stopping point. A few cases are shown in Fig. 37.

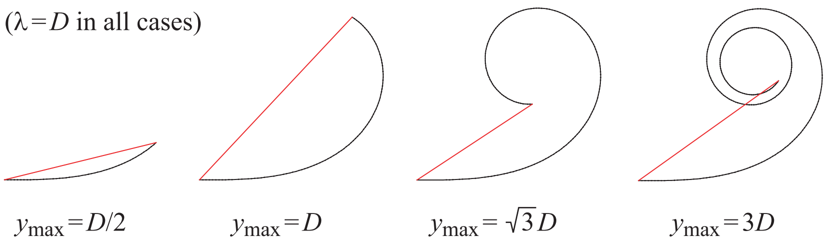


Figure 37

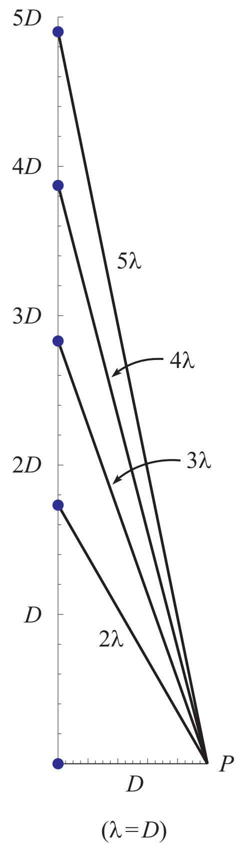


Figure 38

The $y = \sqrt{3}D$ case is an interesting one because it yields a pathlength of $\sqrt{D^2+y^2} = 2D$, which equals 2λ since we're assuming $\lambda = D$. This pathlength is therefore λ more than the pathlength associated with $y = 0$. So the wavelet from $y = \sqrt{3}D$ is in phase with the wavelet from $y = 0$. And this is exactly what we observe in the figure; the slope of the spiral at the $y = \sqrt{3}D$ point equals the slope at the start (both slopes equal zero). A few other values of y that yield pathlengths that are integral multiples of λ are shown in Fig. 38, and the corresponding points in the Cornu spiral are shown in Fig. 39 (eventually the points blend together). The spiral also has zero slope at the top of the "circles" in the spiral. These points correspond to pathlengths of $3\lambda/2$, $5\lambda/2$, $7\lambda/2$, etc. (The $\lambda/2$ is missing here because all the pathlengths are at least $D = \lambda$.) But the associated little vectors in the spiral now point to the left, because the wavelets are exactly out of phase with the wavelet from $y = 0$ (which we defined as pointing to the right).

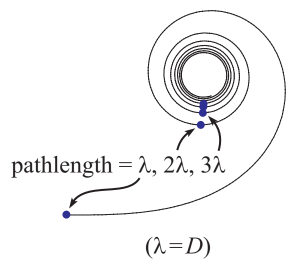


Figure 39

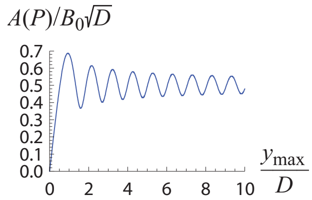


Figure 40

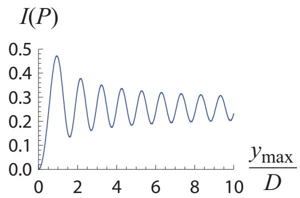


Figure 41

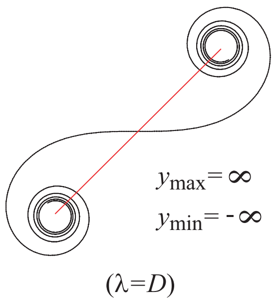


Figure 42

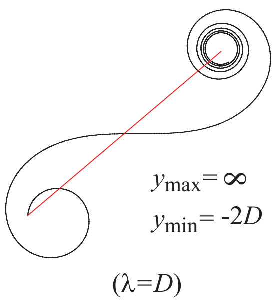


Figure 43

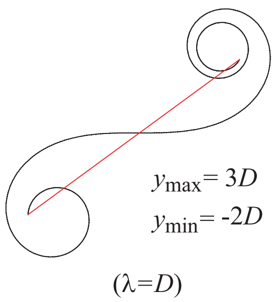


Figure 44

Remarks:

