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.

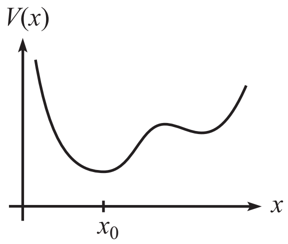


Figure 1

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

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

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

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

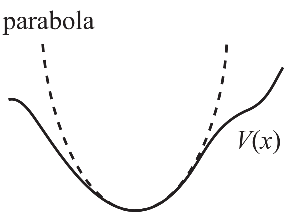


Figure 2

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

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

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

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

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

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

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 $T = 0.1\,\mathrm{s}$, then $\nu = 1/T = 10\,\mathrm{s}^{-1}$, 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.

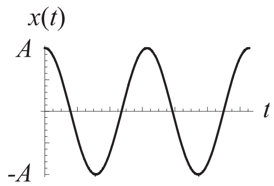


Figure 3

- 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*.

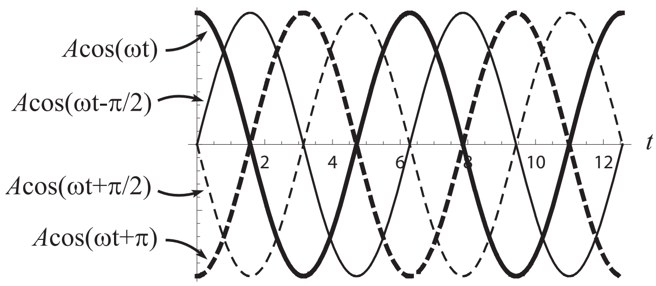


Figure 4

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.

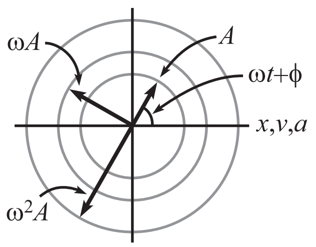
REMARK: The original general solution for $x(t)$ in Eq. (31) contains *four* parameters, namely the real and imaginary parts of C_1, and likewise for C_2 (ω is determined by k and m). Or equivalently, the four parameters are the magnitude and phase of C_1, and likewise for C_2. These four parameters 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

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



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).

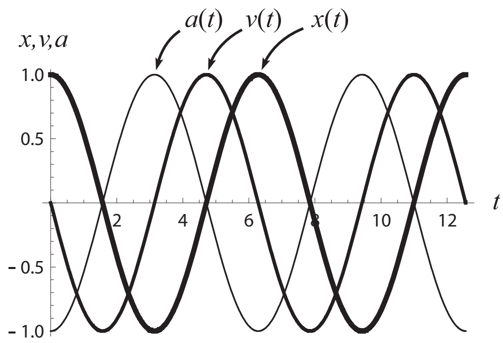


Figure 6

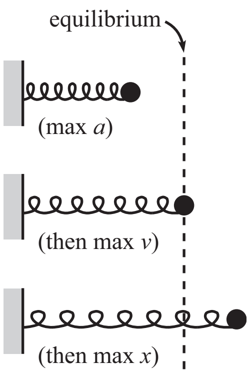


Figure 7

$a(t)$ 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 $x_{\rm i}$ and the final position $x_{\rm f}$. You can quickly show that that work is $\int(-kx)\,dx = kx_{\rm i}^2/2 - kx_{\rm f}^2/2$. 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.

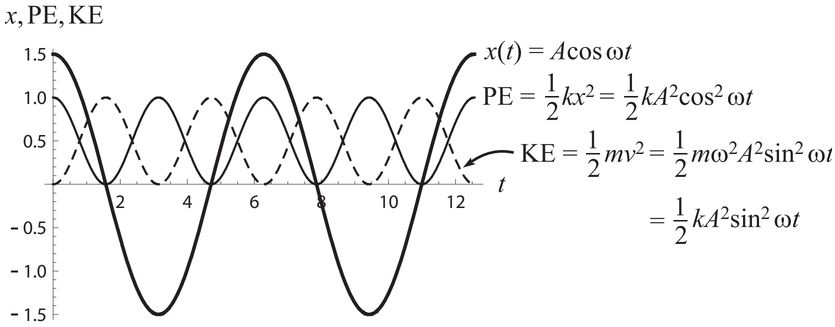


Figure 8

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 $\omega = \sqrt{\text{something}}$.

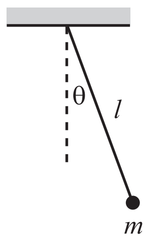


Figure 9

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

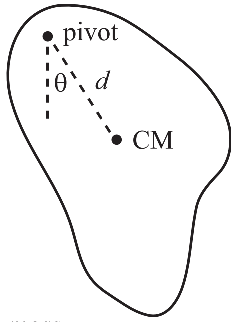
$$\theta(t) = A\cos(\omega t + \phi), \tag{42}$$

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.

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.



mass m,
moment of inertia I

Figure 10

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.

LC circuit

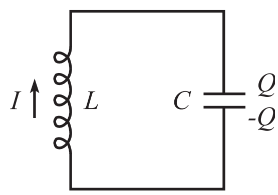


Figure 11

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 $\mu_k N$. 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

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

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

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

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

$$\omega_u \equiv \frac{1}{2}\sqrt{4\omega_0^2 - \gamma^2} \quad \implies \quad \boxed{\omega_u = \omega_0\sqrt{1 - \left(\frac{\gamma}{2\omega_0}\right)^2}} \tag{50}$$

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

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

[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

$$x(t) = e^{-\gamma t/2}\Big(C_1 e^{i\omega_{\mathrm{u}} t} + C_2 e^{-i\omega_{\mathrm{u}} t}\Big). \tag{52}$$

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

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

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 $\omega_0 = 1\,\mathrm{s}^{-1}$, with the initial conditions of $x(0) = 1$ and $v(0) = 0$.

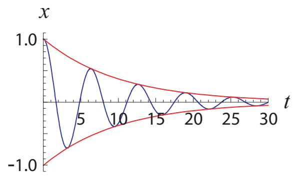


Figure 12

Note that the frequency $\omega_{\mathrm{u}} = \omega_0\sqrt{1-(\gamma/2\omega_0)^2}$ 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 $\omega_{\mathrm{u}} = (0.8)\omega_0$), then you can show that this implies that $\gamma = (1.2)\omega_0$. So after just two cycles (that is, when $\omega_{\mathrm{u}} t = 4\pi$), the damping factor equals

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

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

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

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

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

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

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

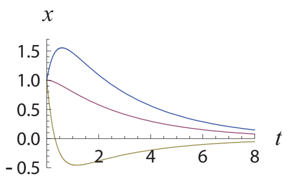


Figure 13

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

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

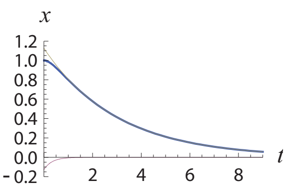


Figure 14

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 $1/\mu_1 \approx 1/\gamma = (1/3)\,$s, and the slow-decay term dies out on a time scale of $1/\mu_2 \approx \gamma/\omega_0^2 = 3\,$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

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

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

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

Therefore,

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

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

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

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

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

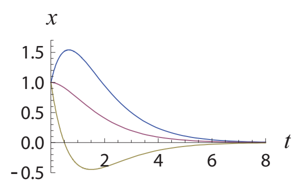


Figure 15

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 $F_\text{d}\cos\omega t$.

The second reason why we choose to consider a force of the form $\cos\omega t$ is that $F(t) = F_\text{d}\cos\omega t$ 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 $x_\text{end}(t) = A_\text{end}\cos\omega t$ (which often happens in real life), then the spring force is

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

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

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,

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

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

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

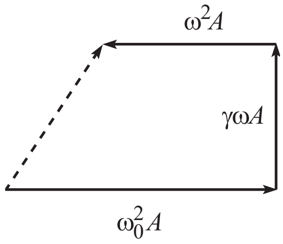


Figure 16

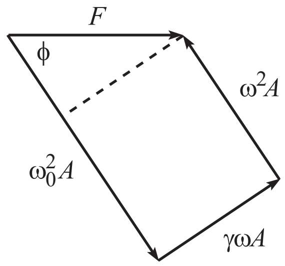


Figure 17

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

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

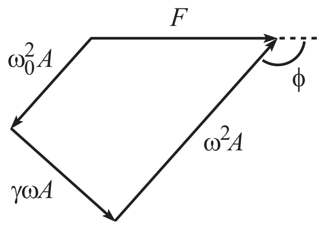


Figure 18

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.