1. Note that the distance between the first two dots along the spiral in Fig. 39 is large, and then it decreases as we march along the spiral. There are two reasons for this. First, there is a large span of y values (from zero up to $y = \sqrt{3}D$) that corresponds to the region between the first two dots on the spiral. This span then gets smaller as y increases, and it eventually approaches the wavelength λ (which we've chosen to equal D). Second, the amplitudes of the wavelets get smaller as y increases (because the amplitude is proportional to $1/\sqrt{r}$), so the little vectors that make up the spiral get shorter as we spiral inward.
2. From Fig. 37, we see that the largest amplitude occurs for a y_{max} that is somewhere around D. It happens to occur at $y_{\text{max}} \approx (0.935)D$. If y_{max} is increased above this value, then apparently the upside of having more light coming through the slit is more than canceled out by the downside of this extra light canceling (due to the relation of the phases) some of the light that was already there. At any local max or min of the amplitude, the line representing the amplitude is perpendicular to the tangent to the spiral.
3. A plot of the amplitude, $A(P)$ (in units of $B_0\sqrt{D}$), as a function of y_{max} is shown in Fig. 40. As the spiral circles around and around, the amplitude oscillates up and down. Since the circles keep getting smaller, the bumps in Fig. 40 likewise keep getting smaller. The plot oscillates around a value that happens to be about 0.5. This is the amplitude associated with $y_{\text{max}} = \infty$. For large y_{max}, the period of the oscillations is essentially λ. This follows from the fact that as we noted in Fig. 38, if y increases by λ (which corresponds to a full circle in the spiral), then the pathlength essentially does also, if the path is roughly parallel to the wall. A plot of the intensity (which is proportional to the amplitude squared) is shown in Fig. 41, with arbitrary units on the vertical axis. ♣

What happens if we put the upper limit y_{max} back at infinity, but now move the top of the wall (the bottom of the slit) downward, so that y runs from some negative value, y_{min}, to infinity? (The point in question on the screen is still the point P directly across from $y = 0$.) To answer this, let's first consider the case where we move the top of the wall all the way down to $y = -\infty$. So we have no wall at all. We claim that the total amplitude at point P is given by the length of the diagonal line in Fig. 42. This is believable, of course, because the length of this line is twice the length of the line in Fig. 36 for the case where the "slit" was half as large. But to be rigorous, you can think of things in the following way.

In Fig. 36 imagine starting at $y = +\infty$ and decreasing down to $y = 0$. This corresponds to starting in the middle of the spiral and "unwrapping" clockwise around it until you reach the origin. The clockwise nature is consistent with the fact that the phase decreases as y decreases (because the pathlength decreases), and we always take positive phase to be counterclockwise. If you then want to keep going to negative values of y, you simply have to keep adding on the little vectors. But now the phase is *increasing*, because the pathlength is increasing. So the spiral wraps around *counterclockwise*. This is indeed what is happening in Fig. 42. (The spiral for the $y < 0$ region has to have the same shape as the spiral for the $y > 0$ region, of course, due to symmetry. The only question is how it is oriented.)

If we want the slit to go down to a finite value of y instead of $y = -\infty$, then we simply need to stop marching along the spiral at the corresponding point. For example, if the wall goes down to $y = -2D$, then the amplitude is given by the diagonal line in Fig. 43.

More generally, if we want to find the amplitude (still at the point P directly across from $y = 0$) due to a slit that goes from a finite y_{min} to a finite y_{max}, then we just need to find the corresponding points on the spiral and draw the line between them. For example, if a slit goes from $y = -2D$ to $y = 3D$, then the amplitude is given by the length of the diagonal line in Fig. 44. In the event that y_{min} and y_{max} are both positive (or both negative), the diagonal line is contained within the upper right (or lower left) half of the full Cornu spiral in Fig. 42. An example of this will come up in Section 9.5.5.

Remarks:

1. Note that in Fig. 44 the slope of the little vector at $y = -2D$ is nonzero. This is because we're still measuring all the phases relative to the phase of the wavelet at $y = 0$. If you want, you can measure all the phases relative to the phase at $y = -2D$ (or any other point). But only the relative phases matter, so this just rotates the whole figure, leaving the length of the diagonal line (the amplitude) unchanged. (The whole figure rotates around in the plane anyway as time goes on, due to the ωt term in the phase, which we've been ignoring since we only care about the amplitude.) By convention, it is customary to draw things as we've done in Fig. 42, with a slope of zero at the middle of the complete spiral.
2. In a realistic situation, the slit location is fixed, and we're concerned with the intensity at various points P on the screen. But instead of varying P, you can consider the equivalent situation where P is fixed (and defined to be across from $y = 0$), and where the slit is moved. This simply involves changing the values of y_{min} and y_{max}, or equivalently the endpoints of the diagonal line on the Cornu spiral representing the amplitude. So the above analysis actually gives the wave at any point P on the screen, not just the point across from $y = 0$.
3. In the earlier far-field case of interference and diffraction, the customary thing to do was to give the intensity relative to the intensity at $\theta = 0$. The most natural thing to compare the near-field amplitude to is the amplitude when there is no wall. This is the amplitude shown in Fig. 42. The Cornu spiral (the shape of which depends on the ratio D/λ in Eq. (49)) completely determines all aspects of the diffraction pattern for any location of the slit. And the length of the diagonal line in Fig. 42 gives the general length scale of the spiral, so it makes sense to compare all other lengths to this one. ♣

9.5.3 The $D \gg \lambda$ limit

When dealing with light waves, it is invariably the case that $D \gg \lambda$. If this relation holds, then the Cornu spiral approaches a particular shape, and we can write down an approximate (and simpler) expression for the integral in Eq. (49). Note that $D \gg \lambda$ does *not* mean that we're in the far-field limit. The far-field limit involves a comparison between D and the span of the slit(s), and it results in the approximation that all of the paths from the various points in the slit(s) to a given point on the screen are essentially equal in length (multiplicatively), and essentially parallel. The wavelength λ has nothing to do with this.