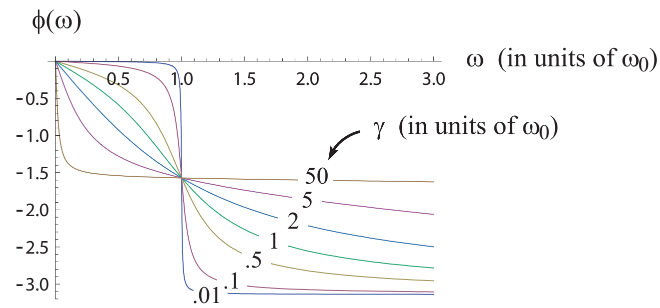


Figure 19

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.

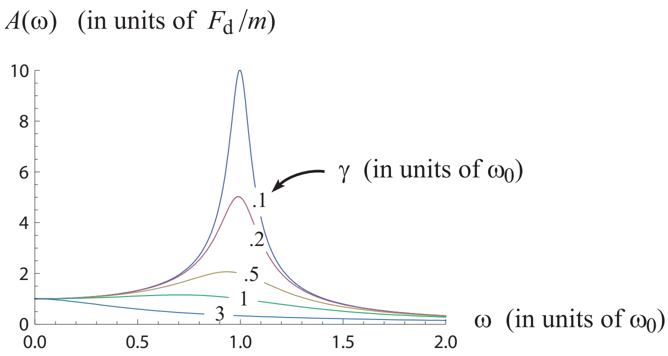


Figure 20

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

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

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,

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

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

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

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

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

we obtain

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

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,

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

Note that the quantity $\mathrm{Re}(Ce^{i\omega t})$ is what matters here, and *not* $\mathrm{Re}(C)\mathrm{Re}(e^{i\omega t})$. 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

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

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

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

  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

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

  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 $\omega_0 \equiv \sqrt{k/m}$, $\gamma \equiv b/m$, $F \equiv F_{\mathrm{d}}/m$, 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 $x = x_{\mathrm{p}} + x_{\mathrm{h}}$)

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

which means that $x = x_{\rm p} + x_{\rm h}$ 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, $x = x_{\rm p} + x_{\rm h}$, is

$$\boxed{x(t) = A_{\rm p}\cos(\omega t + \phi) + A_{\rm h}e^{-\gamma t/2}\cos(\omega_{\rm h}t + \theta)} \qquad (\textit{underdamped}) \tag{109}$$

A word on the various parameters in this result:

- ω is the driving frequency, which can be chosen arbitrarily.
- $A_{\rm p}$ and ϕ are functions of $\omega_0 \equiv \sqrt{k/m}$, $\gamma \equiv b/m$, $F \equiv F_{\rm d}/m$, and ω. They are given in Eqs. (88) and (87).
- $\omega_{\rm h}$ is a function of ω_0 and γ. It is given in Eq. (50).
- $A_{\rm h}$ 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

$$\phi \approx 0, \qquad \text{and} \qquad A \approx \frac{F}{\omega_0^2} \equiv \frac{F_{\rm d}/m}{k/m} = \frac{F_{\rm d}}{k}\,. \tag{110}$$

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

$$x(t) = A\cos(\omega t + \phi) = \frac{F_{\rm d}}{k}\cos\omega t. \tag{111}$$

Note that the spring force is then $F_{\rm spring} = -kx = -F_{\rm d}\cos\omega t$, 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 $\sin\phi = -\gamma\omega A/F \equiv -\gamma m\omega A/F_\mathrm{d}$. So Eq. (123) gives

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

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

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

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

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 $\langle P_\mathrm{driving}\rangle$ 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 $\langle P_\mathrm{driving}\rangle$ in units of $F_\mathrm{d}^2/2\gamma m \equiv F_\mathrm{d}^2/2b$.