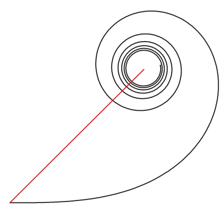


Figure 45

If $D \gg \lambda$, the actual *size* (but not the shape) of the spiral depends on λ; the smaller λ is, the smaller the spiral is. But the size (or the shape) doesn't depend on D. Both of these facts will follow quickly from the approximate expression we'll derive in Eq. (51) below. The fixed shape of the spiral is shown in Fig. 45, and it looks basically the same as Fig. 36.

The size dependence on λ is fairly easy to see physically. Even a slight increase in y from $y = 0$ will lead to a pathlength that increases on the order of λ, if λ is small. This means that the phases immediately start canceling each other out. The wave has no opportunity to build up, because the phase oscillates so rapidly as a function of y. The smaller λ is, the quicker the phases start to cancel each other.

If $D \gg \lambda$, we can give an approximate expression for the wave in Eq. (49). We claim that only small values of z (much less than 1) are relevant in Eq. (49). (These values correspond to y being much less than D.) Let's see what $E_{\text{tot}}(P)$ reduces to under the assumption that $z \ll 1$, then we'll justify this assumption.

If z is small, then we can use the approximation $\sqrt{1+z^2} \approx 1+z^2/2$ in both the exponent and the denominator of Eq. (49). We can ignore the $z^2/2$ term in the denominator, because it is small compared with 1. In the exponent we have $2\pi i D/\lambda + 2\pi i (D/\lambda) z^2/2$. The first of these terms is constant, so it just gives an overall phase in the integral, so we can ignore it. The second term involves a z^2, but we *can't* ignore it because it also contains a factor of

D/λ, which we're assuming is large. Eq. (49) therefore reduces to (recalling $z \equiv y/D$)

$$E_{\text{tot}}(P) \approx \int_0^{z_{\max}} B_0\sqrt{D}\, e^{i\pi(D/\lambda)z^2}\, dz = \int_0^{y_{\max}} \big(B_0/\sqrt{D}\big)e^{i\pi y^2/D\lambda}\, dy, \qquad (51)$$

where $z_{\max}$ is a number much smaller than 1, but also much larger than $\sqrt{\lambda/D}$. And $y_{\max} = Dz_{\max}$. The reason for this lower bound of $\sqrt{\lambda/D}$ comes from the following reasoning that justifies why we need to consider only z values that are much less than 1 in Eq. (49).

If z is much larger than $\sqrt{\lambda/D}$ (which corresponds to y being much larger than $\sqrt{\lambda D}$), but still satisfies our assumption of $z \ll 1$, then the exponent in Eq. (49) is a rapidly changing function of z. This corresponds to being deep inside the spiral where the circles are small. By this point in the spiral, the integral in Eq. (49) has essentially reached its limiting value, so it doesn't matter whether we truncate the integral at this (small) value of z or keep going to the actual upper limit of $z = \infty$. So if you want, you can let the upper bounds in Eq. (51) be infinity:

$$\boxed{E_{\text{tot}}(P) \approx \int_0^{\infty} B_0\sqrt{D}\, e^{i\pi(D/\lambda)z^2}\, dz = \int_0^{\infty} \big(B_0/\sqrt{D}\big)e^{i\pi y^2/D\lambda}\, dy} \qquad (52)$$

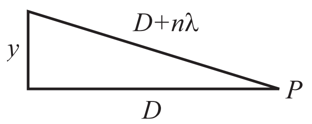


Figure 46

Pictorially, if you want to get a handle on which y values correspond to which points in the spiral, note that an increase in pathlength by one wavelength λ corresponds to a full circle of the spiral. The value of y that yields a pathlength of the form $D + n\lambda$ is found from the right triangle in Fig. 46. The Pythagorean theorem gives

$$D^2 + y^2 = (D + n\lambda)^2 \implies y^2 = 2nD\lambda + n^2\lambda^2 \implies y \approx \sqrt{2nD\lambda}, \qquad (53)$$

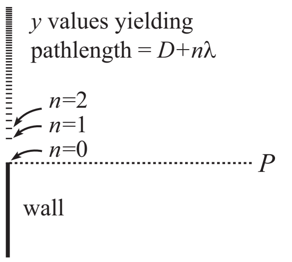


Figure 47

where we have ignored the second-order λ^2 term due to the $D \gg \lambda$ assumption. Fig. 47 shows the first 40 of these values of y for the case where $D/\lambda = 200$, although for actual setups involving light, this ratio is generally much higher, thereby making the approximations even better. This figure is analogous to Fig. 38. As you can see, the y values get closer together as y increases, due to the $\sqrt{n}$ dependence. This is consistent with the above statement that the exponent in Eq. (49) is a rapidly changing function of z.

Remarks:

1. As we noted above, the amplitude is essentially constant for small z, since $\sqrt{1+z^2} \approx 1$. So the little vectors that make up the spiral all have essentially the same length, for small z ($y \ll D$). The size of a "circle" in the spiral is therefore completely determined by how fast the phase is changing. Since the phase changes very quickly for $y \gg D\lambda$, the circles are very small, which means that we have essentially reached the limiting value at the center of the circle.

2. We mentioned above that if $D \gg \lambda$, then the size of the spiral depends on λ but not on D. And the shape depends on neither. These facts follow from Eq. (52) if we make the change of variables, $w \equiv z\sqrt{D/\lambda}$ (which equals $y/\sqrt{D\lambda}$). This turns the integral into

$$E_{\text{tot}}(P) \approx \sqrt{\lambda}\int_0^{\infty} B_0\, e^{i\pi w^2} dw. \qquad (54)$$

There are no D's in this expression, so the size and shape don't depend on D. But the size does depend on λ, according to $\sqrt{\lambda}$ (which decreases as λ decreases, as we argued near the beginning of this subsection). However, the shape doesn't depend on λ, because λ appears only in an overall constant.

3. An interesting fact about the Cornu spiral described by Eq. (52) and shown in Fig. 45 is that the curvature at a given point is proportional to the arclength traversed (starting from the

lower left end) to that point. The curvature is defined to be $1/R$, where R is the radius of the circle that matches up with the curve at the given point.

This property makes the Cornu spiral very useful as a transition curve in highways and railways. If you're driving down a highway and you exit onto an exit ramp that is shaped like the arc of a circle, then you'll be in for an uncomfortable jolt. Even though it seems like the transition should be a smooth one (assuming that the tangent to the circle matches up with the straight road), it isn't. When you hit the circular arc, your transverse acceleration changes abruptly from zero to v^2/R, where R is the radius of the circle. You therefore have to suddenly arrange for a sideways force to act on you (perhaps by pushing on the wall of the car) to keep you in the same position with respect to the car. Consistent with this, you will have to suddenly twist the steering wheel to immediately put it in a rotated position. It would be much more desirable to have the curvature change in a gradual manner, ideally at a constant rate. This way you can gradually apply a larger sideways force, and you can gradually turn the steering wheel. No sudden movements are required. The task of Problem 9.2 is to show that the Cornu spiral does indeed have the property that the curvature is proportional to the arclength. ♣

9.5.4 Diffraction around an object

The Cornu spiral gives the key to explaining the diffraction of light around an object. If we shine light on an object and look at the shadow, something interesting happens near the boundary. Fig. 48 shows the shadow of a razor blade illuminated by laser light.[9] Fig. 49 shows the result of a more idealized setup with an essentially infinite straight edge (oriented vertically on the page).[10]

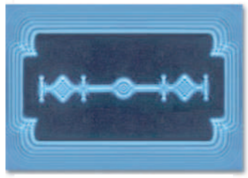


Figure 48

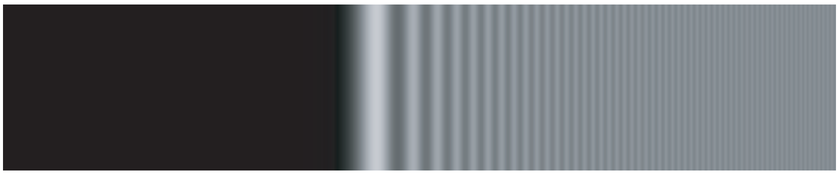


Figure 49

In a normal shadow, we would naively expect to have an abrupt change from a bright region to a dark region. Indeed, if instead of a light wave we had particles (such as baseballs) passing by a wall, then the boundary between the "shadow" and the region containing baseballs would be sharp. Now, even if we realize that light is a wave and can therefore experience interference/diffraction, we might still semi-naively expect to have the same kind of behavior on each side of the boundary, whatever that behavior might be. However, from Fig. 49 we see that there is something fundamentally different between the bright and dark regions. The amplitude oscillates in the bright region, but it appears to (and indeed does) decrease monotonically in the dark region. What causes this difference? We can answer this by looking at the Cornu spiral.

If we scan our eye across Fig. 49, this is equivalent to changing the location of point P in Fig. 35. Points far to the left (right) in Fig. 49 correspond to P being low (high) in Fig. 35. So as we scan our eye from left to right in Fig. 49, this corresponds to P being raised up from a large negative value to a large positive value in Fig. 35. However, as we noted in the second remark at the end of Section 9.5.2, raising the location of point P is equivalent

[9] I'm not sure where this picture originated.