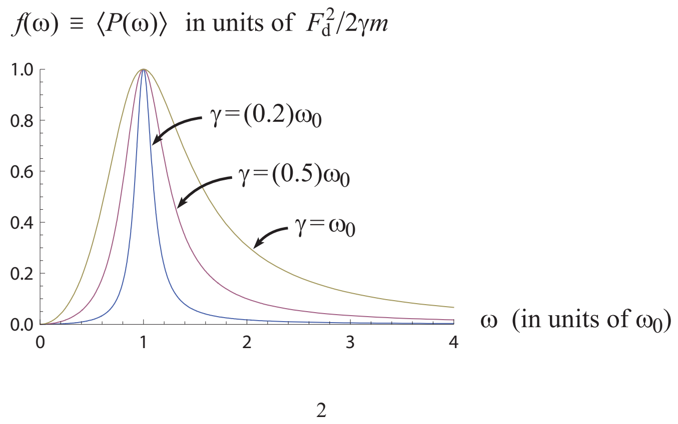


Figure 21

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

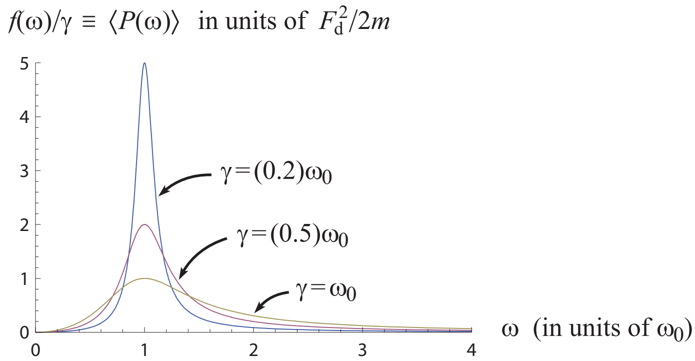


Figure 22

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

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

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

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

(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 $\gamma = 2\omega_0$.

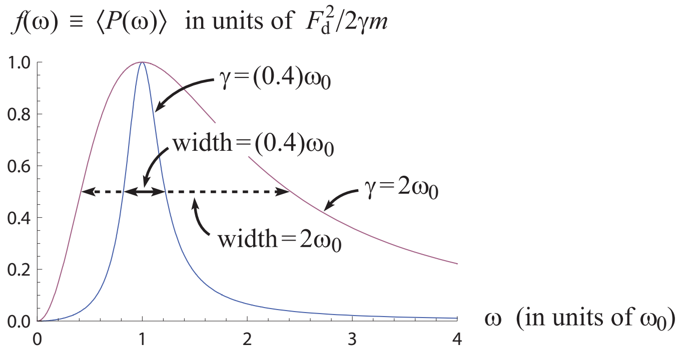


Figure 23

To sum up, the maximum height of the $\langle P_{\text{driving}}\rangle$ curve is proportional to $1/\gamma$ (it equals $F_{\rm d}^2/2\gamma m$), 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 $\omega = \omega_0$.

- $\omega \approx 0$: 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.

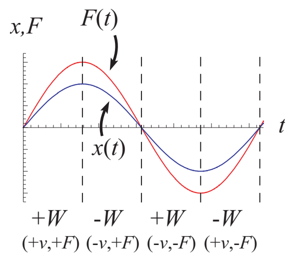


Figure 24

- $\omega \approx \infty$: 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.

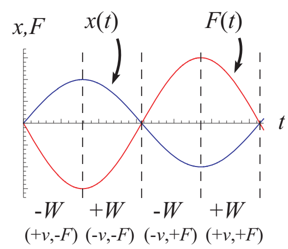


Figure 25

- $\omega = \omega_0$: 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.

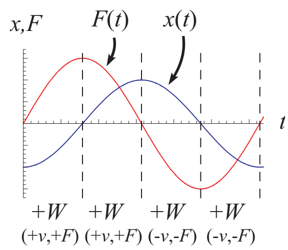


Figure 26

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]

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.

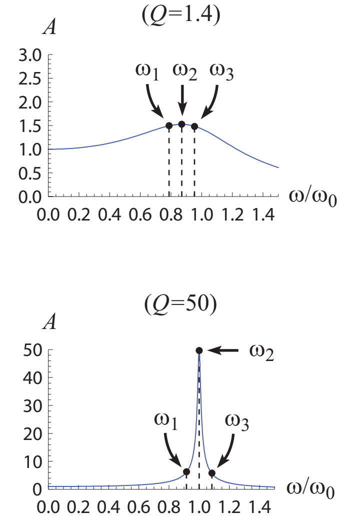


Figure 27

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 $A < F_\mathrm{d}/\gamma m\omega$ (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, $A > F_\mathrm{d}/\gamma m\omega$,

[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 $A = F_{\rm d}/\gamma m\omega$ 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 $F_{\rm d}$ 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.