[10] This image comes from the very interesting webpage, http://spiff.rit.edu/richmond/occult/bessel/bessel.html, which discusses diffraction as applied to lunar occultation.

to keeping P fixed and instead lowering the top of the wall.[11] Therefore, scanning our eye from left to right in Fig. 49 corresponds to lowering the top of the wall from a large positive value to a large negative value. And we can effectively take these values to be $\pm\infty$.

So to determine the intensity of the diffraction pattern as a function of position, we simply need to determine the intensity at P as we lower the wall. In turn, this means that we need to look at the length of the appropriate line in the Cornu spiral (and then square it to go from amplitude to intensity). The line we're concerned with always has one end located at the center of the upper-right spiral in Fig. 42, because in our setup the upper end of the "slit" is always located at $+\infty$. The other end of the line corresponds to the bottom of the slit, and since we're lowering this position down from $+\infty$, this other end simply starts at the center of the upper-right spiral and then winds its way outward in the spiral. When the top of the wall has moved all the way down to $y = 0$ (that is, across from P), the corresponding point on the spiral is as usual the center point between the two halves of the spiral. And when the top of the wall has moved all the way down to $-\infty$, the corresponding point on the spiral is the center of the lower-left spiral.

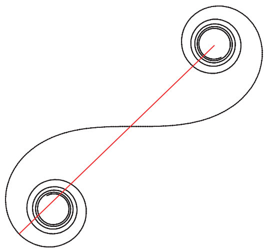


Figure 50

What happens to the amplitude (the length of the line) as we march through this entire process? It start out at zero when the top of the wall is at $+\infty$, and then it *monotonically* increases as we spiral outward in the upper-right spiral. It keeps increasing as we pass through the origin, but then it reaches its maximum possible value, shown in Fig. 50. (This spiral has $D = \lambda$, which undoubtedly isn't the case with light. But the shape of the $D \gg \lambda$ spiral isn't much different from the $D = \lambda$ one.) After this point, the length of the line oscillates up and down as we spiral inward in the lower-left spiral. The size of the oscillations gradually decreases as the circles get smaller and smaller, and the line approaches the one shown in Fig. 42, where the ends are at the centers of the two spirals. This corresponds to the top of the wall being at $y = -\infty$, so there is no wall at all.

The length of the amplitude line at the origin (which corresponds to P being at the edge of the location of the naive sharp shadow) is exactly half the length that it eventually settles down to. Since the intensity is proportional to the square of the amplitude, this means that the intensity at the naive edge is $1/4$ of the intensity far away from the shadow. Numerically, the maximum amplitude associated with Fig. 50 is about 1.18 times the amplitude far away, which means that the intensity is about 1.39 times as large.

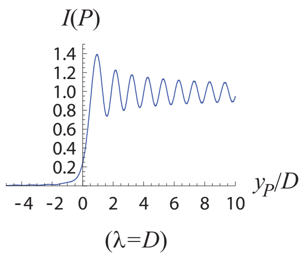


Figure 51

Note that although the two half-spirals in Fig. 50 have the same shape, one of them (the lower-left spiral) produces oscillations in the amplitude, while the other doesn't. The symmetry is broken due to where the starting point of the line is located. It is always located at the center of the upper-right spiral, and this is what causes the different behaviors inside and outside the shadow in Fig. 49.

The plot of the intensity (proportional to the amplitude squared) is shown in Fig. 51, with arbitrary units on the vertical axis. The horizontal axis gives the y coordinate of P, with $y = 0$ being across from the top of the wall. The left part corresponds to P being low in Fig. 35 (or equivalently, keeping P fixed and having the wall be high). In other words, P is in the left part of Fig. 49, in the shadow. The right part corresponds to P being high (or equivalently, keeping P fixed and having the wall be low). So P is in the left part of Fig. 49, outside the shadow. Moving from left to right in Fig. 51 corresponds to moving from left to right in Fig. 49, and also to running around the spiral in the direction we discussed above, starting at the inside of the upper-right spiral. As we mentioned above, you can see in Fig. 51 that the intensity at $y_P = 0$ is $1/4$ of the intensity at large y_P.

In the $D \gg \lambda$ limit (which is generally applicable to any setup involving light), the locations of the bright lines in the diffraction pattern are given by essentially the same reasoning that led to Eq. (53). So we essentially have $y \approx \sqrt{2nD\lambda}$. The $\sqrt{n}$ dependence

[11]With an infinite straight edge, we do indeed have the situation in Fig. 35 with a "half wall." The case of the razor blade is more complicated because it has holes and corners, but the general idea is the same.

implies that the bright lines get closer together as P moves farther away from the shadow (see Fig. 47). This is what we observe in Fig. 49. Note that the *angles* at which he bright lines occur are given by (assuming the angle is small) $\theta \approx y/D = \sqrt{2nD\lambda}/D = \sqrt{2n\lambda/D}$. So although the y values increase with D, the angles decrease with D.

9.5.5 Far-field limit

The expression for the wave in Eq. (48) is an exact one. It holds for arbitrary values of D and k (or equivalently λ), and also for arbitrary values of the limits of integration associated with the endpoints of the slit. Therefore, Eq. (48) and all conclusions drawn from the associated Cornu spiral hold for any setup. There is no need to actually be in the near-field regime; the results hold just as well in the far-field limit. So technically, the title of Section 9.5 should more appropriately be called "Anything-field diffraction" instead of "Near-field diffraction." We should therefore be able to obtain the far-field result as a limiting case of the "near-field" result. Let's see how this comes about.

For concreteness, let the distance to the screen be $D = 100$, and let the width of the slit be $a = 5$. Then $D \gg a$ is a fairly good approximation, so we should be able to (approximately) extract the far-field results from the Cornu spiral. Let's pick the wavelength to be $\lambda = 1$. Fig. 52 shows three possible locations of the slit. The reason for the particular bounds on the highest of these slits will be made clear shortly.

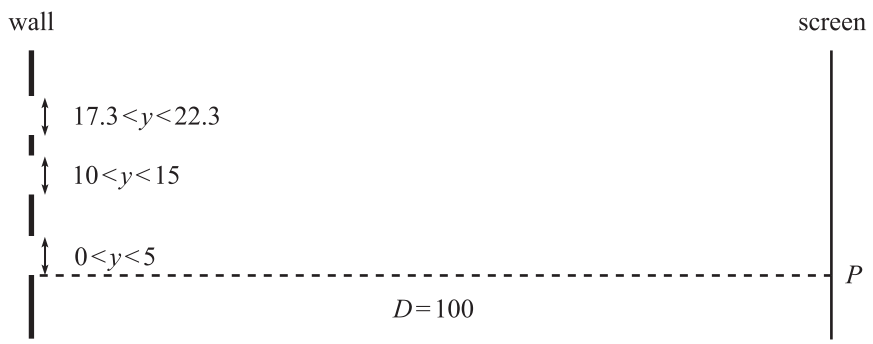


Figure 52

We can geometrically find the amplitudes at point P due to these three slits in the following way. The first spiral in Fig. 53 shows the relevant part of the spiral (the thick part) for the $0 < y < 5$ slit, along with the resulting amplitude (the straight line). The phases from the different points in the slit are roughly equal (because all of the pathlengths are roughly the same), so the wavelets add generally constructively (they mostly point to the right), and we end up with a decent-sized amplitude.

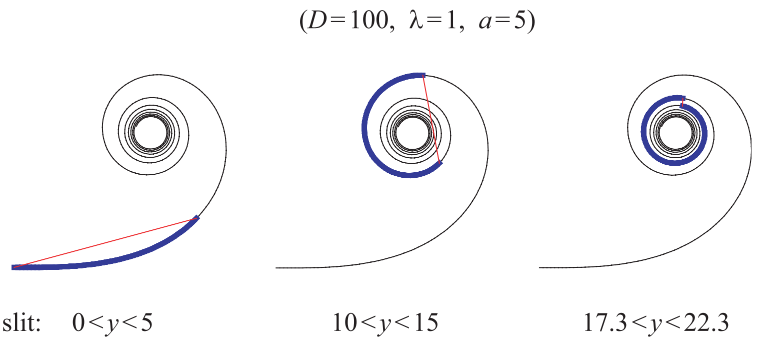


Figure 53

The second spiral shows the situation for the $10 < y < 15$ slit. The phases now differ by a larger amount, so the relevant part of the spiral curls around more, and resulting amplitude isn't as large. As the slit is raised, eventually we get to a point where the relevant part of the spiral forms a complete "circle." (It's not an actual circle, of course, because it doesn't close on itself, but it's close.) The resulting amplitude is then very small. This corresponds to the first zero in the diffraction pattern back in Fig. 23. The reason why the amplitude isn't exactly zero (as it was in the far-field result) is that D/a is only 20 here. This is fairly large, but not large enough to make the far-field approximation a highly accurate one. But remember that the present result is the correct one. The far-field result is an approximation.

If we choose a smaller slit width a, then the relevant part of the spiral (the thick part in Fig. 53) is shorter. It therefore needs to march deeper into the spiral to get to the point where it forms a complete circle (because the circles keep shrinking). Since the circles get closer together as they shrink (eventually they blend together in the figure to form a black blob), the small sideways shift that represents the amplitude in the third spiral in Fig. 53 is very tiny if the circle is deep in the spiral. So it's a better approximation to say that the amplitude there is zero. And consistent with this, the far-field approximation is a better one, because D/a is larger now. Basically, in the far-field limit, the length of the thick section in the spiral is much smaller than the general length scale of the spiral.

Note, however, that since the Cornu spiral never crosses itself, it is impossible to ever get an exactly complete cancelation of the wavelets and thereby a zero amplitude. There will always be a nonzero sideways shift between the two endpoints of the "circle." The zeros in the far-field limit in Fig. 23 are therefore just approximations (but good approximations if $D \gg a$).

Returning to the above case with $a = 5$, let's check that the numbers work out. In the third spiral in Fig. 53, having a complete circle means that the wavelets from the two ends of the slit have the same phase (because they have the same slope in the spiral). So the pathlengths from the two ends differ by one wavelength. (This is consistent with the reasoning in the second bullet point near the beginning of Section 9.3.2.) And indeed, since the slit runs from $y = 17.3$ to $y = 22.3$, and since $D = 100$, the pathlength difference is

$$\sqrt{100^2 + 22.3^2} - \sqrt{100^2 + 17.3^2} = 0.95. \tag{55}$$

This isn't exactly equal to one wavelength (which we chose to be $\lambda = 1$), but it's close. A larger value of D/a would make the difference be closer to one wavelength.

Note that the angle at which the point P is off to the side from the middle of the $17.3 < y < 22.3$ slit (which is located at $y = 19.8$) is given by $\tan\theta = 19.8/100 \Longrightarrow \theta = 11.2^\circ = 0.195$ rad. In the far-field approximation where the paths are essentially parallel, the difference in pathlengths from the ends of the slit is $a\sin\theta = (5)(\sin 11.2^\circ) = 0.97$, which is approximately one wavelength, as it should be.

What if we keep spiraling down into the spiral beyond the position shown in the third case in Fig. 53? This corresponds to raising the slit (while still keeping the width at $a = 5$). Eventually we'll get to a point where the circles are half as big, so the relevant part of the curve (the thick part) will wrap twice around a circle. This corresponds to the second zero in Fig. 23. The difference in the pathlengths from the ends of the slit is now (approximately) 2λ. If we keep spiraling in, the next zero occurs when we wrap three times around a circle. And so on.

However, we should be careful with this "and so on" statement. In the present case with $a = 5$ and $\lambda = 1$, it turns out that the part of the curve corresponding to the slit can wrap around a circle at most 5 times. (And the 5th time actually occurs only in the limit where the slit is infinitely far up along the wall.) This follows from the fact that since $\lambda = 1$, even if the slit is located at $y = \infty$, the pathlength from the far end of the slit is only $a = 5$ longer than the pathlength from the near end. So the phase difference can be at most 5 cycles. In other words, the thick part of the curve can't wrap more than 5 times around in a circle. Without using this physical reasoning, this limit of 5 circles isn't obvious by just looking at the spiral. The circles get smaller and smaller, so you might think that the wrapping number can be arbitrarily large. However, the little vectors corresponding to a given span dy are also getting smaller (because the amplitude is small if the slit is far away), which means that the thick part of the curve gets shorter and shorter. From simply looking at the curve, it isn't obvious which effect wins.

9.6 Problems

9.1. **Non-normal incidence** $*$

A light wave impinges on an N-slit setup at a small angle γ with respect to the normal. Show that for small angles, the interference pattern on a far-away screen has the same form as in Fig. 12, except that the entire plot is shifted by an angle γ. In other words, it's the same interference pattern, but now centered around the direction pointing along a ray of light (or whatever) that passes through the slit region.

9.2. **Cornu curvature** $**$

We stated in the last remark in Section 9.5.3 that the Cornu spiral has the property that the curvature at a given point is proportional to the arclength traversed (starting at the origin) to that point. Prove this. *Hint:* Write down the x and y coordinates associated with Eq. (51), and then find the "velocity" and "acceleration" vectors with respect to $u \equiv z_{\text{max}}$, and then use $a = v^2/R$.

9.7 Solutions

9.1. **Non-normal incidence**

Fig. 54 shows how to obtain the distances from a given wavefront (the left one in the figure) to a distance screen. We see that the lower path is longer than the upper path by an amount $d\sin\theta$, but also shorter by an amount $d\sin\gamma$. So the difference in pathlengths is $d(\sin\theta - \sin\gamma)$. In the derivation in Section 9.2.1 for the $\gamma = 0$ case, the difference in pathlengths was $d\sin\theta$. So the only modification we need to make in the $\gamma \neq 0$ case is the replacement of $d\sin\theta$ in Eq. (11) (and all subsequent equations) with $d(\sin\theta - \sin\gamma)$. So Eqs. (14) and (15) become

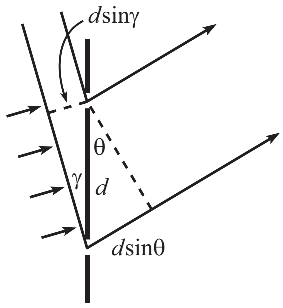


Figure 54

$$A_{\text{tot}}(\theta) = A(\theta)\frac{\sin\left(\frac{1}{2}Nkd(\sin\theta - \sin\gamma)\right)}{\sin\left(\frac{1}{2}kd(\sin\theta - \sin\gamma)\right)} \equiv A(\theta)\frac{\sin(N\alpha/2)}{\sin(\alpha/2)}\,, \tag{56}$$

where

$$\alpha \equiv kd(\sin\theta - \sin\gamma) = \frac{2\pi d(\sin\theta - \sin\gamma)}{\lambda}\,. \tag{57}$$

As before, α is the phase difference between adjacent paths.

For small angles, we can use $\sin\epsilon \approx \epsilon$ to write these results as

$$A_{\text{tot}}(\theta) = A(\theta)\frac{\sin\left(\frac{1}{2}Nkd(\theta - \gamma)\right)}{\sin\left(\frac{1}{2}kd(\theta - \gamma)\right)} \equiv A(\theta)\frac{\sin(N\alpha/2)}{\sin(\alpha/2)}\,, \tag{58}$$

where

$$\alpha \equiv kd(\theta - \gamma) = \frac{2\pi d(\theta - \gamma)}{\lambda}\,. \tag{59}$$

The only difference between this result and the original $\gamma = 0$ result (for small θ) is that the argument is $\theta - \gamma$ instead of θ. So the whole interference pattern is translated by an angle γ. That is, it is centered around $\theta = \gamma$ instead of $\theta = 0$, as we wanted to show.

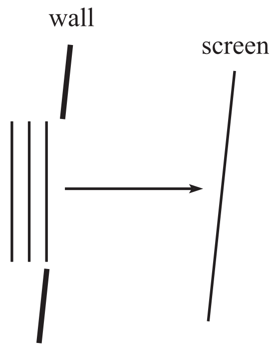


Figure 55

Remark: The same result holds for the diffraction pattern from a wide slit, because this is simply the limit of an N-slit setup, with $N \to \infty$. But Fig. 55 gives another quick way of seeing why the diffraction pattern is centered around the direction of the incident light. Imagine tilting the setup so that the angle of the incident light is horizontal (so the wavefronts are vertical). Then the wall and the screen are tilted. But these tilts are irrelevant (for small angles) because when we use Huygens principle near the slit, the little wavelets are created simultaneously from points on the *wavefronts*, and not in the *slit*. So the setup shown in Fig. 55 is equivalent to having the slit be vertical and located where the rightmost wavefront is at this instant. (Technically, the width of this vertical slit would be smaller by a factor of $\cos\gamma$, but $\cos\gamma \approx 1$ for small γ.) And the tilt of the screen is irrelevant for small angles, because any distances along the screen are modified by at most a factor of $\cos\gamma$. ♣

9.2. **Cornu curvature**

Writing the exponential in Eq. (51) in terms of trig functions tells us that the x and y coordinates of the points on the spiral in the complex plane are given by (with $a \equiv \pi D/\lambda$, and ignoring the factor of $B_0\sqrt{D}$)

$$x(u) = \int_0^u \cos(az^2)\,dz, \qquad \text{and} \qquad y(u) = \int_0^u \sin(az^2)\,dz. \tag{60}$$

The "velocity" vector with respect to u is given by $(dx/du, dy/du)$. But by the fundamental theorem of calculus, these derivatives are the values of the integrands evaluated at u. So we have (up to an overall factor of $B_0\sqrt{D}$)

$$\left(\frac{dx}{du}, \frac{dy}{du}\right) = \left(\cos(au^2), \sin(au^2)\right). \tag{61}$$

The magnitude of this velocity vector is $\cos^2(au^2) + \sin^2(au^2) = 1$. So the speed is constant, independent of the value of u. The total arclength from the origin is therefore simply u.

Since u is the upper limit on the z integral, and since z is proportional to the position y in the slit (from $z \equiv y/D$), we've just shown that if the upper end of the slit is moved up at a constant rate (the bottom end is held fixed at $y = 0$), then the corresponding point on the Cornu spiral moves along the spiral at a constant rate. If you want, you can think of u as the time that an object with constant speed has been moving.

The acceleration vector is the derivative of the velocity vector, which gives

$$\left(\frac{d^2x}{du^2}, \frac{d^2y}{du^2}\right) = \big(-2au\sin(au^2),\, 2au\cos(au^2)\big). \tag{62}$$

The magnitude of this vector is $2au$.

Now, the acceleration, speed, and radius of curvature are related by the usual expression, $a = v^2/R$. So we have $R = v^2/a$, which gives $R = (1)^2/(2au)$. The curvature is then $1/R = 2au$. But u is the arclength, so we arrive at the desired result that the curvature is proportional to the arclength. Note that since $a \propto 1/\lambda$, we have $R \propto \lambda$. So a small value of λ yields a tightly wound (and hence small) spiral. This is consistent with the result in the second remark in Section 9.5.3.